Astronomy & Astrophysics

Dependence of cosmic shear covariances on cosmology

Impact on parameter estimation

T. Eifler[1], P. Schneider[1], and J. Hartlap[1]

Argelander-Institut für Astronomie, Universität Bonn, Auf dem Hügel 71, 53121 Bonn, Germany
e-mail: tim.eifler@astro.uni-bonn.de

ABSTRACT

Context. 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.

Aims. We examine whether the aforementioned assumptions hold and quantify how strong cosmic shear covariances vary within a reasonable parameter range. Furthermore, we examine the impact on likelihood contours when assuming different cosmologies in the covariance. The final goal is to develop an improved likelihood analysis for parameter estimation with cosmic shear.

Methods. We calculate Gaussian covariances analytically for 2500 different cosmologies. 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 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, and to minimize the error due to this approximation, we develop the concept of an iterative likelihood analysis.

Results. Covariances vary significantly within the considered parameter range. The cosmology assumed in the covariance has a non-negligible impact on the size of the likelihood contours. This impact increases with increasing survey size, increasing number density of source galaxies, decreasing ellipticity noise, and when taking non-Gaussianity into account. A proper treatment of this effect is therefore even more important for future surveys. In this paper, we present methods for taking cosmology-dependent covariances into account.

Key words. cosmology: large-scale structure of the Universe – methods: statistical – cosmology: theory – cosmology: cosmological parameters

1. Introduction

Cosmic shear, first detected in 2000 (Bacon et al. 2000; Kaiser et al. 2000; van Waerbeke et al. 2000; Wittman et al. 2000), has become an important tool in cosmology. Latest results (e.g., van Waerbeke et al. 2005; Semboloni et al. 2006; Hoekstra et al. 2006; Schrabback et al. 2007; Hetterscheidt et al. 2007; Massey et al. 2007; Fu et al. 2008) already indicate its great ability to constrain cosmological parameters, which will be enhanced in the future by large upcoming surveys like Pan-STARRS, KIDS, DES, Euclid or LSST. The improved quality of cosmic shear data must be accompanied by an accurate data analysis that is free of assumptions biasing the results. In this context, obtaining appropriate covariances is a crucial issue 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 was derived in Schneider et al. (2002a) and confirmed in Joachimi et al. (2008), who used a power spectrum approach that 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 become important at angular scales ≲10 arcmin. To account for non-Gaussianity, Semboloni et al. (2007) invented 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. 2007). 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 insufficiently large. Third, one can estimate the covariance matrix from ray-tracing simulations, a method that circumvents the aforementioned loss in information. Although, the covariance in this method is again derived by field-to-field variation, we can choose a sufficiently large numerical simulation to create many independent subsamples of adequate size.

We 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 co-variance matrices from ray-tracing simulations seems to be a promising method because it preserves all the information in the data and in addition 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. To date, cosmic shear likelihood analyses have treated the covariance matrix as a constant in parameter space, hence the parameter constraints are assumed to be unaffected by the underlying cosmology. In this paper, we intend to check this assumption and if we find that it does not hold, to present an improved likelihood formalism for future surveys.

This paper is organized as follows. Section 2 summarizes the basic theoretical background of the cosmic shear two-point correlation function (2PCF) and its corresponding covariance. In Sect. 3, we derive a scaling relation for covariances, which can be used for a fast calculation of covariances for arbitrary cosmology. Furthermore, we examine how strongly the covariance depends on its underlying cosmological model. The impact on parameter constraints when assuming a fixed cosmology in the covariance is the subject of Sect. 4, whereas we present improvements on this assumption in Sect. 5. Here, we consider a likelihood analysis with an adaptive covariance matrix, i.e., the covariance is calculated individually for each point in parameter space where the likelihood is evaluated. In addition, we outline the concept of an iterative covariance matrix, i.e., several likelihood analyses are performed, where the covariance is updated in every iteration according to the maximum likelihood parameter set. Here, we also examine the impact of non-Gaussian covariances on the likelihood contours using a ray-tracing covariance matrix derived from the Millennium Simulation. We present our conclusions in Sect. 6.

2. Data vectors and covariances of cosmic shear

We briefly review the basics of the cosmic shear two-point correlation function and its corresponding covariance matrix. For more details of this topic, the reader is referred to Bartelmann & Schneider (2001), Schneider et al. (2002a), Schneider et al. (2002b), Kilbinger & Schneider (2004), and Joachimi et al. (2008).

To measure the shear signal, we define θ as the connecting vector of two points and specify tangential and cross-component of the shear γ as

$$\gamma_t = -\mathrm{Re}\left(\gamma e^{-2i\varphi}\right) \qquad \text{and} \qquad \gamma_\times = -\mathrm{Im}\left(\gamma e^{-2i\varphi}\right), \qquad (1)$$

where φ is the polar angle of θ. The 2PCFs depend only on the absolute value of θ. They are defined in terms of the shear and can be related to the power spectra P_E and P_B (Schneider et al. 2002b)

$$\xi_\pm(\theta) \equiv \langle\gamma_t\gamma_t\rangle(\theta) \pm \langle\gamma_\times\gamma_\times\rangle(\theta) \qquad (2)$$

$$= \int_0^\infty \frac{\mathrm{d}\ell\,\ell}{2\pi} J_{0/4}(\ell\theta)\,\left[P_E(\ell) \pm P_B(\ell)\right], \qquad (3)$$

where J_n denotes the nth order Bessel-function. We only consider E-modes, therefore we define $P_B = 0$. Furthermore, we assume that the 2PCF is estimated in logarithmic bins ϑ of angular width $\Delta\vartheta$. The covariance of the 2PCF is defined as

$$C_\xi\left(\vartheta_i, \vartheta_j\right) := \left\langle\left(\xi_\pm(\vartheta_i) - \hat{\xi}_\pm(\vartheta_i)\right)\left(\xi_\pm(\vartheta_j) - \hat{\xi}_\pm(\vartheta_j)\right)\right\rangle. \qquad (4)$$

For the rest of this paper, we neglect the index ξ in the covariance because we only consider covariances of the 2PCF. As one

can already see from Eq. (4), the 2PCF has four different covariances, denoted by C_{++}, C_{+-}, C_{-+}, C_{--}. Only three of them are independent because $C_{+-}(\vartheta_i, \vartheta_j) = C_{-+}(\vartheta_j, \vartheta_i)$. Assuming a Gaussian shear field, the covariance of the 2PCF can be calculated analytically (Schneider et al. 2002a; Joachimi et al. 2008). There, the covariance is decomposed into three terms, namely the cosmic variance term (V), the pure shot noise term (S), and the mixed term (M)

$$C_{++}(\vartheta_i, \vartheta_j) = V_{++} + M_{++} + S, \qquad (5)$$

$$C_{--}(\vartheta_i, \vartheta_j) = V_{--} + M_{--} + S, \qquad (6)$$

$$C_{+-}(\vartheta_i, \vartheta_j) = V_{+-} + M_{+-}. \qquad (7)$$

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

$$S = \frac{\sigma_\epsilon^4}{2\pi\vartheta_i\Delta\vartheta_i A\bar{n}^2}\,\delta_{\vartheta_i\vartheta_j}, \qquad (8)$$

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

$$V_{\pm\pm} = \frac{1}{\pi A}\int_0^\infty \mathrm{d}\ell\,\ell\,J_{0/4}(\ell\vartheta_i)\,J_{0/4}(\ell\vartheta_j)\,P_E^2(\ell), \qquad (9)$$

$$M_{\pm\pm} = \frac{\sigma_\epsilon^2}{\pi A\bar{n}}\int_0^\infty \mathrm{d}\ell\,\ell\,J_{0/4}(\ell\vartheta_i)\,J_{0/4}(\ell\vartheta_j)\,P_E(\ell). \qquad (10)$$

The corresponding expressions for V and M using the 2PCF are derived in Schneider et al. (2002a). In this paper we only need the expressions for the mixed term, which read

$$M_{++} = \frac{2\sigma_\epsilon^2}{\pi A\bar{n}}\int_0^\pi \mathrm{d}\varphi\,\xi_+(|\phi|), \qquad (11)$$

$$M_{--} = \frac{2\sigma_\epsilon^2}{\pi A\bar{n}}\int_0^\pi \mathrm{d}\varphi\,\xi_+(|\phi|)\,\cos(4\varphi), \qquad (12)$$

$$M_{+-} = \frac{2\sigma_\epsilon^2}{\pi A\bar{n}}\int_0^\pi \mathrm{d}\varphi\left[\sum_{k=0}^4 \binom{4}{k}(-1)^k\vartheta_i^k\,\vartheta_j^{4-k}\cos(k\varphi)\right]$$
$$\times (|\phi|)^{-4}\,\xi_-(|\phi|)\,\cos(4\varphi), \qquad (13)$$

where we denote $|\phi| = \sqrt{\vartheta_i^2 + \vartheta_j^2 - 2\vartheta_i\vartheta_j\cos\varphi}$.

3. Variations in covariances with respect to cosmology

We select a two-dimensional parameter grid with 50×50 grid-points 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 Eqs. (5)−(10). The shear power spectra P_E are obtained from the density power spectra P_δ by 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 paper, 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 (π_0), which we have chosen to be similar to the cosmology of the Millennium Simulation
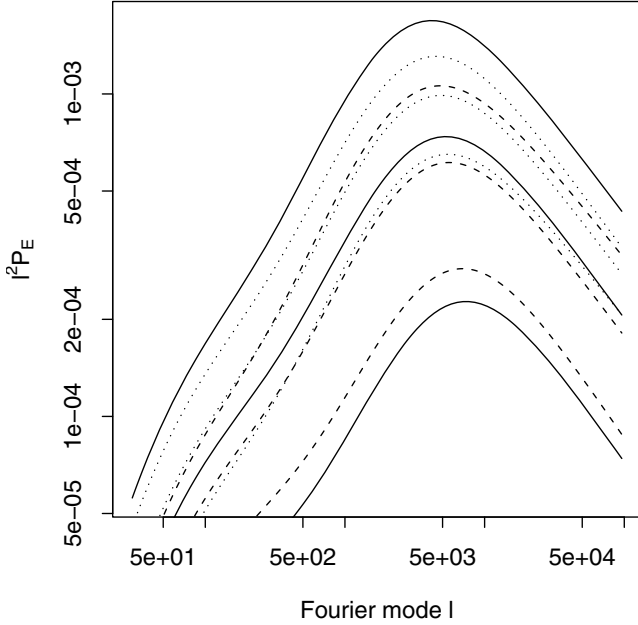
Fig. 1. The dimensionless shear power spectrum $\ell^2 P_E$. The solid curves correspond to variation in Ω_m and σ_8: $\Omega_m = 0.2$, $\sigma_8 = 0.6$ (*lower*), $\Omega_m = 0.3$, $\sigma_8 = 0.8$ (*middle*), $\Omega_m = 0.4$, $\sigma_8 = 1.0$ (*top*). The dashed curves show variation in σ_8 with $\Omega_m = 0.25$: $\sigma_8 = 0.6$ (*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*).

(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. 3 are generally valid and independent of survey parameters. For the likelihood analyses in Sects. 4 and 5, we choose, unless stated otherwise, an intrinsic ellipticity noise of $\sigma_\epsilon = \sqrt{\langle \epsilon^{(s)} \epsilon^{(s)} \rangle} = 0.4$, 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 that 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 50 logarithmic bins. The results of this section are limited by the accuracy of the non-linear fitting formula of Smith et al. (2003). It will be the subject of future work to refine the results presented here using an improved version of this fit-formula.

3.1. A fast method for calculating covariances of arbitrary Ω_m and σ_8

From Eqs. (9) and (10), one can directly see that the covariance matrix depends on the cosmological model, which depends in turn on the power spectrum P_E. Figure 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 Eqs. (5)−(10). 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 π to a reference model π_0 (π_0 being the fiducial cosmology described above). A basic theorem in statistics states (e.g., Anderson 2003) that if there is a relation between

two data vectors x and y such as $y = Ax$ (A being a matrix), the relation between the covariances of x and y can be written as

$$
\begin{aligned}
\mathbf{C}_y &= \left\langle (y - \langle y \rangle)(y - \langle y \rangle)^t \right\rangle \\
&= \left\langle (Ax - \langle Ax \rangle)(Ax - \langle Ax \rangle)^t \right\rangle \\
&= A\mathbf{C}_x A^t.
\end{aligned}
\tag{14}
$$

In this derivation, A must be independent of the ensemble average. If we apply the above ansatz to the 2PCF (where we denote a 2PCF calculated for a cosmology π as ξ_π), it seems reasonable to define a scaling relation for parameter-dependent covariances as

$$
\mathbf{C}_{\xi_\pi} = A\mathbf{C}_{\xi_{\pi_0}} A^t,
\tag{15}
$$

where we can calculate the scaling matrices A using the 2PCF

$$
\text{diag}(A) = \xi_\pi / \xi_{\pi_0}.
\tag{16}
$$

In contrast to a covariance matrix, the 2PCF can be calculated extremely rapidly for many different cosmologies by means of Eq. (3). Hence, it would be a fast and convenient method to calculate the covariance for a reference cosmology and then apply Eq. (15) to obtain covariances for arbitrary cosmological parameters. Unfortunately, we cannot transfer this method directly to the cosmic shear case. We recall that the 2PCF is derived from the measured ellipticities of galaxies. Schneider et al. (2002b) demonstrate 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 not only consists of terms originating from the shear, but has additional noise terms that originate from the intrinsic ellipticity of galaxies. The pure shot noise term in Eq. (8) is independent of cosmology and, as can be seen from Eqs. (11)−(13), the mixed term cannot be scaled with the relation in Eq. (15), 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 Eq. (15) exists. We prove this explicitly below by, in particular, showing that the scaling matrices are independent of the ensemble average. The cosmic variance term can be calculated via Eq. (9). Cosmology is important only in terms of the power spectrum, hence the relation between π and π_0 can be described as $P_E(\ell, \pi) = a(\ell, \pi)P_E(\ell, \pi_0)$. Using this relation, we transform the cosmic variance term in Eq. (9) for given bins ϑ_i, ϑ_j as follows

$$
\begin{aligned}
V_{\pm\pm}(\pi) &= \frac{1}{\pi A} \int_0^\infty \mathrm{d}\ell\, \ell\, J_{0/4}(\ell\vartheta_i)\, J_{0/4}(\ell\vartheta_j)\, P_E^2(\ell, \pi) \\
&= \frac{1}{\pi A} \sum_{\bar{\ell}} \Delta\bar{\ell}\, \bar{\ell}\, J_{0/4}(\bar{\ell}\vartheta_i)\, J_{0/4}(\bar{\ell}\vartheta_j)\, a^2(\bar{\ell}, \pi) P_E^2(\bar{\ell}, \pi_0),
\end{aligned}
\tag{17}
$$

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

$$
\left\langle \Delta P_E(\bar{\ell}) \Delta P_E(\bar{\ell}') \right\rangle = \frac{4\pi}{A\bar{\ell}\Delta\bar{\ell}}\, P_E^2(\bar{\ell})\, \delta_{\bar{\ell}\bar{\ell}'},
\tag{18}
$$

to rewrite Eq. (17) as

$$
\begin{aligned}
V_{\pm\pm}(\pi) = \frac{1}{4\pi^2} \sum_{\bar{\ell},\bar{\ell}'} \Delta\bar{\ell}^2 \bar{\ell}' \bar{\ell}\, J_{0/4}(\bar{\ell}\vartheta_i)\, J_{0/4}(\bar{\ell}'\vartheta_j)\, a(\bar{\ell}, \pi)\, a(\bar{\ell}', \pi) \\
\times \left\langle \Delta P_E(\bar{\ell}, \pi_0) \Delta P_E(\bar{\ell}', \pi_0) \right\rangle.
\end{aligned}
\tag{19}
$$

The mean value theorem guarantees that there exist values $\bar{a}(\vartheta_i, \boldsymbol{\pi})$, $\bar{a}(\vartheta_j, \boldsymbol{\pi})$ such that Eq. (19) becomes

$$
\begin{aligned}
V_{\pm\pm}(\boldsymbol{\pi}) &= \frac{\bar{a}(\vartheta_i, \boldsymbol{\pi})\, \bar{a}(\vartheta_j, \boldsymbol{\pi})}{4\pi^2} \left\langle \int_0^\infty \mathrm{d}\ell\, \ell\, J_{0/4}(\ell\vartheta_i)\, \Delta P_{\mathrm{E}}(\ell, \boldsymbol{\pi}_0) \right. \\
&\quad \left. \times \int_0^\infty \mathrm{d}\ell'\, \ell'\, J_{0/4}(\ell'\vartheta_j)\, \Delta P_{\mathrm{E}}(\ell', \boldsymbol{\pi}_0) \right\rangle \\
&= \bar{a}(\vartheta_i, \boldsymbol{\pi})\, \bar{a}(\vartheta_j, \boldsymbol{\pi})\, V_{\pm\pm}(\boldsymbol{\pi}_0),
\end{aligned}
\tag{20}
$$

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 to be

$$
\bar{a}_{\pm\pm}(\vartheta_i, \boldsymbol{\pi}) = \frac{\int_0^\infty \mathrm{d}\ell\, \ell\, J_{0/4}(\ell\vartheta_i) P_{\mathrm{E}}(\ell, \boldsymbol{\pi})}{\int_0^\infty \mathrm{d}\ell\, \ell\, J_{0/4}(\ell\vartheta_i) P_{\mathrm{E}}(\ell, \boldsymbol{\pi}_0)} = \frac{\xi(\vartheta_i, \boldsymbol{\pi})}{\xi(\vartheta_i, \boldsymbol{\pi}_0)},
\tag{21}
$$

where we inserted Eq. (3) in the last step. This provides a fast and convenient method for scaling the cosmic variance term in parameter space, because we can use a computationally efficient Hankel transformation to calculate the 2PCF.

From Eqs. (11)−(13) we see that the mixed term $M_{\pm\pm}$ scales linearly with the 2PCF, which prevents a scaling relation similar to Eq. (15). Fortunately, the direct calculation of the linear term with Eq. (10) is comparatively fast, and the scaling relation for the cosmic variance term therefore already reduces the computational costs significantly.

Nonetheless, we numerically derive a fit-formula for the linear term based on the following expression

$$
M_{\pm\pm}(\vartheta_i, \vartheta_j, \boldsymbol{\pi}) = M_{\pm\pm}(\vartheta_i, \vartheta_j, \boldsymbol{\pi}_0) \left(\frac{\Omega_{\mathrm{m}}}{0.25}\right)^\alpha \left(\frac{\sigma_8}{0.9}\right)^\beta.
\tag{22}
$$

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 differ for the different parts of the covariance matrix, $\mathbf{C}_{++}$, $\mathbf{C}_{--}$, and $\mathbf{C}_{+-}$. The tables with α and β are available on the internet[1]. The fit formula for the cosmic variance term in Eq. (20) and the linear term in Eq. (22) are valid for any set of survey parameters. This can be seen directly from Eqs. (9) and (10), respectively, where survey parameters enter as prefactors. In the next section, we consider the trace of the *inverse* covariance to illustrate the strength of the CDC-effect. We note that deriving a fit for the cosmology dependence of the inverse covariance is not very useful since this formula would change depending on σ_ϵ and $\bar{n}$. The reason for this is that the covariance is a sum of three terms (see Eqs. (5)−(7)) and only the survey size A enters similarly in all terms, implying that $\mathbf{C} \propto 1/A$ and $\mathbf{C}^{-1} \propto A$. However, for σ_ϵ and $\bar{n}$, the behavior of $\mathbf{C}^{-1}$ must be evaluated numerically.

3.2. Variations in the inverse covariance with respect to Ω_m and σ_8

From the variations in the power spectrum with Ω_{m} and σ_8 (Sect. 3.1), it is clear that covariances vary with respect to comological parameters. For simplicity and to improve the readability, we refer to this variation as the CDC-effect (CDC ≡ Cosmology

[1] http://www.astro.uni-bonn.de/~teifler/fit-parameters.pdf

Dependent Covariances). To examine the CDC-effect more closely, we recall that the structure of the covariance is given by

$$
\mathbf{C} = \begin{pmatrix} \mathbf{C}_{++} & \mathbf{C}_{+-} \\ \mathbf{C}_{+-}^{\mathrm{t}} & \mathbf{C}_{--} \end{pmatrix}
\tag{23}
$$

and the individual parts are calculated from Eqs. (5)−(10). From these equations, we can see that the covariances are filtered versions of the power spectrum, filtered either by a product of J_0's (in the 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, since 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. 1), therefore, the CDC-effect for the individual parts of $\mathbf{C}$ is also similar. However, this might change when considering different cosmological parameters, such as the shape parameter Γ. A change in Γ rotates the power spectrum. The covariances are integrals over P_{E}, and depending on the filter function, the change in P_{E} can average out. A second argument for why the individual covariance parts have different sensitivity to the CDC-effect is that $\mathbf{C}_{+-}$ is unaffected by shot noise, hence a change in cosmology has a stronger impact on $\mathbf{C}_{+-}$ compared to $\mathbf{C}_{++}$ and $\mathbf{C}_{--}$.

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 of this effect, as it depends on the binning, which can be seen from Eq. (8). The trace of $\mathbf{C}$ becomes arbitrarily large when the bin width decreases. 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. More precisely, once the bin width of the 2PCF data vector is small enough for discretization effects to be unimportant, the trace of $\mathbf{C}^{-1}$ hardly changes for different binning. For more details on how the binning affects covariances, their inverse, or parameter constraints derived from these covariances, the reader is referred to chapter 4 of Eifler (2009).

Figure 2 shows how the trace of the inverse covariance matrix depends on Ω_{m} for various constant values of σ_8 (top) and vice versa (bottom). Here, we normalize the survey size to $A = 1$ deg^2, and the other survey parameters are $\sigma_\epsilon = 0.4$ and $\bar{n} = 10/\text{arcmin}^2$. We postpone a detailed analysis of how survey parameters influence the CDC-effect to Sect. 4. 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 ith 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.
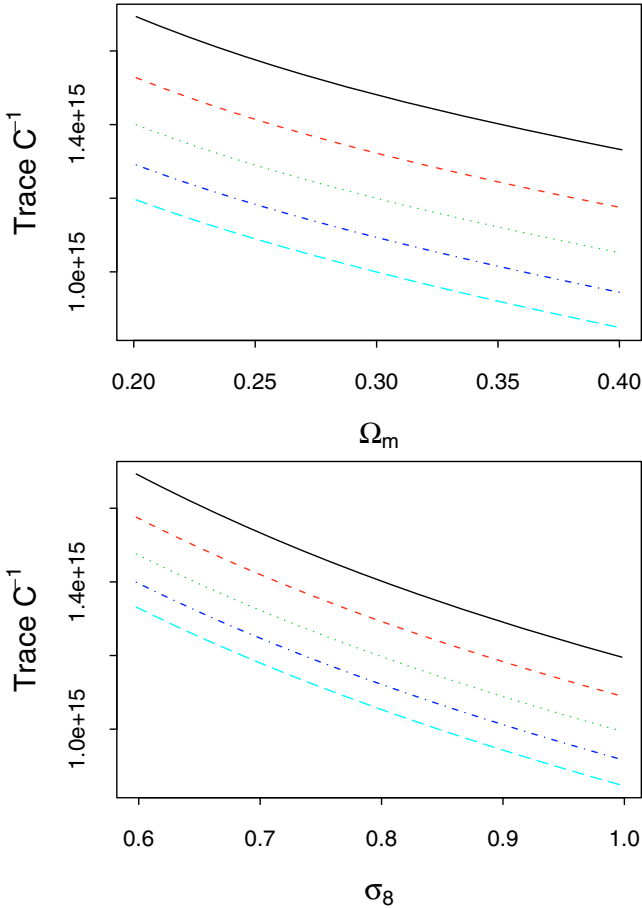
4. Impact of the CDC-effect on parameter estimation

4.1. Basics of the likelihood analysis

Throughout the entire likelihood analysis, we assume the ΛCDM model. We define the posterior likelihood $p(\boldsymbol{\pi}|\boldsymbol{\xi})$ for the case of a 2PCF data vector as

$$
p(\boldsymbol{\pi}|\boldsymbol{\xi}) = \frac{p(\boldsymbol{\xi}|\boldsymbol{\pi})}{p(\boldsymbol{\xi})}\, p(\boldsymbol{\pi}),
\tag{24}
$$

Fig. 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*) $\Omega_m = [0.2, 0.25, 0.3, 0.35, 0.4]$.

where $p(\boldsymbol{\pi})$ denotes the prior probability density, $p(\boldsymbol{\xi}|\boldsymbol{\pi})$ is the likelihood, and $p(\boldsymbol{\xi})$ denotes the evidence. The prior contains information on the parameter vector $\boldsymbol{\pi}$ which comes from former experiments. Here, we assume flat priors with cutoffs, which means that $p(\boldsymbol{\pi})$ 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 $p(\boldsymbol{\pi}) = 0$ elsewhere. All other parameters are fixed to those of our fiducial cosmology (see Sect. 3). We assume that $\boldsymbol{\xi}$ is normally distributed in parameter space, hence our likelihood $p(\boldsymbol{\xi}|\boldsymbol{\pi})$ can be written as

$$p(\boldsymbol{\xi}|\boldsymbol{\pi}) = \frac{\exp\left[-\frac{1}{2}\left((\boldsymbol{\xi}_\pi - \hat{\boldsymbol{\xi}})^{\mathrm{t}}\, \mathbf{C}^{-1}\,(\boldsymbol{\xi}_\pi - \hat{\boldsymbol{\xi}})\right)\right]}{(2\pi)^{d/2}\,|\mathbf{C}|^{\frac{1}{2}}}, \tag{25}$$

where $\hat{\boldsymbol{\xi}}$ denotes the mean data vector, $\boldsymbol{\xi}_\pi$ the model data vector, d is the dimension of the data vectors, hence $|\mathbf{C}|$ is the determinant of a $d \times d$ covariance matrix. We note that the 2PCF data vector consists of two parts ($\boldsymbol{\xi}_+$ and $\boldsymbol{\xi}_-$), each with $d/2$ bins. The evidence is a normalization obtained by integrating the likelihood over the considered parameter space

$$p(\boldsymbol{\xi}) = \int \mathrm{d}\boldsymbol{\pi}\, \frac{\exp\left[-\frac{1}{2}\left((\boldsymbol{\xi}_\pi - \hat{\boldsymbol{\xi}})^{\mathrm{t}}\, \mathbf{C}^{-1}\,(\boldsymbol{\xi}_\pi - \hat{\boldsymbol{\xi}})\right)\right]}{(2\pi)^{d/2}\,|\mathbf{C}|^{\frac{1}{2}}}. \tag{26}$$

In our case, we calculate $\hat{\boldsymbol{\xi}}$ from P_{E} with Eq. (3) 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 using the determinant of the second-order moment of the posterior likelihood (see Kilbinger & Schneider 2004)

$$Q_{ij} \equiv \int \mathrm{d}^2\boldsymbol{\pi}\, p(\boldsymbol{\pi}|\boldsymbol{\xi})\, (\pi_i - \pi_i^{\mathrm{f}})(\pi_j - \pi_j^{\mathrm{f}}), \tag{27}$$

where π_i, π_j are the varied parameters, and π_i^{f}, π_j^{f} are the parameter of the fiducial model. Here, the indices i and j can take the values 1 (corresponding to Ω_m) or 2 (corresponding to σ_8). The square root of the determinant is given by

$$q = \sqrt{|Q_{ij}|} = \sqrt{Q_{11}Q_{22} - Q_{12}^2}, \tag{28}$$

and can be considered as our figure-of-merit quantity. Smaller credible regions in parameter space correspond to a lower value of q. In this paper, all q are given in units of 10^{-4}.

4.2. Results of the likelihood analysis

In Sect. 3, we calculate 2500 covariances covering a parameter range of $\Omega_m \in [0.2, 0.4]$ and $\sigma_8 \in [0.6, 1.0]$. We examine how the CDC-effect influences the likelihood contours, and hence, for each of the 2500 covariance matrices, we perform a likelihood analysis. In these analyses, we consider the same parameter space, similar priors, and similar $\hat{\boldsymbol{\xi}}$ and $\boldsymbol{\xi}_\pi$, only the covariance in Eq. (25) is changed. The left panel of Fig. 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. 3), the contours broaden for increasing Ω_m and σ_8.

Without any information about which cosmology to choose in our covariance matrix, it is reasonable to include prior information from other cosmological probes. The middle panel of Fig. 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-year analysis (Komatsu et al. 2008). Compared to the left panel, the deviation in the contours reduces significantly, although it remains 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. We calculate the values of q (Sect. 4.1) for all contour plots and summarize them in Table 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. 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. 3, the strength of the CDC-effect, i.e., the gradient
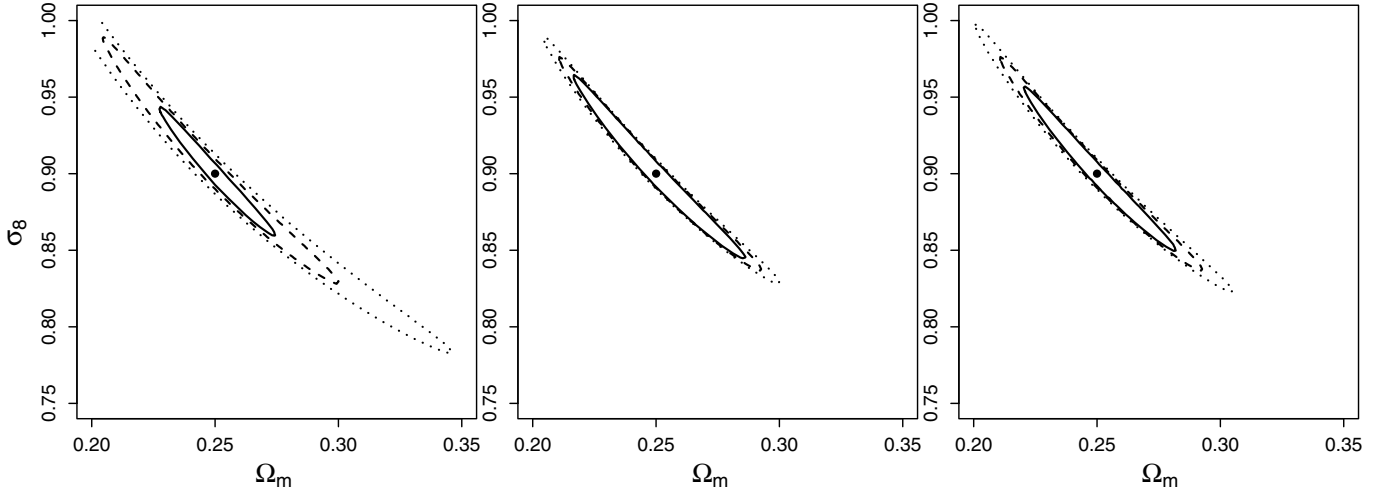
Fig. 3. The 95%-credible intervals obtained from likelihood analyses with different cosmological models assumed in their covariance matrix. *The left panel* corresponds to the following covariance parameters: $\Omega_m = 0.2$, $\sigma_8 = 0.6$ (*solid*), $\Omega_m = 0.25$, $\sigma_8 = 0.9$ (*dashed*), and $\Omega_m = 0.4$, $\sigma_8 = 1.0$ (*dotted*). The middle panel shows the deviation that 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*).

Table 1. Values of q for different covariance models.

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

of the curves in Fig. 4, depends on the considered point in parameter space. For the fiducial model we calculate $(\partial q/\partial \Omega_m)_{fid} = 7.5$ and $(\partial q/\partial \sigma_8)_{fid} = 3.5$.

4.3. Impact of survey parameters on the CDC-effect

We have shown, that the CDC-effect non-negligibly affects the likelihood contours. However, we have only quantified 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 the survey size A, the ellipticity dispersion σ_ϵ, and the number density of source galaxies $\bar{n}$, where in the case of the latter two, only the combination $\sigma_\epsilon^2/\bar{n}$ is of interest. We perform likelihood analyses 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 Ω_m and σ_8, defined to be $\Delta q = q(\pi_{max})/q(\pi_{min})$. The minimum value of q is obtained when choosing the minimum parameter set in the calculation of the covariance, i.e., $\pi_{min} = (\Omega_m = 0.2, \sigma_8 = 0.6)$. Correspondingly, choosing the maximum parameter set $\pi_{max} = (\Omega_m = 0.4, \sigma_8 = 1.0)$ results in the maximal value of 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 Eqs. (8)–(10) 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, Eqs. (8)–(10) 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 that the strength of CDC-effect is independent of A. However, when considering the likelihood, the inverse covariance enters in the exponent, and the values of q are furthermore an integral over the posterior likelihood. This non-linearity in the inverse covariance causes the strength of the CDC-effect to vary 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.
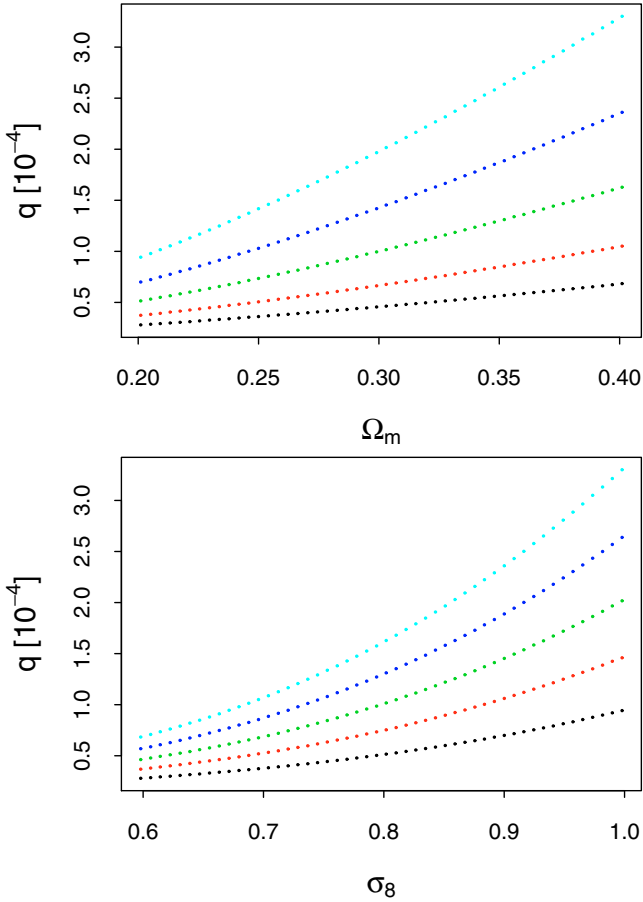
The upper panel of Fig. 5 shows $\Delta q = q(\pi_{max})/q(\pi_{min})$ as a function of $\sigma_\epsilon^2/\bar{n}$. The ratio Δq changes from 4 to 18 over the considered interval of $\sigma_\epsilon^2/\bar{n}$. When increasing the survey size A (Fig. 5, lower panel), we find that the impact of the CDC-effect increases from $\Delta q = 2.5$ (for a 25 deg^2 survey) to $\Delta q = 12.3$ (for a 2500 deg^2 survey). We note that the size of the likelihood contours, and 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 $\sigma_\epsilon^2/\bar{n}$.

5. Likelihood analysis with a model-dependent covariance

5.1. Adaptive covariance matrix

For a given cosmological model, we can calculate the covariance directly from Eqs. (5)–(10). 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 (25) as
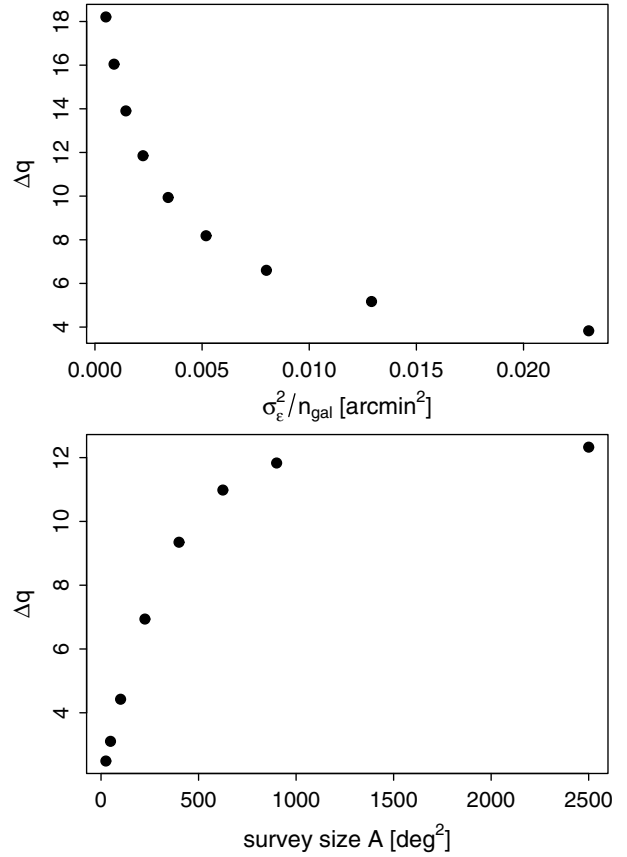
$$p(\xi|\pi) = \frac{\exp\left[-\frac{1}{2}\left((\xi_\pi - \hat{\xi})^t \, \mathbf{C}_\pi^{-1} \, (\xi_\pi - \hat{\xi})\right)\right]}{(2\pi)^{d/2} \, |\mathbf{C}_\pi|^{\frac{1}{2}}} . \tag{29}$$

Fig. 4. The values of q depending on Ω_{m} (*top*), the individual lines in each figure corresponding 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*) $\Omega_{\mathrm{m}} = [0.2, 0.25, 0.3, 0.35, 0.4]$.

Fig. 5. The ratio of maximum to minimum value of q depending on the ratio $\sigma_\epsilon^2/\bar{n}$ (*upper panel*) and depending on the survey size A (*lower panel*).

Compared to the case of a constant covariance, there are two main differences. First, the covariance in the exponential term of Eq. (29) 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 (see Eq. (26)). As a consequence, the posterior likelihood does not only depend on the exponential terms, which basically compare ξ_π and $\hat{\xi}$, 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. 6 shows the likelihood contours for a 84 deg² survey, where the posterior probability is calculated with the new likelihood in Eq. (29). For comparison, the right panel shows the likelihood contours when neglecting the parameter dependence in the determinant terms, and 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. To explain this shift, we overlay the right panels of Fig. 6 with the contours of constant $|\mathbf{C}_\pi|^{-1/2}$ (for numerical reasons, we plot $\ln |\mathbf{C}_\pi|^{-1/2}$). We can 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, $|\mathbf{C}_\pi|^{-1/2}$ induces a parameter-dependent weighting, which increases the likelihood

at small Ω_{m} and σ_8 and vice versa decreases it for large values of Ω_{m} and σ_8.

In general, the exponential term dominates the likelihood, $|\mathbf{C}_\pi|^{-1/2}$ has a significant impact only on the 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. $\times \Omega_{\mathrm{m}}^{-0.6}$. With respect to these curves, the contours of constant $|\mathbf{C}_\pi|^{-1/2}$ are rotated slightly, which allows 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 higher $|\mathbf{C}_\pi|^{-1/2}$ compared to the right panel. We note that for a different parameter combination this bias might not cause such a large shift in the best-fit value.

The second row of Fig. 6 shows the same analysis but for a 900 deg² survey. Comparing the left and right panel, we see that the likelihood contours are, similar to the 84 deg² survey, shifted and stretched towards regions of higher $|\mathbf{C}_\pi|^{-1/2}$. 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

$$p(\pi|\xi) = \frac{\exp\left[-\frac{1}{2}\left((\xi_\pi - \hat{\xi})^{\mathrm{t}}\, \mathbf{C}_\pi^{-1}\, (\xi_\pi - \hat{\xi})\right)\right]}{\int \mathrm{d}\pi' |\mathbf{C}_{\pi'}^{-1}\, \mathbf{C}_\pi|^{\frac{1}{2}} \exp\left[-\frac{1}{2}\left((\xi_{\pi'} - \hat{\xi})^{\mathrm{t}}\, \mathbf{C}_{\pi'}^{-1}\, (\xi_{\pi'} - \hat{\xi})\right)\right]} \cdot (30)$$

Compared to the case of a constant covariance the above expression has an additional factor in the denominator, i.e. $|\mathbf{C}_\pi \mathbf{C}_{\pi'}^{-1}|^{1/2}$. We 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
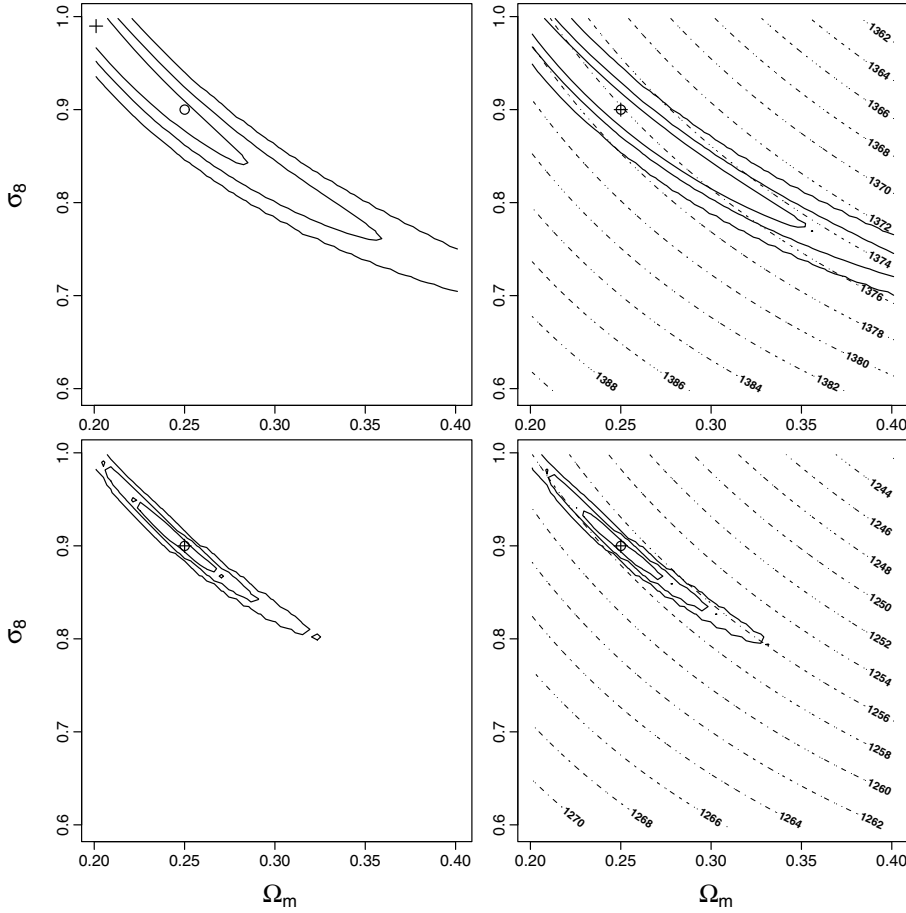
Fig. 6. The left plots shows the likelihood contours when using a model-dependent covariance, more explicitly, when calculating the posterior from Eq. (29). The cross illustrates the best-fit value, whereas the circle indicates our fiducial model. The panels on the right-hand side show the likelihood contours obtained when neglecting the determinant-terms in Eq. (29). 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.

covariance determinant becomes negligible for sufficiently large surveys.

5.2. Fisher matrix analysis

We expect tighter constraints on cosmological parameters if the cosmology dependences of both the mean data vector and the covariance matrix are incorporated into the likelihood analysis, instead of the mean data vector alone (Tegmark et al. 1997). The Fisher information matrix can be used to illustrate this effect; its definition reads (Kendall & Stuart 1979; Tegmark et al. 1997)

$$\mathbf{F}_{ij} = \left\langle \frac{\partial^2 L(\boldsymbol{\pi})}{\partial \pi_i\, \partial \pi_j} \right\rangle = \left(\frac{\partial^2 L(\boldsymbol{\pi})}{\partial \pi_i\, \partial \pi_j} \right)_{\boldsymbol{\pi} = \boldsymbol{\pi}_{\mathrm{ML}}}, \qquad (31)$$

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.

In addition, if we complete a Taylor expansion of $L(\boldsymbol{\pi})$ in parameter space at the fiducial parameters, we derive

$$L(\boldsymbol{\pi}) = L(\boldsymbol{\pi}_{\mathrm{fid}}) + 0 + \frac{1}{2}(\boldsymbol{\pi} - \boldsymbol{\pi}_{\mathrm{fid}})^{\mathrm{t}}\, \mathbf{T}^{-1}(\boldsymbol{\pi} - \boldsymbol{\pi}_{\mathrm{fid}}) + O(\Delta_\pi^3), \qquad (32)$$

where

$$(\mathbf{T}^{-1})_{ij} = \left(\frac{\partial^2 L(\boldsymbol{\pi})}{\partial \pi_i\, \partial \pi_j} \right)_{\boldsymbol{\pi} = \boldsymbol{\pi}_{\mathrm{fid}}}. \qquad (33)$$

The first-order term vanishes since $(\partial L/\partial \boldsymbol{\pi})_{|\boldsymbol{\pi}_{\mathrm{fid}}}$ is zero, and hence Eq. (32) is dominated by second-order terms. In this analysis. we consider only Ω_{m} and σ_8; $\boldsymbol{\pi}_{\mathrm{ML}}$ corresponds to our fiducial model,

i.e., $\boldsymbol{\pi}_{\mathrm{fid}} = (\Omega_{\mathrm{m}} = 0.25, \sigma_8 = 0.9)$. Comparing Eqs. (33) and (31) one can see that the Fisher matrix and the inverse parameter covariance matrix $\mathbf{T}^{-1}$ are equal. We rewrite Eq. (32) as

$$L(\boldsymbol{\pi}) = L(\boldsymbol{\pi}_{\mathrm{fid}}) + \frac{1}{2}(\boldsymbol{\pi} - \boldsymbol{\pi}_{\mathrm{fid}})^{\mathrm{t}}\, \mathbf{F}\, (\boldsymbol{\pi} - \boldsymbol{\pi}_{\mathrm{fid}}) + O(\Delta_\pi^3). \qquad (34)$$

For a given Fisher matrix, this equation enables us to calculate lower bounds to $L(\boldsymbol{\pi})$, hence we can derive lower bounds to the likelihood contours. When $p(\boldsymbol{\xi}|\boldsymbol{\pi})$ 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 mean data vector and the data covariance matrix (e.g., Tegmark et al. 1997)

$$\mathbf{F}_{ij} = \frac{1}{2}\mathrm{tr}\left[\mathbf{C}^{-1}\mathbf{C}_{,i}\mathbf{C}^{-1}\mathbf{C}_{,j} + \mathbf{C}^{-1}\mathbf{M}_{ij} \right], \qquad (35)$$

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 $\mathbf{M}_{ij} \equiv \boldsymbol{\xi}_{,i}\boldsymbol{\xi}_{,j}^{\mathrm{t}} + \boldsymbol{\xi}_{,j}\boldsymbol{\xi}_{,i}^{\mathrm{t}}$. The first term of Eq. (35) vanishes in case the covariance matrix is constant in parameter space, the second term vanishes when the mean is constant. For cosmic shear, we have seen that neither the mean data vector, nor the covariance matrix are independent of cosmological parameters, and hence when calculating the Fisher matrix both terms must be taken into account. We recall that $\mathbf{C} \propto 1/A$, which holds also for the derivatives $\mathbf{C}_{,i}$, and hence the first term is independent of the survey size. The second term increases proportionally to the survey volume, and therefore the information gain in cosmological parameters, by incorporating the cosmology dependence of covariances, becomes less important for large surveys (see also Kilbinger & Munshi 2006, for a similar result).
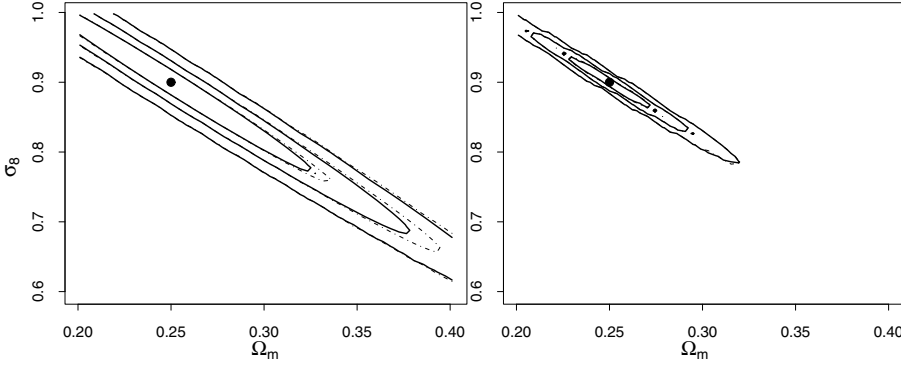
Table 2. The ML-parameter sets which occur when choosing different starting cosmologies in the iterative likelihood analysis.

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

Figure 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 the 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 greater for the case of non-Gaussian covariances for the following reason. Non-Gaussianity increases the cosmic variance term, becoming important in particular on small scales, which remain dominated by shot noise in the pure Gaussian case. Since the CDC-effect results mainly from the cosmic variance term, its strength also increases in the non-Gaussian case. A stronger dependence of the covariance on the parameters would increase the first term in Eq. (35), which implies that for the case of truly non-Gaussian covariances the improvement in parameter constraints is more significant than shown in Fig. 7.

5.3. Iterative likelihood analysis

In Sect. 5.1, we introduced the adaptive covariance, which is a proper way of incorporating cosmology-dependent covariances into a likelihood analysis. Its disadvantage is the large computational effort required, which is already high for Gaussian covariances. 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.

We now 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), neglecting the CDC-effect and approximating the covariance as a constant in parameter space. The error in the posterior likelihood caused by this approximation increases with increasing distance from the cosmology of the ray-tracing simulation. Since we are mainly interested in regions around the maximum likelihood parameter set, π_{ML}, this suggests the following strategy for a likelihood analysis. First, we perform an iterative likelihood analysis using Gaussian covariances to

derive π_{ML}. Then, we start a numerical simulation with this cosmology, derive a ray-tracing covariance, and perform the final likelihood analysis. This ansatz minimizes the errors caused by the CDC-effect in the region of interest and additionally incorporates non-Gaussianity.

To derive π_{ML} iteratively, we start from an arbitrary cosmology, calculate a Gaussian covariance matrix therefrom using Eqs. (5)–(10), and perform a likelihood analysis. Throughout this first iteration step, the covariance matrix remains 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. 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, because the model data vector (of the fiducial model) fits the fiducial data vector exactly, $\xi_{\pi_{fid}} = \hat{\xi}$. Table 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 that started from the fiducial model deviates from the others. Although the suggested values of π_{ML} are close to π_{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, this bias occurs if the function $\xi_\pi - \hat{\xi}$ does not fall off steeply enough around the ML-parameter set, which applies especially to higher-dimensional likelihood analyses.

Our iterative pre-analysis has converged to $\Omega_m = 0.26$, $\sigma_8 = 0.882$, however we "only" have a ray-tracing simulation with $\Omega_m = 0.25$, $\sigma_8 = 0.9$ available. Figure 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
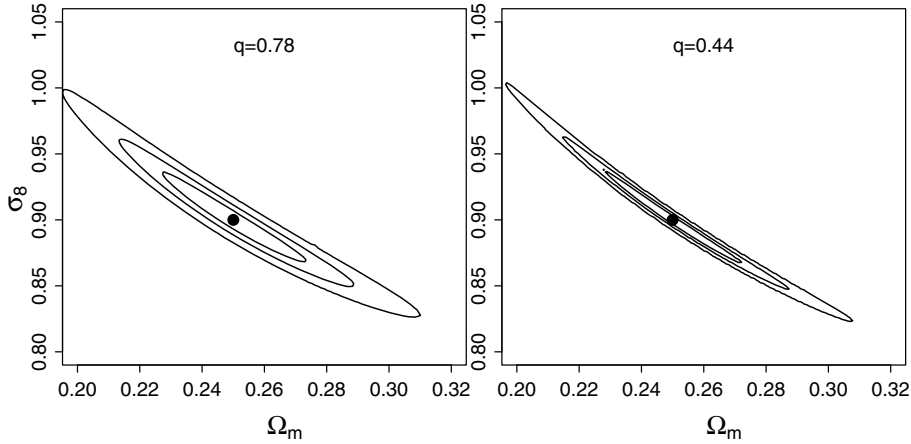
Fig. 8. The likelihood contours when using a ray-tracing covariance derived from the Millennium Simulation via field-to-field variation (*left panel*), compared to the case of a Gaussian covariance (*right panel*). Although the original size of each field is only 16 deg^2, we extrapolated the covariance to a 900 deg^2 survey. The values of q are given in units of 10^{-4}.

the Gaussian to 0.78×10^{-4} in the non-Gaussian case. We note that the value of q in the Gaussian case does not correspond to that in Table 1, because we use different survey parameters (here, $\sigma_\epsilon = 0.3$, $\bar{n} = 15/\text{arcmin}^2$) and a different data vector (here, 30 logarithmic bins from 0.2–130 arcmin) 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, and the impact is therefore 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 that we described in Sect. 4.2. However, the strength of the latter will most likely increase for non-Gaussian covariances, as we explained at the end of the last section.

6. Conclusions

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 for deriving covariances. First, one can calculate **C** analytically assuming a Gaussian shear field. This assumption breaks down on small angular scales (<10 arcmin), where non-linearities in the matter density field begin 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 by assuming a specific cosmology. In the first case, this cosmology is reflected in the power spectrum from which **C** is calculated; in the second case, we estimate **C** from numerical simulations, which are also based on a given cosmology. Past cosmic shear data analyses approximate the covariance as a constant in parameter space and assume that its underlying cosmology does not influence the result of a likelihood analysis significantly.

In this paper, 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 calculated 2500 Gaussian covariance matrices for various parameters of $\Omega_m \in [0.2, 0.4]$ and $\sigma_8 \in [0.6, 1.0]$; all other cosmological parameter were 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, in relation the CDC-effect will become 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, **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 cost. Using the analytic expression for Gaussian covariances is already time-consuming, and 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, because 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, although it minimizes the error around the maximum likelihood parameter set (π_{ML}). The method consists of two steps, first deriving π_{ML} with an iterative process, then deriving a ray-tracing covariance with π_{ML} as underlying cosmology and incorporating this into 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 that the starting point of the iteration might bias π_{ML}. This must be checked carefully before employing this method, otherwise the approximation of a constant covariance fails around π_{ML}. We note, that the results of this paper strongly depend on the non-linear fitting formula of Smith et al. (2003). It will be the subject to future work to repeat this analysis using numerical simulations directly to account fully for the non-linearities in the power spectrum.

References

Anderson, T. W. 2003, an Introduction to Multivariate Statistical Analysis (Wiley-Interscience), 623/624

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

Bartelmann, M., & Schneider, P. 2001, Phys. Rep., 340, 291

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

Eifler, T. 2009, Theoretical aspects of cosmic shear and its ability to constrain cosmological parameters, University of Bonn, Ph.D. Thesis, urn:nbn:de:hbz:5N-16908,
http://hss.ulb.uni-bonn.de/diss_online/math_nat_fak/2009/eifler_tim

Fu, L., Semboloni, E., Hoekstra, H., et al. 2008, A&A, 479, 9

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

Hetterscheidt, M., Simon, P., Schirmer, M., et al. 2007, A&A, 468, 859

Hilbert, S., Hartlap, J., White, S. D. M., & Schneider, P. 2008 [arXiv:0809.5035H]

Hoekstra, H., Mellier, Y., van Waerbeke, L., et al. 2006, ApJ, 647, 116

Joachimi, B., Schneider, P., & Eifler, T. 2008, A&A, 477, 43

Kaiser, N. 1998, ApJ, 498, 26

Kaiser, N., Wilson, G., & Luppino, G. A. 2000 [arXiv:0003338K]

Kendall, M., & Stuart, A. 1979, The Advanced Theory of Statistics (Charles Griffin & Company Limited)

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

Komatsu, E., Dunkley, J., Nolta, M. R., et al. 2008 [arXiv:0803.0547K]

Massey, R., Rhodes, J., Leauthaud, A., et al. 2007, ApJS, 172, 239

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

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

Smith, R. E., Peacock, J. A., Jenkins, A., et al. 2003, MNRAS, 341, 1311

Springel, V., White, S. D. M., Jenkins, A., et al. 2005, Nature, 435, 629

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

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

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