
Statistical Properties of the Cosmic Density Field beyond 2-point Statistics: Covariance Matrices and Density Split Statistics

Oliver Friedrich



München 2018

Statistical Properties of the Cosmic Density Field beyond 2-point Statistics: Covariance Matrices and Density Split Statistics

Oliver Friedrich

Dissertation
an der Fakultät für Physik
der Ludwig-Maximilians-Universität
München

vorgelegt von
Oliver Friedrich
aus Potsdam

München, den 20. August 2018

Erstgutachter: Prof. Dr. Bender

Zweitgutachter: Prof. Dr. Weller

Tag der mündlichen Prüfung: Freitag, 16. November 2018

Contents

| | |
|---|-----------|
| Zusammenfassung | xi |
| Abstract | xii |
| 1 Units and conventions | 1 |
| 2 The cosmological standard model | 3 |
| 3 Subject and outline of this work | 7 |
| I Background: Theory of cosmic structure formation | 13 |
| 4 Cosmology in the Newtonian approximation | 15 |
| 4.1 Expansion and collapse of a homogeneous, spherical dust cloud | 16 |
| 4.2 A universe filled with dust ($R \rightarrow \infty$) | 17 |
| 4.3 Equations of motion for density fluctuations | 18 |
| 4.3.1 Coordinate transformations | 19 |
| 5 Cosmology in General Relativity | 21 |
| 5.1 Friedmann Universe | 22 |
| 5.1.1 FLRW-metric and Friedmann equations | 22 |
| 5.1.2 Thermodynamics | 24 |
| 5.1.3 Dust: $\epsilon \sim a^{-3}$, $w = 0$ | 25 |
| 5.1.4 Radiation: $\epsilon \sim a^{-4}$, $w = \frac{1}{3}$ | 26 |
| 5.1.5 Vacuum energy density: $\epsilon \equiv \text{const.}$, $w = -1$ | 26 |
| 5.1.6 The Λ CDM model | 26 |
| 5.2 The inhomogeneous universe | 27 |
| 5.2.1 The Einstein tensor with perturbations | 28 |
| 5.2.2 Perturbations in T_μ^ν for cold matter | 28 |
| 5.2.3 The Einstein equations for perturbations and non-relativistic limit . | 30 |

| | |
|--|-----------|
| 6 Theory of cosmic structure formation | 33 |
| 6.1 Initial conditions | 33 |
| 6.2 Linear evolution | 35 |
| 6.3 Lagrangian coordinates and symmetric initial conditions | 38 |
| 6.3.1 Spherical collapse | 38 |
| 6.3.2 Cylindrical and planar collapse | 40 |
| 6.4 Non-linear perturbation theory | 40 |
| 6.5 N-point functions | 43 |
| 7 The cosmic density PDF | 45 |
| 7.1 From the cumulant generating function to the PDF | 46 |
| 7.2 Path integral approach for the CGF | 47 |
| 7.3 Non-linear rescaling of the CGF | 48 |
| II Background: Observations of the large scale Structure | 51 |
| 8 Distances and redshift estimation | 53 |
| 8.1 Distance measures | 53 |
| 8.2 Redshift estimation | 57 |
| 9 2D power spectra and correlation functions | 59 |
| 9.1 2-point statistics on the sky | 59 |
| 9.2 Galaxy clustering 2-point function and galaxy bias | 62 |
| 10 Gravitational lensing | 67 |
| 10.1 Gravitational light deflection | 67 |
| 10.2 Lensing potential, convergence and shear | 69 |
| 10.3 Shapes of background galaxies as tracers of foreground density | 73 |
| 10.4 Cosmic Shear | 75 |
| 10.4.1 Lensing by the Cosmos | 75 |
| 10.4.2 Sources with a redshift distribution and convergence power spectrum | 77 |
| 10.4.3 Cosmic shear correlation functions | 77 |
| 10.4.4 Galaxy-galaxy lensing | 80 |
| 10.4.5 Intrinsic alignments | 83 |
| 11 The Dark Energy Survey | 87 |
| III Estimation of Covariance Matrices for 2-point correlation functions | 93 |
| 12 Covariance matrices and the derivation of parameter constraints | 95 |
| 12.1 Covariance matrices | 95 |

| | | |
|-----------|--|------------|
| 12.2 | Frequentist approach on constraining model parameters | 97 |
| 12.2.1 | Choice of the figure-of-merit as a contract between scientists | 97 |
| 12.2.2 | Prior knowledge and model parametrization | 98 |
| 12.2.3 | Criticism | 98 |
| 12.3 | Bayesian approach on constraining model parameters | 99 |
| 12.3.1 | Probability as degree of personal belief | 99 |
| 12.3.2 | Criticism | 100 |
| 13 | Performance of internal covariance estimators for cosmic shear correlation functions | 101 |
| 13.1 | Introduction | 102 |
| 13.2 | Cosmic Shear Basics | 104 |
| 13.2.1 | Cosmic Shear Correlation Functions | 104 |
| 13.2.2 | Covariance of the Correlation Functions | 105 |
| 13.3 | Log-normal Simulations | 108 |
| 13.3.1 | Setup and Validation of the Simulations | 109 |
| 13.4 | Internal Covariance Estimation for two-point correlation functions | 112 |
| 13.4.1 | Subsample Covariance | 114 |
| 13.4.2 | Correlation of sub-samples | 115 |
| 13.4.3 | Galaxy pairs crossing between sub-samples | 118 |
| 13.4.4 | Jackknife | 119 |
| 13.4.5 | Bootstrap Covariance | 120 |
| 13.4.6 | Stability and Inversion of the Covariance Estimate | 121 |
| 13.5 | Testing internal Covariance Estimators on simulated cosmic Shear Surveys | 122 |
| 13.5.1 | Constraints on cosmological Parameters | 124 |
| 13.5.2 | Matching the procedure to DES science verification and year 5 Data | 131 |
| 13.6 | Conclusions | 134 |
| 13.A | Correlation Matrices and Constraints from empirical Covariance | 138 |
| 13.B | Likelihood Contours | 138 |
| 14 | Precision matrix expansion – efficient use of numerical simulations in estimating errors on cosmological parameters | 145 |
| 14.1 | Introduction | 146 |
| 14.2 | Parameter constraints from noisy covariance estimates | 147 |
| 14.3 | Precision matrix expansion | 151 |
| 14.3.1 | Estimating the expansion of Ψ | 151 |
| 14.4 | Examples: parameter errors for LSST weak lensing and DES weak lensing and multi-probe analyses | 152 |
| 14.4.1 | Performance for DES weak lensing data vector | 156 |
| 14.5 | Conclusions | 162 |
| 14.A | Influence of noisy covariance estimates on the scatter of best fitting cosmological parameters | 166 |
| 14.B | Unbiased estimator of the square of a Wishart matrix | 166 |

| | |
|--|------------|
| 14.C General properties and convergence of the power series | 167 |
| 14.C.1 General properties | 167 |
| 14.C.2 Special cases | 167 |
| 14.C.3 Convergence in the General Case | 172 |
| 14.D Data vectors | 173 |
| 14.D.1 Weak lensing data vectors | 173 |
| 14.D.2 Lens galaxies | 173 |
| 14.D.3 Binning and scales | 174 |
| IV Density Split Statistics | 175 |
| 15 Idea and formalism | 177 |
| 16 Density split statistics: | |
| joint model of counts and lensing in cells | 185 |
| 16.1 Introduction | 187 |
| 16.2 Density split statistics: data vector, modeling and forecasts | 189 |
| 16.2.1 Measuring density split statistics | 189 |
| 16.2.2 Modeling density split statistics | 193 |
| 16.2.3 Data vector and forecasts on parameter constraints | 198 |
| 16.3 Simulated data and covariance matrix | 204 |
| 16.3.1 Buzzard mock galaxy catalogs | 204 |
| 16.3.2 Simulated density and convergence fields from FLASK and covariance matrix | 205 |
| 16.4 Modeling details and comparison to simulations | 206 |
| 16.4.1 Projected density PDF | 206 |
| 16.4.2 Convergence profile around lines-of-sight of fixed density contrast $\delta_{m,T}$ | 212 |
| 16.4.3 Shot-noise, stochasticity and Counts-in-Cells | 216 |
| 16.4.4 Summary of fiducial model and approximations therein | 224 |
| 16.5 Recovering cosmology in N-body simulations | 225 |
| 16.5.1 Simulated likelihood analysis | 225 |
| 16.6 Discussion & conclusions | 227 |
| 16.A Friedmann equations, linear growth, spherical collapse and cylindrical collapse | 229 |
| 16.B Λ CDM perturbation theory | 230 |
| 16.B.1 Second order of δ in Einstein-de Sitter universe | 231 |
| 16.B.2 Second order of δ in Λ CDM universe | 232 |
| 16.B.3 Bispectrum and 3-point function at leading order | 232 |
| 16.B.4 Variance and skewness of the density contrast in long cylinders at leading order in perturbation theory | 233 |
| 16.B.5 The moment $\langle \delta_{R_A,L}^2 \delta_{R_B,L} \rangle_{\text{tree}}$ | 236 |
| 16.B.6 The moment $\langle \delta_{R_A,L}^n \delta_{R_B,L} \rangle_{\text{tree}}$ | 237 |
| 16.C Comparison with Millennium simulation | 237 |

| | |
|--|------------|
| 16.D Galaxy stochasticity | 237 |
| 16.E Validation of alternative shot-noise model | 240 |
| 17 Density split statistics: cosmological constraints from counts and lensing in cells in DES Y1 and SDSS | 243 |
| 17.1 Analysis overview | 243 |
| 17.2 Constraints on the skewness of the density PDF | 250 |
| 17.3 Comparing different models of the matter-galaxy relation | 251 |
| V Summary and open questions | 253 |
| 18 Contributions of my thesis to the field of large scale structure cosmology | 255 |
| 19 Open questions in covariance estimation | 257 |
| 19.1 Non-Gaussian likelihoods | 257 |
| 19.2 Covariance estimation for analyses that combine different cosmological probes | 258 |
| 20 Open questions and tasks for density split statistics | 261 |
| 20.1 Improving the analysis of matter-galaxy relation | 261 |
| 20.2 Improving the analysis of gravitational non-linearities | 262 |
| 21 Cosmology and thermodynamics of spacetime | 265 |
| A The Einstein tensor in conformal Newtonian gauge | 269 |
| A.1 Without Perturbations | 269 |
| A.2 With scalar Perturbations | 270 |
| A.3 Perturbations of $T_{\mu\nu}$ | 276 |
| B 2-point correlation function on curved sky | 277 |
| C Use case for combining internal covariance estimates and PME | 279 |
| D Full author list of Friedrich et al. (2018) | 283 |
| Note of Thanks | 306 |

Zusammenfassung

Dunkle Materie, dunkle Energie, kosmische Inflation - unser Verständnis der drei Hauptzutaten des kosmologischen Standardmodels ist nach wie vor gering. Das deutet auf allgemeine Lücken in unserem Verständnis der Physik hin. Eine Observable mit dem Potential, diese Lücken zu schließen ist die großskalige Struktur von Dichtefluktuationen im Universum. Beobachtungen der Entwicklung dieser Struktur bei unterschiedlichen Rotverschiebungen können Aufschluss über das genaue Verhalten von dunkler Materie und dunkler Energie geben. Außerdem erlauben solche Beobachtungen Rückschlüsse über die Anfangsbedingungen des Universums, was uns Hinweise auf den genauen Mechanismus der kosmischen Inflation geben kann.

Die späten Stadien in der Entwicklung der großskaligen Struktur sind besonders schwer zu analysieren. Ein Grund dafür ist, dass die Differenzialgleichungen, die das Wachstum von Dichtefluktuationen beschreiben, im späten Universum nicht mehr gut durch lineare Gleichungen approximiert werden können. Mit dieser Arbeit präsentiere ich in zweierlei Hinsicht Fortschritte in der Behandlung der späten Strukturbildung. Zum einen verbessere ich Techniken zur Schätzung der statistischen Unsicherheiten in Messungen von 2-Punkt Statistiken des kosmischen Dichtefeldes. Das geschieht in einem ersten Schritt, indem ich die Leistung vorhandener Methoden untersuche und verbessere. Und in einem zweiten Schritt, indem ich eine eine völlig neue Methode präsentiere, die den exorbitanten Rechenaufwand einer weit verbreiteten Prozedur umgeht.

Zum zweiten entwickle ich ein theoretisches Modell für eine neue kosmologische Untersuchungsmethode namens Density Split Statistics. Damit wird es ermöglicht die lokale Wahrscheinlichkeitsdichtefunktion von Fluktuationen des Materiedichtefeldes zu studieren. Das eröffnet ein reiches Spektrum an Informationen über die großskalige Struktur des Universums, die mit rein auf 2-Punkt Statistik basierenden Analysen nicht zu erhalten wäre.

Abstract

Dark matter, dark energy, cosmic inflation - the three main ingredients of the cosmological standard model remain poorly understood. This points to general gaps in our understanding of physics. An observable that has the potential to fill these gaps is the large scale structure of density fluctuations in the universe. Observations of the evolution of that large scale structure over a range of different redshift can be used to study the exact behavior of dark matter and dark energy. Also, such observations can improve our understanding of the initial conditions of the universe and hence point us to the exact mechanism of inflation.

The late stages in the evolution of the large scale structure are particularly difficult to understand. One reason for this is that the differential equations governing the growth of density fluctuations are not well approximated by linear equations in the late time universe. In this work I advance the treatment of late time structure formation in two ways. First, I improve techniques to estimate the statistical uncertainties in measurements of the 2-point statistics of the cosmic density field. This is done in a first step by investigating and refining the performance of existing estimators. And in a second step, by proposing an entirely new method that can bypass the exorbitant computational needs of a prominent existing procedure.

Secondly, I develop a theoretical model for a new cosmological probe called density split statistics. This enables the study of the probability density function (PDF) of matter density fluctuations in the universe. This opens up a rich amount of information about the large scale structure of the universe that is otherwise missed if one only analyses 2-point statistics.

Chapter 1

Units and conventions

Units

Unless stated otherwise I use the natural unit system in this work, i.e. the units where the vacuum speed of light c , the gravitational constant G , the reduced Planck constant $\hbar = h/2\pi$ and the Boltzmann constant k_B are all set to 1 (which means that everything is measured in units of these constants).

Vector and matrix notation

Lowercase bold letters such as $\mathbf{v}, \mathbf{w}, \mathbf{d}$ denote vectors. These might be vectors in 3-dimensional space or vectors of data points in a data analysis. Uppercase bold letters such as $\mathbf{A}, \mathbf{B}, \mathbf{C}$ but also uppercase bold Greek symbols such as $\boldsymbol{\Psi}$ denote matrices and \mathbf{A}^{-1} etc. are the corresponding inverse matrices.

Fourier transform

If $f(\mathbf{x})$ is a function in real space, then its Fourier transform (if it exists) is defined as

$$\hat{f}(\mathbf{k}) = \int d^3x \ f(\mathbf{x}) \ e^{-i\mathbf{x}\cdot\mathbf{k}} . \quad (1.1)$$

The inverse transform is then given by

$$f(\mathbf{x}) = \int \frac{d^3k}{(2\pi)^3} \ f(\mathbf{k}) \ e^{i\mathbf{k}\cdot\mathbf{x}} . \quad (1.2)$$

Special and General Relativity and 4-vector notation

In the context of special relativity symbols like x^μ, u^μ (resp. $x^\nu, x^\alpha, u^\nu, u^\alpha$) etc. denote the components of contra-variant 4-vectors and any Greek indices can have values from 0 to 3. Latin indices (as in x^i, u^j etc.) are used to reference spatial components of 4-vectors and range from 1 to 3. The Minkowsky metric is given by $\eta = \text{diag}(1, -1, -1, -1)$ and its components are denoted by $\eta_{\mu\nu}$ (or with any other Greek indices). Latin indices as in η_{ij} reference the spatial part of η . The components of the inverse metric are denoted by $\eta^{\mu\nu}$ (and are within special relativity obviously identical to the metric components.)

When an index appear once as a subscript and once as a super script in an expression (as, e.g. , in $\eta_{\mu\nu}x^\nu$) than a summation over the possible values of that index is implied (Einstein's sum convention). For Greek indices this sum ranges from 0 to 3 and for Latin indices it ranges from 1 to 3. The position of a Greek index can be lowered or raised by contraction with the metric elements, i.e.

$$x_\mu \equiv \eta_{\mu\nu}x^\nu, \quad x^\mu = \eta^{\mu\nu}x_\nu. \quad (1.3)$$

Equivalent conventions apply in the context of general relativity, but with the metric and inverse metric components $\eta_{\mu\nu}$ and $\eta^{\mu\nu}$ being replaced the metric tensor $g_{\mu\nu}$ and its inverse $g^{\mu\nu}$.

Chapter 2

The cosmological standard model

Before I discuss the topic and outline of this work I want to briefly summarize four major phases of the evolution of the universe in the standard model of cosmology. This model is also called the Λ CDM-*model* where Λ is the cosmological constant in Einstein's equations and CDM stands for *Cold Dark Matter* (see the explanations below). A detailed description of the model is, e.g. , given in the books of Peebles (1993) and Mukhanov (2005) or in the reviews of Bernardeau et al. (2002); Bartelmann & Schneider (2001). Parts of the Λ CDM-model will also be reviewed in more detail in chapters 4 and 5 of this work.

Cosmic Inflation

13 Billion years ago the universe was filled by the condensate of a scalar quantum field ψ (the *Inflaton field*; cf. chapter 5 of Mukhanov 2005 for the following statements). The energy density ϵ and the pressure p of this condensate can be expressed in terms of the Hamiltonian and Lagrangian density of ψ as

$$\begin{aligned}\epsilon &\equiv \mathcal{H}(\psi, \partial_\mu \psi) \\ &= \frac{1}{2} \partial_\mu \psi \partial^\mu \psi + V(\psi) \\ p &\equiv \mathcal{L}(\psi, \partial_\mu \psi) \\ &= \frac{1}{2} \partial_\mu \psi \partial^\mu \psi - V(\psi) ,\end{aligned}\tag{2.1}$$

where V is the potential energy density of the field. The observable universe originates from a region of this condensate where due to quantum fluctuations $\frac{1}{2} \partial_\mu \psi \partial^\mu \psi \ll V(\psi)$ such that ϵ and p are related by $p \approx -\epsilon$. According to the theory of General Relativity (GR) such an equation of state leads to an exponential expansion of spacetime in that particular region of the condensate. This so-called *Inflation* continues for at least 60 *e-folds*, i.e. until that region has expanded by at least a factor of $> e^{60}$. During the expansion, the kinetic energy of the Inflaton field grows and its potential energy decreases until $\frac{1}{2} \partial_\mu \psi \partial^\mu \psi \sim V(\psi)$ at which point the exponential expansion stops. The Inflaton field now decays into particles of other quantum fields thus creating the matter content we observe in the universe today. Due to quantum fluctuations in the initial field ψ the end of Inflation is not reached at the

same time everywhere in the expanding patch. Instead, some parts of the universe expand for a slightly longer or shorter period of time. This leads to curvature perturbations in spacetime, which via the equations of GR propagate into inhomogeneities of the cosmic density field (cf. chapter 8 of Mukhanov 2005 or section 14.8 of Padmanabhan 2010a).

It should be stressed that this is a direct interaction between quantum physics and General Relativity!

Radiation dominated epoch

After the end of Inflation the universe is still expanding but at a much slower rate. This further slows down continuously due to the gravity of the universe's matter and radiation contents. For the next $\sim 50.000^1$ years the energy content of the universe was dominated by relativistic particles (i.e. by particles whose kinetic energy is much higher than their rest frame mass) in plasma of tightly coupled matter and radiation. For most of this time, the matter-radiation-plasma consists of photons and electrons as well as Hydrogen and Helium nuclei that are coupled through electromagnetic scattering. Initial density fluctuations in this plasma propagate in the form of sound waves. Especially, these fluctuations oscillate and cannot grow through gravitational attraction. The cosmic microwave background (CMB) that reaches us today can be thought of as the last snapshot of this epoch (even though it is produced ~ 380.000 years after the end of inflation, i.e. in the matter dominated epoch of the universe that is explained below). It is created when the universe cools down sufficiently to let electrons and Hydrogen nuclei combine to Hydrogen atoms - a process called *recombination* (see, e.g. , section 3.6 of Mukhanov, 2005).

Matter dominated epoch

The temperature fluctuations in the CMB indicate that the typical amplitude of density fluctuations in the matter-radiation-plasma wrt. its mean density $\bar{\rho}$ is (Tegmark & Rees, 1998)

$$\delta \equiv \frac{\rho - \bar{\rho}}{\bar{\rho}} \approx 10^{-5} \quad (2.2)$$

at the time of recombination. After recombination the newly formed Hydrogen (and also the previously formed Helium) is not coupled to radiation anymore and its density fluctuations can grow through gravitational interaction. The energy budget in the universe is now mostly made up off non-relativistic massive particles and the growth of density perturbations is approximately given by

$$\delta \sim \frac{1}{1+z} , \quad (2.3)$$

where z is the cosmological redshift at a given age of the universe (cf. chapter 5 of this thesis, or chapter 6 of Mukhanov (2005)). This means that until today ($z = 0$) the initial fluctuations at recombination ($z \approx 1100$) can only have grown by a factor of ≈ 1100 to typical amplitudes of $\sim 10^{-2}$. This is in stark contradiction to the density contrast

¹<http://www.oxfordreference.com/view/10.1093/oi/authority.20110803100140589>

observed on scales of galaxies and galaxy clusters today which can reach values of up to $\delta \approx 500$!

The cosmological standard model resolves this contradiction by postulating a substance called *dark matter* that interacts with other forms of matter and energy only via its gravitational force (see Plehn, 2017, and references therein). Hence, this substance would not have been coupled to the matter-radiation-plasma through electromagnetic interactions and its density fluctuations could have started to grow already during the radiation dominated epoch. It is then assumed that after recombination the ordinary matter (Hydrogen, Helium etc.) fell into the already existing gravitational potentials of dark matter density perturbations which boosted the structure formation of the visible matter. In order for this scenario to work, dark matter must make up about 80% of all non-relativistic matter in the universe.

Phase of accelerated expansion

If the energy content of the universe consists only of the known form of matter and radiation (and possibly dark matter) then the gravitational forces of these substances should ever slow down the cosmic expansion. Observations have however shown that $\sim 3 \cdot 10^9$ years ago the universe entered a new phase of accelerated expansion (e.g. Riess et al., 1998; Wang et al., 2017). The cosmological standard model assumes that this is caused by an additional substance called dark energy. The equation of state that relates the energy density ϵ and pressure p of this substance is postulated to be $p \approx -\epsilon$, i.e. similar to that of the Inflaton field at the beginning of Inflation. As cosmic expansion progresses the density of matter decreases until it falls below the energy density of dark energy. At this point the equation of state $p \approx -\epsilon$ of dark energy starts to dominate and accelerates the expansion of the cosmos again.

The nature of dark energy is as of yet unclear - it could be an additional term to be added in the Einstein equations (the *cosmological constant* term), a non-zero energy density of the vacuum, a dynamical field itself or any other deviation from the standard theories of physics (though the latter two possibilities would in fact go beyond the cosmological standard model).

Chapter 3

Subject and outline of this work

The gravitational forces within our solar system are very well described by the theory of General Relativity (GR). It correctly derives the almost perfectly Keplerian motion of most planets but also successfully predicts deviations from Newtonian physics such as the perihelion shift of mercury, the time dilation of satellites around earth or the shift of stellar positions due to gravitational lensing by the sun (see chapter 6 of Fließbach, 1990). Also, it accurately describes the motion of stars close to the central black hole of our galaxy (Parsa et al., 2017).

The interactions of atomic and subatomic particles on the other hand are astonishingly well described by quantum theory. Predictions of quantum theory and observations agree to a high precision, e.g. , regarding the magnetic dipole moment of the electron (Odom et al., 2006) or the emission and absorption spectra of atoms (including such subtleties as the Lamb shift of the ground state energy of the Hydrogen atom, Bethe 1947).

Despite the successes of both quantum theory and General Relativity it has been difficult to build up a theory that describes both quantum and general relativistic effects at the same time. It is often believed that this should be accomplished by constructing a linear space of geometries on which quantum superpositions of spacetimes could be defined (Dewitt, 1967; Hartle & Hawking, 1983; Rovelli, 2000). Apart from theoretical difficulties, progress on combining the two theories is also hindered by an experimental obstacle: so far no laboratory experiment could be realized in which a quantum theory of the gravitational field would be relevant.

Astronomical observations on cosmological scales ($\gtrsim 10$ Mpc) have revealed a number of shortcomings of our understanding of both GR and the standard model of particle physics that might fill in this experimental gap. The standard model of cosmology that was summarized in chapter 2 has three major ingredients that are not explained by our standard theories of gravity and particle physics. These are:

- **Inflation**

A phase of cosmic inflation is, e.g. , needed to explain the large scale homogeneity of the early universe as seen in CMB temperature fluctuations or the fact that the universe is almost perfectly flat. It is also the only mechanism that can explain the

generation of density fluctuations on small scales (see chapters 5 and 8 of Mukhanov, 2005). The Inflaton field is however not predicted by the standard model of particle physics. But it must interact with the ingredients of that model since the Inflaton field must decay into standard model particles at the end of Inflation. So - under the assumption that Inflation indeed took place - the standard model of particle physics must be an incomplete description of subatomic physics.

- **Dark Matter**

To explain the amplitude of density fluctuations today compared to those observed in the CMB (but also to explain other observations such as galaxy rotation curves) gravity must be stronger on the scales of galaxies and galaxy clusters than would be expected from GR and the amount of Baryonic matter (i.e. ordinary matter that interacts with the electromagnetic field). This could either be caused by the existence of a dark matter component or by deviations of gravity from Einstein's theory (see Plehn 2017 for a review about dark matter; see Milgrom 2001 for alternative considerations). In any case, this would represent a shortcoming of either GR or the standard model of particle physics. Recent hypotheses that link this excess gravity to the thermodynamics of quantum degrees of freedom of spacetime might even signify that this problem is directly relevant to the problem of a combined description of quantum and gravitational physics (Verlinde, 2016, see also chapter 21 of this thesis for a brief review).

- **Dark Energy**

If GR is valid on cosmological scales, then the accelerated expansion of the universe (that has, e.g. , been observed by measuring the luminosities of Type Ia supernovae or the angular size of the BAO feature as a function of redshift, Riess et al. 1998; Wang et al. 2017) requires the existence of a substance with the exotic equation of state $p \approx -\epsilon$. A candidate for this dark energy would be a constant energy density of the vacuum. The standard model of particle physics indeed predicts such a vacuum energy but with a value that is many orders of magnitudes higher than the observed value of dark energy density (Adler et al., 1995). To correct this, a so-called cosmological-constant-term would have to be added to the Einstein equations with an amplitude that must be tuned to almost perfectly cancel the contributions of the vacuum energy density. Such a fine tuning is ill motivated. Especially, it would mean that Einstein's equations know about the results of quantum Field theory.

Any of the above phenomena point to a possible breakdown of at least one of our two fundamental theories. This makes cosmology a rich subject for advancing our understanding of fundamental physics (and possibly for answering the question of how to combine our theories of gravity and quantum physics).

Large scale structure cosmology is the part of cosmology that analyses the properties and evolution of density fluctuations in the universe. This includes the study of distinct objects such as groups or clusters of galaxies or of the density contrast field $\delta(\mathbf{x})$ in the

observable universe as a whole. To confront the observed large scale structure (LSS) of the universe with cosmological theory one often (but not always) has to measure statistical properties of $\delta(\mathbf{x})$ as opposed to its exact configuration. The reason for this is that the theory of Inflation does not predict the exact initial conditions of the universe but only their statistical properties. One quantities that can be used to characterize the LSS is the *2-point correlation function*

$$\xi(\Delta\mathbf{x}, t) = \langle \delta(\mathbf{x}, t) \delta(\mathbf{x} + \Delta\mathbf{x}, t) \rangle. \quad (3.1)$$

Here the expectation value $\langle \cdot \rangle$ can be equivalently understood as a spatial average over all positions \mathbf{x} or as an ensemble average over all possible initial conditions (i.e. the quantum fluctuations of the Inflaton field). Since Inflation predicts a perfectly homogeneous and isotropic ensemble if initial conditions, the above definition depends neither in the location \mathbf{x} nor on the direction of $\Delta\mathbf{x}$, such that ξ is only a function of the absolute value $\Delta x = |\Delta\mathbf{x}|$. In the late universe $\xi(\Delta x, t)$ is typically decreasing with increasing Δx , indicating that more distant regions of the universe are less correlated wrt. their density contrast. Instead of the correlation function ξ one often considers its Fourier transform

$$P(\mathbf{k}, t) = \int d^3 \Delta x \xi(\Delta\mathbf{x}, t) e^{-i\Delta\mathbf{x} \cdot \mathbf{k}} \quad (3.2)$$

which is called the *power spectrum*. Because of the isotropy of the initial conditions it also only depends on the absolute value $k = |\mathbf{k}|$.

The exact shape of $P(\mathbf{k}, t)$ in the early universe can be used to test predictions of Inflation (Mukhanov, 2013) or to determine the abundance of ordinary matter, dark matter and radiation. A comparison of the amplitudes of $P(\mathbf{k}, t)$ at early and late times of the universe can be used to characterize the growth of structure and hence to test how gravity acts on cosmological scales. The power spectrum at the time of recombination can, e.g. , be inferred from the temperature anisotropies in the cosmic microwave background (CMB). To infer the late time power spectrum one can, e.g. , observe the distribution of galaxies. Figure 3.1 shows a reconstruction of the power spectrum in the late universe derived by Tegmark & SDSS Collaboration (2002). The galaxy density contrast (which I will denote as δ_g) is however only a tracer of the total matter density contrast δ . Galaxies form in small scale density peaks of the total density field and - depending on their mass - the clustering of these peaks can be either enhanced or reduced compared to the total matter density. This leads to a shift in the amplitude of the galaxy power spectrum wrt. the matter power spectrum which is called *galaxy bias*. Further complications are that galaxies are only sparse and stochastic tracers of the matter density field and also the fact that their selection function can be quite complex, i.e. whether or not a galaxy appears in a galaxy survey does not just depend on its mass and redshift but also on its optical properties.

An alternative way to measure the late time power spectrum is via the effect of *gravitational lensing*. As the light from distant galaxies travels towards us through the large scale structure of the universe the gravitational tical forces of the intervening density fluctuations distort the corresponding light cones. This leads to coherent distortion patterns

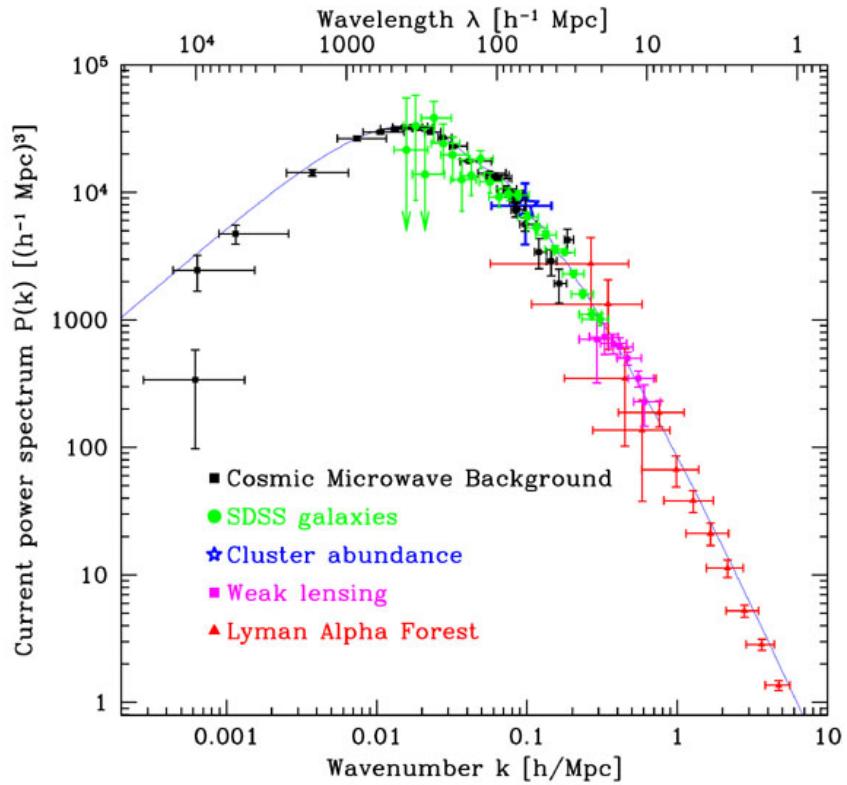


Figure 3.1: Taken from Tegmark & SDSS Collaboration (2002) who reconstructed the shape of the 3D matter power spectrum $P(\mathbf{k}, t)$ today from different large scale structure observables. The value of P at a given wavenumber k indicates the average squared amplitude of density fluctuations with a wave length of $\lambda \approx 2\pi/k$.

of the corresponding galaxy images - the so-called *cosmic shear field*. Since the gravitation of both dark and ordinary matter is assumed to act alike, the cosmic shear field is a tracer of the total matter density field. Measurements of cosmic shear 2-point correlation functions have been successfully used to determine the parameters of the cosmological standard model (Kilbinger et al., 2013; Heymans et al., 2013; Abbott et al., 2016; Hildebrandt et al., 2017; Troxel et al., 2017). Combined analyses of the 2-point correlation functions of the cosmic shear field, the galaxy density field and the cross-correlation function of the two fields have been used to improve these parameter constraints and to simultaneously determine the bias of the used galaxy samples (van Uitert et al., 2018; DES Collaboration et al., 2017). In these analyses, the 2-point functions have also been measured at a range of redshifts. DES Collaboration et al. (2017) used this to determine the equation of state of dark energy (DE) as a function of redshift and find no significant deviation from an exact relation $p_{\text{DE}} \equiv -\epsilon_{\text{DE}}$.

A fundamental limit to such analyses is the fact that we can only observe one realization of the universe. Future experiments such as *Euclid* (Laureijs et al., 2011) or *LSST* (Ivezic et al., 2008) will already perform analyses based on 2-point statistics with a close to all-sky coverage. At this point, if we want to extend the cosmological information obtained from the low redshift universe, we need to probe the large scale structure beyond its 2-point correlations. One way to do so is to study the abundance of density peaks in the universe as a function of their mass. This has, e.g. , been done by measuring the abundance of peaks in the cosmic shear field (Kacprzak et al., 2016; Shan et al., 2018; Martinet et al., 2018) or by studying the mass function of galaxy clusters (Mantz et al., 2016). These peak statistics are potentially very powerful but it is difficult to understand them theoretically since they probe the density field far into its non-linear regime.

An alternative way to explore the large scale structure beyond its 2-point statistics is to analyze higher order moments (i.e. moments beyond 2nd order) of the density field on mildly non-linear scales, i.e. on scales that can still partly be addressed analytically. Probes that are sensitive to the higher order moments of the density field have several advantages over pure 2-point statistics (Bernardeau et al., 2002; Takada & Jain, 2002; Pires et al., 2012; Uhlemann et al., 2018a): They can reduce the cosmic variance of LSS observation because gravitational non-linearities constantly transfer information about the initial conditions from the power spectrum to higher order spectra. They can be used to test generic, parameter independent predictions of the cosmological standard model (the Λ CDM-model) such as certain scaling relations between the variance of the density contrast δ and its higher order moments. They can break degeneracies between Λ CDM parameters and properties of the bias-relation between galaxies and total matter density. And, at high redshifts and on large scales, they can be used to search for non-Gaussianities in the initial density field which are predicted by models of Inflation that postulate more than just one Inflaton field (see Jeong & Komatsu, 2009, and references therein).

Another reason why an understanding of the higher order moments of the density field is important for a successful cosmological program is the fact that they appear in the covariance matrix of any measure of 2-point statistics (Schneider et al., 2002; Joachimi et al., 2008; Eifler et al., 2014; Crocce et al., 2016; Krause & Eifler, 2016). Estimating

these covariances from simulated realizations of the 2-point function measurements (e.g. in N-body simulations) is inherently noisy and can significantly deplete the obtained cosmological information even when a large number of simulations is available (Dodelson & Schneider, 2013; Taylor & Joachimi, 2014; Friedrich & Eifler, 2018). Future experiments hence require either an accurate analytical modeling of 4th order moments of the density field or a more efficient class of covariance estimators from simulated data.

This work considers both aspects of higher order moments of the density field: as a cosmological probe of their own and as covariance matrices for measures of 2-point statistics. Part I summarizes key concepts of large scale structure cosmology. I start with an overview of cosmology in the Newtonian approximation (chapter 4) followed by the full, relativistic treatment (chapter 5). Then I describe the evolution of density fluctuations within cosmological perturbation theory (chapter 6) and summarize the theory of computing the cosmic density PDF (chapter 7).

Part II gives an overview on how the theory of the large scale structure can be confronted with observable quantities. This includes a summary of different cosmological distance measures (chapter 8), a description of angular power spectra and correlation functions as line-of-sight projections of the corresponding 3-dimensional quantities (chapter 9) and an introduction to the gravitational lensing effect (chapter 10). Also, I give a brief introduction of the *Dark Energy Survey* (DES, chapter 11) which is the main source of data for this work.

In part III I consider the task of estimating the covariance matrix for measures of angular 2-point statistics of the density field. I summarize the role of covariance matrices in cosmological analyses in chapter 12 and present an analysis of the performance of different internal covariance estimators for cosmic shear 2-point correlation functions in chapter 13 which is based on the journal article Friedrich et al. (2016). Chapter 14 presents a completely new way of estimating the inverse covariance matrix (the *precision matrix*) from simulated data. This chapter is based on the journal article Friedrich & Eifler (2018) in which we call this method *precision matrix expansion*. In this method, the precision matrix is expanded as a power series around an analytical guess of the true matrix. We show there that the leading terms of this expansion can be estimated much more efficiently than the entire matrix.

Part IV introduces *Density Split Statistics* (DSS) - a new cosmological probe based on the theory of the cosmic density PDF. Chapter 15 summarizes the general idea behind this probe as well as the results of a relevant precursor study. Chapter 16 is based on the journal article of Friedrich, Gruen et al., 2018 and presents the modeling of density split statistics that I have developed based on perturbation theory and a cylindrical collapse approach. Chapter 17 summarizes our cosmological analysis of this probe in year-1 data of the Dark Energy Survey and is based on the journal article by Gruen, Friedrich et al., 2018.

Part V discusses conclusions from this work and gives a number of outlooks.

Part I

Background: Theory of cosmic structure formation

Chapter 4

Cosmology in the Newtonian approximation

In this chapter I will motivate important aspects of cosmology solely from considering the expansion and collapse of a cloud of pressure less matter within the Newtonian theory of gravity. The presentation mostly follows along the lines of Bernardeau et al. (2002); Mukhanov (2005).

In classical hydrodynamics a fluid without shear and viscosity can be described by the quantities

- $\rho(\mathbf{t}, \mathbf{r})$, the mass density of the fluid,
- $\mathbf{v}(\mathbf{t}, \mathbf{r})$, the streaming velocity of the fluid,
- $p(\mathbf{t}, \mathbf{r})$, the pressure of the fluid.

The evolution of such a fluid is governed by the continuity equation,

$$\frac{\partial \rho}{\partial t} + \nabla(\rho \mathbf{v}) = 0 , \quad (4.1)$$

and the Euler equation,

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla) \mathbf{v} + \frac{1}{\rho} \nabla p = \mathbf{g} . \quad (4.2)$$

Here $\mathbf{g} = \mathbf{g}(t, \mathbf{r})$ is the acceleration due to external forces. If the only external force acting on a fluid element is the gravitational force caused by the rest of the fluid then \mathbf{g} becomes

$$\mathbf{g}_{\text{grav}}(t, \mathbf{r}) = -\nabla \phi_{\text{grav}}(t, \mathbf{r}) , \quad (4.3)$$

where the gravitational potential ϕ_{grav} is given in terms of the fluid density as

$$\phi(t, \mathbf{r}) = -G \int d^3x \frac{\rho(t, \mathbf{x})}{|\mathbf{r} - \mathbf{x}|} \Leftrightarrow \Delta \phi(t, \mathbf{r}) = 4\pi G \rho(t, \mathbf{r}) . \quad (4.4)$$

Here G is the gravitational constant. The second equality of 4.4 is called the Poisson equation.

4.1 Expansion and collapse of a homogeneous, spherical dust cloud

Consider a spherical, homogeneous cloud of dust (where dust denotes any fluid with vanishing pressure, $p \equiv 0$). In the absence of external forces the Euler equation for this cloud reduces to

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla) \mathbf{v} = -\nabla \phi_{\text{grav}} , \quad (4.5)$$

while the continuity equation is still given by 4.1. If R_0 is the radius of the cloud at time t_0 , then its matter density is described by

$$\rho(t_0, \mathbf{r}) = \begin{cases} \rho_0 & \text{if } |\mathbf{r}| \leq R_0 \\ 0 & \text{else} . \end{cases} \quad (4.6)$$

Let us assume that the cloud has a homogeneous density distribution at all times. Looking at 4.1 this means that $\nabla \mathbf{v}$ within the radius of the cloud can only depend on t , i.e. within $|\mathbf{r}| \leq R(t)$ we have

$$\nabla \mathbf{v}(t, \mathbf{r}) = \vartheta(t) \mathbf{r} \quad (4.7)$$

end hence

$$\mathbf{v}(t, \mathbf{r}) = \frac{1}{3} \vartheta(t) \mathbf{r} + \boldsymbol{\alpha}(t) . \quad (4.8)$$

Here $\boldsymbol{\alpha}$ is a overall movement of the cloud that in the absence of external forces will be independent of t . We set $\boldsymbol{\alpha} \equiv \mathbf{0}$ to keep the cloud centered around $\mathbf{r} = \mathbf{0}$, arriving at

$$\mathbf{v}(t, \mathbf{r}) = \begin{cases} \frac{1}{3} \vartheta(t) \mathbf{r} & \text{if } |\mathbf{r}| \leq R_0 \\ 0 & \text{else} . \end{cases} \quad (4.9)$$

The velocity divergence ϑ describes an isotropic contraction or expansion of the cloud. As a consequence, the matter density of the cloud evolves as

$$\rho(t, \mathbf{r}) = \begin{cases} \rho_0 \left(\frac{R_0}{R(t)} \right)^3 & \text{if } |\mathbf{r}| \leq R(t) \\ 0 & \text{else} . \end{cases} \quad (4.10)$$

with a function $R(t)$ for which $R(t_0) = R_0$. To determine $R(t)$ consider a mass elements dm at the very edge of the dust cloud. The gravitational force acting on dm due to the matter inside the cloud is the same as the force that would be caused by a point mass at $\mathbf{r} = 0$ with mass

$$M = \frac{4\pi}{3} R_0^3 \rho_0 . \quad (4.11)$$

This force is

$$dF_{\text{grav}} = -\frac{dm MG}{R(t)^2} , \quad (4.12)$$

where the direction of the force is towards the center of the cloud. For the radial acceleration of the element dm this gives

$$\begin{aligned} \frac{d^2R}{dt^2} &= -\frac{MG}{R^2} \\ &= -\frac{4\pi}{3} G\rho R . \end{aligned} \quad (4.13)$$

Multiplying the first equality with $\dot{R} \equiv dR/dt$ and integrating it gives

$$(\dot{R})^2 = 2\frac{MG}{R} - K . \quad (4.14)$$

Here K is an integration constant that depends on the initial conditions of the cloud or, equivalently, on its total energy. The total energy of our dust cloud can be shown to equal

$$\begin{aligned} E_{\text{tot}} &= E_{\text{kin}} + E_{\text{grav}} \\ &= \frac{3M}{10} \left(\dot{R}^2 - 2\frac{GM}{R} \right) \\ &= -\frac{3M}{10} K . \end{aligned} \quad (4.15)$$

This shows that for $K > 0$ the cloud is gravitationally bound while $K < 0$ corresponds to an unbound cloud. For both cases, parametric solutions for $R(t)$ exist (Fosalba & Gaztanaga, 1998; Valageas, 2002a). For $K = 0$ an analytic solution is readily obtained by the ansatz $R \sim t^x$ as

$$R_{K=0}(t) = R_0 \left(\frac{t}{t_0} \right)^{2/3} . \quad (4.16)$$

Note that in Newtonian theory K is only determining the total Energy of the collapsing/expanding cloud while in General Relativity it also obtains a geometric meaning (cf. chapter 5). To conclude this section, consider the expansion rate $H \equiv \dot{R}/R$ with which equation 4.14 can be transformed to

$$H^2 = \frac{8\pi G}{3}\rho - \frac{K}{R^2} . \quad (4.17)$$

This emphasizes the structural similarity between the expansion of a cloud of dust in Newtonian theory and the Friedmann equations of relativistic cosmology (though the latter are much more general, cf. chapter 5).

4.2 A universe filled with dust ($R \rightarrow \infty$)

For simplicity, let us from now on assume that $K \equiv 0$ (for a more general discussion see chapter 5). Also, let us define the *scale factor* $a(t)$ by

$$a(t) = \frac{R(t)}{R_0} = \left(\frac{t}{t_0} \right)^{2/3} . \quad (4.18)$$

The time evolution of the cloud density is then given by

$$\rho(t) = \frac{\rho_0}{a(t)^3} \quad (4.19)$$

and the expansion rate of the cloud can be equivalently defined as

$$H(t) = \frac{\dot{a}}{a} . \quad (4.20)$$

Using the continuity equation, 4.1, one can furthermore express the velocity divergence ϑ as

$$\begin{aligned} \vartheta &= -\frac{\dot{\rho}}{\rho} \\ &= 3H . \end{aligned} \quad (4.21)$$

Keeping t_0 and ρ_0 constant but taking the limit $R_0 \rightarrow \infty$ changes the time evolution of neither a , ρ , H nor ϑ . This limit describes a universe that is completely filled with dust and whose velocity field is everywhere given by

$$\mathbf{v}(t, \mathbf{r}) = H(t)\mathbf{r} . \quad (4.22)$$

Such a universe is the only one consistent with the so-called *cosmological principle* (Mukhanov, 2005). This principle states that (at least on large scales) the universe is homogeneous and isotropic. Especially, changing the coordinate system to the inertial frame of any mass element dm in the universe leaves the form of 4.22 unchanged. The velocity given by 4.22 is called the *Hubble flow*.

Note that the gravitational potential as given by 4.4 diverges in the limit $R_0 \rightarrow \infty$. This is a conceptual short coming of the Newtonian approximation to cosmology that can only be fully resolved within general relativity using Birkhoff's theorem (see Birkhoff & Langer, 1923; Padmanabhan, 2010a). For now, let us only note that the potential

$$\phi_{\text{grav}}(t, \mathbf{r}) = \frac{2\pi G}{3}\rho(t)|\mathbf{r}|^2 \quad (4.23)$$

is a solution of the Poisson equation (second relation in 4.4) which is consistent with the Hubble flow described above.

4.3 Equations of motion for density fluctuations

The matter in the universe is obviously not distributed completely homogeneously. We now consider density perturbations on top of a homogeneous background density, i.e. we consider the density field

$$\rho = \bar{\rho} (1 + \delta) , \quad (4.24)$$

where $\bar{\rho}$ is given by 4.19 and depends only on t while $\delta = \delta(t, \mathbf{r})$ describes a spatially varying density contrast with respect to $\bar{\rho}$.

The density fluctuations will also cause the velocity field to deviate from the Hubble flow, 4.22. Let us parametrize these deviation as

$$\mathbf{v}(t, \mathbf{r}) = \bar{\mathbf{v}}(t, \mathbf{r}) + \mathbf{v}(t, \mathbf{r}) , \quad (4.25)$$

where $\bar{\mathbf{v}}(t, \mathbf{r}) = H(t)\mathbf{r}$ is the background Hubble flow and $\mathbf{v}(t, \mathbf{r})$ are velocity perturbations.

The perturbed gravitational potential can be parametrized as

$$\phi_{\text{grav}}(t, \mathbf{r}) = \frac{2\pi G}{3}\rho(t)|\mathbf{r}|^2 + \varphi(t, \mathbf{r}) \quad (4.26)$$

where the first term again represents the background potential governing the overall Hubble flow. The potential perturbations φ and the density contrast δ are connected through the modified Poisson equation (cf. equation 4.4)

$$\Delta\varphi(t, \mathbf{r}) = 4\pi G\bar{\rho}(t)\delta(t, \mathbf{r}) . \quad (4.27)$$

4.3.1 Coordinate transformations

To investigate how the fluctuations δ and \mathbf{v} evolve, it is convenient to introduce *co-moving coordinates* that move with the Hubble flow, i.e. with the background expansion of the dust cloud (or rather: the universe filled with dust). Following chapter 6 of Mukhanov (2005) such coordinates can be defined in terms of the physical coordinates as

$$\mathbf{x}(t, \mathbf{r}) = \frac{1}{a(t)}\mathbf{r} \quad (4.28)$$

where the scale factor $a(t)$ is defined in 4.18. Note that the definition in 4.28 is such that co-moving coordinates \mathbf{x} and physical coordinates \mathbf{r} coincide at time t_0 . The differential operators ∇ and Δ in both coordinate systems are related via

$$\nabla_{\text{co}} = a\nabla_{\text{ph}} , \quad \Delta_{\text{co}} = a^2\Delta_{\text{ph}} . \quad (4.29)$$

Similarly, it is convenient to introduce the so-called *conformal time* η which can, e.g. , be defined as

$$\eta = \int_{t_0}^t dt' \frac{1}{a(t')} , \quad (4.30)$$

such that (cf. Mukhanov 2005)

$$\left. \frac{\partial}{\partial\eta} \right|_{\mathbf{x}=\text{const.}} = a \left\{ \left. \frac{\partial}{\partial t} \right|_{\mathbf{r}=\text{const.}} + \bar{\mathbf{v}} \cdot \nabla_{\text{ph}} \right\} . \quad (4.31)$$

Here, the subscripts $\mathbf{x}, \mathbf{r} = \text{const.}$ indicate that the derivatives should be taken along trajectories of constant \mathbf{x} resp. \mathbf{r} (I will drop those subscripts now and imply this notion unless explicitly stated otherwise).

In the new coordinates (\mathbf{x}, η) the continuity becomes

$$\begin{aligned} 0 &= a \left\{ \frac{\partial \rho}{\partial t} + \nabla_{\text{ph}}(\rho \bar{\mathbf{v}} + \rho \mathbf{v}) \right\} \\ &= a \left\{ \frac{\partial \rho}{\partial t} + \bar{\mathbf{v}} \cdot \nabla_{\text{ph}} \rho + \rho \nabla_{\text{ph}} \bar{\mathbf{v}} + \nabla_{\text{ph}}(\rho \mathbf{v}) \right\} \\ &= \frac{\partial \rho}{\partial \eta} + 3\mathcal{H}\rho + \nabla_{\text{co}}(\rho \mathbf{v}) . \end{aligned} \quad (4.32)$$

Here we have defined $\mathcal{H} \equiv \dot{a} = a'/a$ where $'$ denotes derivation wrt. η . From the fact that $\bar{\rho} \sim a^{-3}$ it can be seen that the background density fulfills $\partial \bar{\rho} / \partial \eta + 3\mathcal{H}\bar{\rho} = 0$. Hence, using the definition of δ in 4.24 the continuity equation can be recast as

$$\frac{\partial \delta}{\partial \eta} + \nabla_{\text{co}}[(1 + \delta)\mathbf{v}] = 0 , \quad (4.33)$$

where we have divided the equation by an overall factor of $\bar{\rho}$.

The Euler equation in the new coordinates reads

$$\begin{aligned} -\nabla_{\text{co}}\phi_{\text{grav}} &= a \left\{ \frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla_{\text{ph}})\mathbf{v} \right\} \\ &= \frac{\partial \mathbf{v}}{\partial \eta} + (\mathbf{v} \cdot \nabla_{\text{co}})\mathbf{v} \\ &= \frac{\partial \mathbf{v}}{\partial \eta} + \mathcal{H}\mathbf{v} + (\mathbf{v} \cdot \nabla_{\text{co}})\mathbf{v} \\ &= \frac{\partial \mathbf{v}}{\partial \eta} + \mathcal{H}\mathbf{v} + (\mathbf{v} \cdot \nabla_{\text{co}})\mathbf{v} + \frac{\partial \mathcal{H}}{\partial \eta} \mathbf{x} , \end{aligned} \quad (4.34)$$

where we have used the fact that $H\mathbf{r} = \mathcal{H}\mathbf{x}$. From 4.13 and the definition of \mathcal{H} it can be shown that $\partial \mathcal{H} / \partial \eta = -4\pi G / 3 \bar{\rho} a^2$ such that

$$\phi_{\text{grav}}(\eta, \mathbf{x}) = -\frac{1}{2} \frac{\partial \mathcal{H}}{\partial \eta} |\mathbf{x}|^2 + \varphi(\eta, \mathbf{x}) . \quad (4.35)$$

Thus the Euler equation becomes

$$\frac{\partial \mathbf{v}}{\partial \eta} + \mathcal{H}\mathbf{v} + (\mathbf{v} \cdot \nabla_{\text{co}})\mathbf{v} = -\nabla_{\text{co}}\varphi . \quad (4.36)$$

In chapter 6 I will summarize how to solve equations 4.33 and 4.36 perturbatively. This perturbation theory is a corner stone of the model for density split statistics that I present in chapter 16.

Chapter 5

Cosmology in General Relativity

This chapter of the thesis is based on the chapters 1 and 7 of Mukhanov (2005). See also chapters 10 and 13 of Padmanabhan (2010a) for a similar presentation.

To describe the cosmic expansion within General Relativity (GR) we need to specify a coordinate system x^μ ($\mu = 0, 1, 2, 3$) and a metric tensor with components $g_{\mu\nu}$ in this coordinate system. The proper time interval $d\tau$ that an observer experiences along an infinitesimal path dx^μ is given by

$$c^2 d\tau^2 = g_{\mu\nu} dx^\mu dx^\nu , \quad (5.1)$$

where here and subsequently we will apply the convention that indices appearing once in subscript and once in superscript are summed over (Einstein's sum convention, cf. chapter 1). The quantity $ds = cd\tau$ is called the *line element*. For the rest of this chapter we will return to the natural unit system in which $c \equiv 1$, such that the line element is equivalent to the infinitesimal proper time interval.

Changing the coordinate system to new coordinates $\tilde{x}^\alpha = \tilde{x}^\alpha(x^\mu)$ changes the components of the metric tensor as

$$\tilde{g}_{\alpha\beta} = g_{\mu\nu} \frac{\partial x^\mu}{\partial \tilde{x}^\alpha} \frac{\partial x^\nu}{\partial \tilde{x}^\beta} . \quad (5.2)$$

In a coordinate system that is freely falling at a point x_0^μ the metric reduces to the usual Minkowsky metric at this location, i.e.

$$g_{\mu\nu}(x_0^\mu) = \eta_{\mu\nu} = (\text{diag}[1, -1, -1, -1])_{\mu\nu} . \quad (5.3)$$

In the following Greek indices will always range from 0 to 3 and denote spacetime indices, while Latin indices will always run from 1 to 3 and denote spatial indices. $g^{\mu\nu}$ denotes the inverse of the metric $g_{\mu\nu}$ and indices of the components of any tensor are lowered and raised by contraction with the components of the metric or inverse metric respectively, e.g.

,

$$A_\mu = g_{\mu\nu} A^\nu \text{ and } A^\mu = g^{\mu\nu} A_\nu \quad (5.4)$$

for the components of covariant and contravariant vectors.

If the universe is filled with "materials" (fluids, fields, particles) whose combined energy-momentum-tensor has components $T_{\mu\nu}$ then according to GR the metric components have to satisfy the *Einstein equations*

$$G_{\mu\nu} = 8\pi T_{\mu\nu}. \quad (5.5)$$

Here $G_{\mu\nu}$ is the so-called *Einstein tensor* which contains derivatives of the metric components up to the second order. We will not introduce any further concepts of GR and instead refer the reader to textbook literature (e.g. Fließbach, 1990).

5.1 Friedmann Universe

This section is based on chapter 2 of Mukhanov (2005).

5.1.1 FLRW-metric and Friedmann equations

Based on the cosmological principle we postulate that the universe is homogeneous and isotropic. Up to arbitrary coordinate transformations, the only metrics satisfying this principle are of the form

$$ds^2 = dt^2 - a(t)h_{ij}(\mathbf{x})x^i x^j \quad (5.6)$$

where we introduced the notation $x^\mu = (t, \mathbf{x})^\mu$ and where h_{ij} is a time independent metric describing a 3-dimensional space of uniform curvature. Such a spatial metric can always be put into the form (cf. chapter 1 of Mukhanov 2005)

$$ds^2 = dt^2 - a^2(t) (\mathrm{d}\chi^2 + \Phi_k^2(\chi) \{ \mathrm{d}\theta^2 + \sin^2(\theta) \mathrm{d}\phi^2 \}), \quad (5.7)$$

where the function $\Phi_k(\chi)$ depends on the curvature of the 3D subspaces and is given by

$$\begin{aligned} \Phi_k(\chi) &= \chi && \text{for zero 3D curvature ("flat" universe)} \\ &= \sin \chi && \text{for positive 3D curvature ("closed" universe)} \\ &= \sinh \chi && \text{for negative 3D curvature ("open" universe).} \end{aligned} \quad (5.8)$$

The meaning of the subscript k will become clear below. A universe with positive 3D curvature is called *closed* because its $t = \text{const.}$ subspaces are topologically closed 3D spheres. For a flat universe the above line element becomes

$$ds^2 = dt^2 - a^2(t) \mathrm{d}\mathbf{x}^2, \quad (5.9)$$

where \mathbf{x} are the Cartesian coordinates of a flat 3-dimensional space and (χ, θ, ϕ) would be a corresponding spherical coordinate system. The scale factor $a(t)$ describes an isotropic expansion of that space and the coordinates \mathbf{x} (or χ, θ, ϕ in the general case) are co-moving with this expansion. At any time t , a distance Δx in the co-moving coordinates corresponds to a physical distance $\Delta l = a\Delta x$. The metric in equation 5.7 is called the Friedmann-Lemaître-Robertson-Walker metric (FLRW metric, but we will sometimes simply call it the Friedmann metric).

The time coordinate t describes exactly the proper time measured by an observer at rest in the above coordinate system. It is however convenient to introduce a different time coordinate

$$\eta(t) = \int_{t_0}^t \frac{dt'}{a(t')} \quad (5.10)$$

with which the line element becomes

$$\begin{aligned} ds^2 &= a^2(t) (d\eta^2 - dx^2 + \Phi_k^2(\chi) \{d\theta^2 + \sin^2(\theta)d\phi^2\}) \\ &(\text{ = } a^2(t) (d\eta^2 - dx^2) \text{ in the flat case. }) \end{aligned} \quad (5.11)$$

This is the conformal form of the FLRW metric and η is the *conformal time* (which was already introduced in the context of a collapsing Newtonian gas cloud in section 4.3.1).

The evolution of the a universe with the FLRW metric is entirely encoded in the scale factor $a(t)$ resp. $a(\eta)$. To determine the time dependence of the scale factor via Einstein's equations we approximate the energy content of the universe as a perfect fluid. The energy-momentum-tensor of such a fluid is given by

$$T_{\mu\nu} = (\epsilon + p)u_\mu u_\nu - pg_{\mu\nu} , \quad (5.12)$$

where ϵ is the energy density of the fluid (in the rest frame of each mass element of the fluid), p is its pressure and u^μ is its four-velocity. To satisfy the cosmological principle the fluid should rest in the expanding coordinates, such that in u^μ in conformal time is given by (see the text around equation 7.20 of Mukhanov, 2005)

$$\begin{aligned} u^0 &\equiv \frac{d\eta}{d\tau} = g^{00} = \frac{1}{a^2} \\ u^i &\equiv \frac{dx^i}{d\tau} = 0 . \end{aligned} \quad (5.13)$$

The Einstein equations in this case simplify to the *Friedmann equations*

$$\begin{aligned} \mathcal{H}' &= -\frac{4\pi}{3}(\epsilon + 3p)a^2 \\ \mathcal{H}^2 + k &= \frac{8\pi}{3}\epsilon a^2 . \end{aligned} \quad (5.14)$$

Here we introduced the conformal expansion rate $\mathcal{H} = a'/a$ and $'$ denotes derivative wrt. η . Furthermore, k/a^2 with $k \in \{-1, 0, 1\}$ is exactly the 3D-curvature of the $\eta \equiv \text{const.}$ subspaces. Using the physical time t as time coordinate the Friedmann equations can be recast as

$$\begin{aligned} \ddot{a} &= -\frac{4\pi}{3}(\epsilon + 3p)a \\ H^2 + \frac{k}{a^2} &= \frac{8\pi}{3}\epsilon , \end{aligned} \quad (5.15)$$

where $H = \dot{a}/a$ and \cdot denotes derivative wrt. t . The first Friedmann equation is similar to equation 4.13 which described the acceleration of a mass element dm , with the difference that in GR the pressure of a fluid also acts as a source of gravity. The second Friedmann equation coincides with Eq. 4.17 which described the expansion rate of a dust cloud in the Newtonian approximation. The parameter k however, which in Newtonian theory was linked to the total energy of the dust cloud, obtains a geometric meaning in GR.

We can see from equation 5.15 that there is a critical energy density

$$\epsilon_{\text{crit}} = \frac{3H^2}{8\pi} \quad (5.16)$$

for which the universe must be flat. Hence, the parameter

$$\Omega = \frac{\epsilon}{\epsilon_{\text{crit}}} \quad (5.17)$$

is a measure for the curvature of the universe. Observations of the present day Hubble flow indicate that the universe is flat (Riess et al., 1998), i.e. $\Omega = 1$. We will hence consider only flat universes and set $k = 0$ for the rest of this chapter.

5.1.2 Thermodynamics

To solve the Friedmann equations we need to know how the pressure p and energy density ϵ evolve under a changing scale factor. To do so, let us consider a fixed volume V_{co} in the co-moving coordinates. The corresponding physical volume at time t is

$$V = a^3 V_{\text{co}} . \quad (5.18)$$

If the universe is filled with a homogeneous fluid with mass density ρ and energy density ϵ then the total energy within the above volume is given by

$$\begin{aligned} E &= M + Q \\ \Rightarrow a^3 \epsilon &= a^3 \rho + Q , \end{aligned} \quad (5.19)$$

where we denote the heat energy contained in V as Q . The first law of thermodynamics tells us that

$$\begin{aligned} \frac{dQ}{dV} &= -p \\ \Rightarrow \frac{dQ}{da} &= -3pa^2 . \end{aligned} \quad (5.20)$$

Here p again denotes the pressure of the fluid. To use equations 5.19 and 5.20 for deriving ϵ and p as functions of a we need two more assumptions. First, we assume that there is a linear relation between the energy density and the pressure of the fluid which can be written as

$$\epsilon = wp . \quad (5.21)$$

For an ideal gas such a relation always exists and is called the *equation of state* of the gas while w is called the equation of state parameter. Our second assumption is that ϵ is a power law in a which we can write in the form

$$\epsilon = \epsilon_0 a^{-\gamma} . \quad (5.22)$$

Equation 5.19 is then equivalent to

$$\begin{aligned} \epsilon_0 a^{3-\gamma} &= \frac{M}{V_{\text{co}}} + Q \\ \Rightarrow (3-\gamma)a^{2-\gamma} &= \frac{dQ}{da} \\ &= -3w a^{2-\gamma} \\ \Rightarrow \gamma &= 3(1+w) , \end{aligned} \quad (5.23)$$

where we have used the fact that M/V_{co} is a constant. I will now derive the equation of state parameter w for three important special cases and analyze how this affects the evolution of the scale factor (following along the lines of Mukhanov, 2005; Padmanabhan, 2010a).

5.1.3 Dust: $\epsilon \sim a^{-3}$, $w = 0$

The energy density ϵ of cold matter is almost exclusively given by the density of its rest mass. For this type of matter ϵ scales as a^{-3} since the physical volume of any fixed region in co-moving coordinates is proportional to a^3 . The density can then be written as

$$\epsilon(t) \approx \epsilon_0 \left(\frac{a_0}{a(t)} \right)^3 , \quad (5.24)$$

where a_0 and ρ_0 are the scale factor and density at some time t_0 . Without loss of generality we can choose t_0 as the present time and then rescale the co-moving coordinates such that $a_0 = 1$. Then we have

$$\epsilon(t) \approx \frac{\epsilon_0}{a(t)^3} . \quad (5.25)$$

Comparing this to equations 5.22 and 5.23 we see that for cold matter $w \approx 0$ and $\gamma \approx 3$. This means that indeed $p \ll \epsilon$ as is expected for cold matter (or *dust*, cf. section 4.1).

Based on equation 5.25 we can now solve the Friedmann equations 5.15 for a universe that is filled purely with dust. Using the power law ansatz $a(t) \sim t^\alpha$ gives

$$a(t) = \left(\frac{t}{t_0} \right)^{\frac{2}{3}} . \quad (5.26)$$

(Remember that we consider only flat universe, i.e. $k = 0$, for the rest of this chapter.)

5.1.4 Radiation: $\epsilon \sim a^{-4}$, $w = \frac{1}{3}$

The energy density of radiation or any gas of ultra-relativistic particle scales as a^{-4} because the density of density of particle in such a gas is still proportional to $\sim a^{-3}$ while the energy of each individual particle is proportional to a^{-1} . The reason for that is that the expansion of the universe stretches the wavelength of each individual particle by a factor of a . Keeping the same conventions as before regarding our choice of t_0 and a_0 this gives

$$\epsilon(t) = \frac{\epsilon_0}{a(t)^4} . \quad (5.27)$$

Comparing this to equations 5.22 and 5.23 we see that for radiation $w = \frac{1}{3}$ and $\gamma = 4$.

Using equation 5.27 we can again solve the Friedmann equations 5.15 and find that in a flat universe filled with radiation the scale factor evolves as

$$a(t) = \left(\frac{t}{t_0} \right)^{\frac{1}{2}} . \quad (5.28)$$

5.1.5 Vacuum energy density: $\epsilon \equiv \text{const.}$, $w = -1$

If the universe is filled with a constant vacuum energy density then obviously

$$\epsilon(t) \equiv \epsilon_0 . \quad (5.29)$$

This means that $\gamma = 0$ and $w = -1$. For such an equation of state a power law ansatz $a(t) \sim t^\alpha$ does not solve the Friedmann equations any more. Instead, we have to use the exponential ansatz $a(t) \sim e^{\alpha t}$ to derive

$$a(t) = e^{H_\Lambda(t-t_0)} . \quad (5.30)$$

Here H_Λ is equal to the expansion rate $H(t)$ which is now independent of time. The symbol Λ is usually denotes the so-called cosmological constant in Einstein's equations. It is used here, because a non-zero, positive value of that constant would influence the expansion of the universe in the same way as a constant vacuum energy density (cf. chapters 2 and 3).

5.1.6 The Λ CDM model

The energy content of the universe can be approximated as a mixture of cold matter, radiation and dark energy (where the latter can be caused either by a positive cosmological or by a non-vanishing energy density of the vacuum). The total energy density of the universe can then be decomposed as

$$\epsilon_{\text{tot}}(t) = \epsilon_m(t) + \epsilon_r(t) + \epsilon_\Lambda(t) , \quad (5.31)$$

where the subscript m stands for matter, the subscript r stands for radiation and Λ denotes dark energy. Inserting the dependence of the different energy density contributions on the

scale factor and denoting today's expansion rate of the universe by H_0 this can be written as

$$\begin{aligned}\epsilon_{\text{tot}}(t) &= \frac{\epsilon_{m,0}}{a(t)^3} + \frac{\epsilon_{r,0}}{a(t)^4} + \epsilon_{\Lambda,0} \\ &= \frac{3H_0^2}{8\pi} \left(\frac{\Omega_{m,0}}{a(t)^3} + \frac{\Omega_{r,0}}{a(t)^4} + \Omega_{\Lambda,0} \right) .\end{aligned}\quad (5.32)$$

Here Ω_X is the ratio of the energy density of component X to the critical energy density (cf. equation 5.16) and $\Omega_{X,0}$ is the value of that ratio today. Since we are considering a flat universe we have $\Omega_{\text{tot}} = 1$ at all times.

Using the different equation of state parameter of the different energy components we can decompose the total pressure in a similar way:

$$\begin{aligned}p_{\text{tot}}(t) &= \frac{1}{3} \frac{\epsilon_{r,0}}{a(t)^4} - \epsilon_{\Lambda,0} \\ &= \frac{3H_0^2}{8\pi} \left(\frac{1}{3} \frac{\Omega_{r,0}}{a(t)^4} - \Omega_{\Lambda,0} \right) .\end{aligned}\quad (5.33)$$

Using equations 5.32 and 5.33 we can rewrite the Friedmann equations for a flat universe as

$$\begin{aligned}\ddot{a} &= -\frac{H_0^2}{2} \left(\frac{\Omega_{m,0}}{a^2} + 2 \frac{\Omega_{r,0}}{a^3} - 2a\Omega_{\Lambda,0} \right) \\ H^2 &= H_0^2 \left(\frac{\Omega_{m,0}}{a^3} + \frac{\Omega_{r,0}}{a^4} + \Omega_{\Lambda,0} \right) .\end{aligned}\quad (5.34)$$

Given the parameters $\Omega_{m,0}$, $\Omega_{r,0}$, $\Omega_{\Lambda,0}$ and H_0 and initial conditions for a and H these equations can be integrated numerically. The radiation fraction in the total energy budget of the universe today is negligible so that the expansion of the universe today is influenced mainly by the dark energy (Λ) and the cold dark matter component (CDM, cf. chapters 2 and 3). Hence, the above model for the expansion history of the universe is called the Λ CDM model.

5.2 The inhomogeneous universe

In the last sections we have studied a perfectly homogeneous and isotropic universe. However, the observed universe is homogeneous and isotropic only on very large scales (~ 100 Mpc). To obtain a more realistic description of the universe we have to consider perturbations to the perfect Friedmann metric. For simplicity, we will do this only for the spatially flat Friedmann universe.

In cosmology the most important perturbations are the so-called scalar perturbations, since they inhibit gravitational instability and can grow during cosmic evolution. As shown, e.g. , in chapter 7 of Mukhanov (2005) the line element with these perturbations becomes

$$ds^2 = a(\eta)^2 \left\{ (1 + 2\phi)d\eta^2 - (1 - 2\psi)d\mathbf{x}^2 \right\} , \quad (5.35)$$

where ϕ and ψ are two scalar functions. Equation 5.35 uses the so-called conformal Newtonian gauge to fix fictitious degrees of freedom (cf. Mukhanov 2005). In the following, we will derive the Einstein equations for the perturbations ϕ and ψ (this can also be found in chapter 7 of Mukhanov 2005). The purpose of this section is to consistently derive the equations of motions of the Newtonian approximation in a non-relativistic limit from these Einstein equations. Especially, I want to demonstrate why some terms that are of 2nd order in perturbations of the velocity field can be neglected while other such terms can't. This is important, because we need terms of second order in the velocity to develop non-linear perturbation theory in chapter 6. I did not find references that do this derivation as consistent as is presented here.

5.2.1 The Einstein tensor with perturbations

We want to express the Einstein tensor corresponding to the metric 5.35 as

$$G_\mu^\nu = \bar{G}_\mu^\nu + \delta G_\mu^\nu , \quad (5.36)$$

where \bar{G}_μ^ν is the Einstein tensor of a perfectly homogeneous Friedmann universe and δG_μ^ν is a perturbation to that background universe. As shown in appendix A, at linear order in ψ and ϕ these perturbations are given by

$$\begin{aligned} a^2 \delta G_0^0 &= -6\mathcal{H}\psi' - 6\mathcal{H}^2\phi + 2\Delta\psi \\ a^2 \delta G_0^i &= -2\partial_j(\psi' + \mathcal{H}\phi) \\ a^2 \delta G_i^j &= -\delta_{ij} [2(\mathcal{H}^2 + 2\mathcal{H}')\phi + 2\psi'' + 2\mathcal{H}(\phi' + 2\psi') + \Delta(\phi - \psi)] \\ &\quad - \partial_i\partial_j(\psi - \phi) . \end{aligned} \quad (5.37)$$

On cosmological scales $\psi \ll 1$ and $\phi \ll 1$ is always fulfilled. In Fourier space we can see that this also holds for the spatial derivatives of the perturbations since $\partial/\partial x^i \sim k_i \sim 1/L$ where L is the typical size of cosmological density fluctuations (which is $\gg 1$ in natural units). Also, $\partial/\partial\eta \sim \omega \sim aH_0$, where the age of the universe was approximated by $1/H_0 \gg 1$. As a consequence, the linear approximation of the Einstein tensor is an excellent approximation in the regime relevant for this work.

5.2.2 Perturbations in T_μ^ν for cold matter

The energy-momentum-tensor of the (cold) matter component is given by

$$T_\mu^\nu = \rho_m u_\mu u^\nu , \quad (5.38)$$

where ρ_m is the mass density of matter in its local rest frame and $u^\mu = dx^\mu/ds$ is the 4-velocity field of the matter component. We would like to express this tensor in the form

$$T_\mu^\nu = \bar{T}_\mu^\nu + \delta T_\mu^\nu , \quad (5.39)$$

where \bar{T}_μ^ν is the energy momentum tensor of a perfectly homogeneous Friedmann universe and δT_μ^ν are perturbations to that. In order to do so, let us write ρ_m as

$$\rho_m = \bar{\rho}_m(1 + \delta_m) . \quad (5.40)$$

Here $\bar{\rho}_m \sim 1/a^3$ is the mean density of the universe and δ_m is the density contrast (which we already encountered in the Newtonian treatment in section 4.3). Also, let $v^\mu = dx^\mu/d\eta$ denote the coordinate velocity of individual fluid elements (where the word *fluid* refers to the cold matter component) and let $\mathbf{v} = dx/d\eta$ be the spatial part of v^μ . With this definition the contravariant components of u are given by

$$u^\mu = \frac{dx^\mu}{d\tau} = \frac{d\eta}{d\tau} \frac{dx^\mu}{d\eta} = u^0 v^\mu = \frac{v^\mu}{a} \frac{1}{\sqrt{[1+2\phi] - [1-2\psi]\mathbf{v}^2}} , \quad (5.41)$$

while the covariant components are given by

$$\begin{aligned} u_0 &= \frac{v^0 g_{00}}{a} \frac{1}{\sqrt{[1+2\phi] - [1-2\psi]\mathbf{v}^2}} = \frac{a(1+2\phi)}{\sqrt{[1+2\phi] - [1-2\psi]\mathbf{v}^2}} \\ u_i &= \frac{v^i g_{ii}}{a} \frac{1}{\sqrt{[1+2\phi] - [1-2\psi]\mathbf{v}^2}} = v^i \frac{-a(1-2\psi)}{\sqrt{[1+2\phi] - [1-2\psi]\mathbf{v}^2}} . \end{aligned} \quad (5.42)$$

(Note that in the last line I temporarily ignored the Einstein sum convention.)

The perturbations of the energy momentum tensor are caused by ϕ , ψ , δ_m and \mathbf{v} . As argued above, on the scales of interest for this work the perturbations ϕ and ψ are always small and we only need to consider them in up to linear order. The amplitudes of the perturbations δ and \mathbf{v} behave more complicated. For now, I will treat \mathbf{v}^2 to have the same order of magnitude as the metric perturbations, which at linear order gives

$$\begin{aligned} \delta T_0^0 &\approx \bar{\rho}_m [\mathbf{v}^2 + \delta_m(1 + \mathbf{v}^2)] \\ \delta T_0^i &\approx \bar{\rho}_m (1 + \delta_m) v^i \\ \delta T_i^j &\approx \bar{\rho}_m (1 + \delta_m) v^i v^j . \end{aligned} \quad (5.43)$$

Now the equations of motion for the perturbations can be obtained from the Einstein equations. Since by definition the background Einstein tensor and background energy momentum tensor satisfy their own Einstein equations, this leads to

$$\delta G_\mu^\nu = 8\pi \delta T_\mu^\nu . \quad (5.44)$$

I will investigate these equations in detail in section 5.2.3, but first we need to discuss the significance of terms $\sim \mathcal{O}(\mathbf{v}^2)$.

Magnitude of \mathbf{v} , \mathbf{v}^2 and δ

As we will see in chapter 6, we need to keep terms of the order $\sim \delta_m v^i$ or $\sim v^i v^j$ in order to describe the evolution of δ_m and \mathbf{v} in non-linear perturbation theory. In fact, terms

of such order were present in the exact equations 4.33 and 4.36 that were derived in the Newtonian limit. Hence, in order to show that these equations follow from the Einstein equations in the appropriate limit, one cannot simply ignore terms that are quadratic in \mathbf{v} . Nevertheless, we can simplify the expressions for the energy momentum tensor obtained in equations 5.43 by arguing as follows:

- At a sufficiently early time, the perturbations δ_m and \mathbf{v} will be small enough to ensure that $\mathbf{v}^2 \ll \delta_m$.
- Non-linear perturbation theory becomes important as soon as $\mathbf{v}^2 \sim \phi, \psi$, i.e. as soon as terms $\sim \mathbf{v}^2$ cannot be ignored in the non-linear equations of motion anymore (see chapter 6 or section 2 of Bernardeau et al., 2002). However, in this regime we also have $\delta_m \sim \mathcal{O}(1)$. Hence, even at late time we can still assume that $\mathbf{v}^2 \ll \delta_m$.

Using this we can see that at any time

$$\delta T_0^0 \approx \bar{\rho}_m \delta_m \quad \text{and} \quad \delta T_i^j \ll \delta T_0^0. \quad (5.45)$$

Note especially, that the above reasoning is not sufficient to set δT_0^i to 0.

5.2.3 The Einstein equations for perturbations and non-relativistic limit

Using our above considerations for the energy momentum tensor we can now write the Einstein equations for the perturbations as (cf. equations 7.38 to 7.40 of Mukhanov 2005)

$$\begin{aligned} 4\pi a^2 \bar{\rho}_m \delta_m &= \Delta\psi - 3\mathcal{H}\psi' - 3\mathcal{H}^2\phi \\ -4\pi a^2 \bar{\rho}_m (1 + \delta_m) v^i &= \partial_i(\psi' + \mathcal{H}\phi) \\ -4\pi a^2 \delta T_i^j &= \delta_{ij} \left[(\mathcal{H}^2 + 2\mathcal{H}')\phi + \psi'' + \mathcal{H}(\phi' + 2\psi') + \frac{\Delta(\phi - \psi)}{2} \right] + \frac{\partial_i \partial_j (\psi - \phi)}{2}. \end{aligned} \quad (5.46)$$

Small scale, non-relativistic limit

On co-moving scales $L \ll 1/\mathcal{H}$ the term $\Delta\psi$ will dominate the right hand-side of the 00-component of the Einstein equations. These scales are the only relevant ones for this work. Hence we can set

$$4\pi a^2 \bar{\rho}_m \delta_m \approx \Delta\psi. \quad (5.47)$$

From the ij -component of the Einstein equations we can see that

$$\partial_i \partial_j (\psi - \phi) \sim 4\pi a^2 \delta T_i^j \ll 4\pi a^2 \delta T_0^0 \sim \Delta\psi. \quad (5.48)$$

Hence, within the approximations discussed in the end of section 5.2.2 we can set $\Delta\psi \approx \Delta\phi$ which gives

$$4\pi a^2 \bar{\rho}_m \delta_m \approx \Delta\phi. \quad (5.49)$$

This is equivalent to the Poisson equation 4.27 we encountered in the Newtonian approximation. Taking the spatial divergence of the $0i$ -component of the Einstein equations equations we furthermore get

$$\begin{aligned}
-4\pi a^2 \bar{\rho}_m \nabla \cdot ([1 + \delta_m] \mathbf{v}) &= \Delta(\psi' + \mathcal{H}\phi) \\
&\approx \Delta(\phi' + \mathcal{H}\phi) \\
&= \frac{(a\Delta\phi)'}{a} \\
\Rightarrow -a^2 \bar{\rho}_m \nabla \cdot ([1 + \delta_m] \mathbf{v}) &\approx \frac{(a^3 \bar{\rho}_m \delta_m)'}{a} \\
&\approx \bar{\rho}_{m,0} \frac{\delta_m'}{a} \\
\Rightarrow \delta_m' + \nabla \cdot ([1 + \delta_m] \mathbf{v}) &\approx 0 . \tag{5.50}
\end{aligned}$$

Here $\bar{\rho}_{m,0} = a^3 \bar{\rho}_m$ is the average matter density today. The last line of the above equations is exactly the continuity equation that was derived for perturbations in the Newtonian approximation (cf. 4.33).

To complete our recovery of the Newtonian equations derived in chapter 4 we are still missing the Euler equation. Since we are considering only cold matter without pressure, the trajectories of individual fluid elements are geodesics. This means that the 4-velocity field satisfies the geodesic equation

$$u^\alpha \nabla_\alpha u^\mu \equiv u^\alpha \partial_\alpha u^\mu + \Gamma_{\alpha\beta}^\mu u^\alpha u^\beta = 0 , \tag{5.51}$$

where the first equality defines the so-called covariant derivative ∇_α of the components of a vector and $\Gamma_{\alpha\beta}^\mu$ are the Christoffel symbols which are derived in appendix A. Remember that $u^\mu = u^0 v^\mu$ with $v^0 = 1$. Hence, the spatial part of the geodesic equation reads

$$\begin{aligned}
0 &= u^\alpha \nabla_\alpha u^i \\
&= u^\alpha \partial_\alpha u^i + \Gamma_{\alpha\beta}^i u^\alpha u^\beta \\
&= u^\alpha \partial_\alpha (u^0 v^i) + \Gamma_{\alpha\beta}^i u^\alpha u^\beta \\
&= v^i u^\alpha \partial_\alpha u^0 + u^0 u^\alpha \partial_\alpha v^i + \Gamma_{\alpha\beta}^i u^\alpha u^\beta \\
&= v^i u^\alpha (\partial_\alpha u^0 + \Gamma_{\alpha\beta}^0 u^\beta) + u^0 u^\alpha \partial_\alpha v^i + \Gamma_{\alpha\beta}^i u^\alpha u^\beta - v^i \Gamma_{\alpha\beta}^0 u^\alpha u^\beta \\
&= v^i u^\alpha \nabla_\alpha u^0 + u^0 u^\alpha \partial_\alpha v^i + \Gamma_{\alpha\beta}^i u^\alpha u^\beta - v^i \Gamma_{\alpha\beta}^0 u^\alpha u^\beta . \tag{5.52}
\end{aligned}$$

The first term on the right hand-side vanishes because it is proportional to the time component of the geodesic equation. Dividing the overall equation by $(u^0)^2$ we get

$$\begin{aligned}
0 &= v^\alpha \partial_\alpha v^i + \Gamma_{\alpha\beta}^i v^\alpha v^\beta - v^i \Gamma_{\alpha\beta}^0 v^\alpha v^\beta \\
&= (v^i)' + (\mathbf{v} \nabla) v^i + \Gamma_{00}^i + 2\Gamma_{0k}^i v^k + \Gamma_{kl}^i v^k v^l - \Gamma_{00}^0 v^i - 2\Gamma_{0k}^0 v^k v^i - \Gamma_{kl}^0 v^k v^l . \tag{5.53}
\end{aligned}$$

In appendix A I show that (see also Mukhanov, 2005)

$$\begin{aligned}
\Gamma_{00}^i &= \partial_i \phi \\
\Gamma_{0j}^i &= \delta_{ij} (\mathcal{H} - \psi') \\
\Gamma_{ij}^k &= \delta_{ij} \partial_k \psi - \delta_{ki} \partial_j \psi - \delta_{jk} \partial_i \psi \\
\Gamma_{00}^0 &= \mathcal{H} + \phi' \\
\Gamma_{0i}^0 &= \partial_i \phi \\
\Gamma_{ij}^0 &= \delta_{ij} (\mathcal{H} - \psi')
\end{aligned} \tag{5.54}$$

at linear order in ϕ and ψ . Inserting this above and ignoring all terms $\sim \mathcal{O}(\phi \mathbf{v}, \psi \mathbf{v})$ (which are $\sim \mathcal{O}(\mathbf{v}^2)$ in the early universe and $\sim \mathcal{O}(\mathbf{v}^3)$ in the late universe, cf. the discussion at the end of section 5.2.2) we get

$$0 \approx (v^i)' + (\mathbf{v} \nabla) v^i + \partial_i \phi + \mathcal{H} v^i. \tag{5.55}$$

This is exactly the Euler equation 4.36 for the perturbations which was derived in the Newtonian approximation.

To summarize, on co-moving scales $L \ll 1/\mathcal{H}$ and for a universe filled with cold matter we have derived the following equations for the evolution of the perturbations:

$$\begin{aligned}
4\pi a^2 \bar{\rho}_m \delta_m &= \Delta \phi \\
\delta_m' + \nabla([1 + \delta_m] \mathbf{v}) &= 0 \\
\mathbf{v}' + (\mathbf{v} \nabla) \mathbf{v} + \nabla \phi + \mathcal{H} \mathbf{v} &= 0.
\end{aligned} \tag{5.56}$$

These are exactly the equations we derived within the Newtonian frame work. In chapter 6 I will use these to build up a perturbative scheme to calculate the non-linear evolution of the perturbations δ_m and \mathbf{v} . Since I have only considered scalar perturbations, only the scalar part of \mathbf{v} will play a role in this expansion (which is, e.g. , given by the divergence $\theta \equiv \nabla \cdot \mathbf{v}$).

Chapter 6

Theory of cosmic structure formation

As discussed in sections 4.3 and 5.2 the evolution of perturbations in a background Friedmann universe filled with dust (i.e. pressure less matter) is described by the matter density contrast δ_m , the co-moving velocity perturbations \mathbf{v} and the scalar metric perturbation ϕ . The equations of motion for these perturbations are the Poisson equation,

$$4\pi a^2 \bar{\rho}_m \delta_m = \Delta \phi , \quad (6.1)$$

the continuity equation,

$$\delta'_m + \nabla \cdot ([1 + \delta_m] \mathbf{v}) = 0 , \quad (6.2)$$

and the Euler equation,

$$\mathbf{v}' + (\mathbf{v} \nabla) \mathbf{v} + \nabla \phi + \mathcal{H} \mathbf{v} = 0 . \quad (6.3)$$

Here $\bar{\rho}_m$ is the mean matter density in the universe at any given time, $'$ denotes derivation wrt. the conformal time η and $\mathcal{H} = a'/a$ is the conformal expansion rate.

In this chapter I will analyze different exact and approximate solutions to these equations. In section 6.1 I describe the initial conditions of the perturbation as predicted by the theory of inflation and in section 6.2 I solve the linearized version of the equations of motion. In section 6.4 I present the standard scheme to solve the non-linear equations of motion perturbatively. And in section 6.3 I explore exact solutions of the equations of motion for a set of highly symmetric initial conditions. As we will see in chapter 7, these symmetric collapse solutions are intimately related to the leading order of the perturbative solutions. The derivations in this chapter follow closely to the review of Bernardeau et al. (2002) and the book of Mukhanov (2005).

6.1 Initial conditions

Consider a field $f(\mathbf{x})$ that is a realization of a random process (such a random process could, e.g. , be the generation of density fluctuations $\delta_m(\mathbf{x})$ from quantum fluctuations of the inflaton field). f is called a *Gaussian random field* if for any number of locations $\mathbf{x}_1, \dots, \mathbf{x}_N$

the joint probability density function (joint PDF) of the values $[f(\mathbf{x}_1), \dots, f(\mathbf{x}_N)]$ is of the form (see, e.g., appendix A of Hilbert et al. 2011)

$$p(f(\mathbf{x}_1), \dots, f(\mathbf{x}_N)) = \frac{1}{\sqrt{(2\pi)^N |\mathbf{C}|}} \exp \left(-\frac{1}{2} \sum_{i,j=1}^N [f(\mathbf{x}_i) - \bar{f}(\mathbf{x}_i)] [f(\mathbf{x}_j) - \bar{f}(\mathbf{x}_j)] (C^{-1})_{ij} \right). \quad (6.4)$$

Here $\bar{f}(\mathbf{x}) = \langle f(\mathbf{x}) \rangle$ and the elements of the matrix \mathbf{C} are given by the covariance $(C)_{ij} \equiv \xi_f(\mathbf{x}_i, \mathbf{x}_j) = \langle [f(\mathbf{x}_i) - \bar{f}(\mathbf{x}_i)] [f(\mathbf{x}_j) - \bar{f}(\mathbf{x}_j)] \rangle$, where ξ_f is the 2-point correlation function of f . Equation 6.4 is simply the definition of a multivariate Gaussian PDF. A Gaussian random field f is said to be *homogeneous* if $\bar{f}(\mathbf{x}) \equiv \bar{f}$ does not depend on \mathbf{x} and if $\xi_f(\mathbf{x}_i, \mathbf{x}_j)$ only depends on $\mathbf{x}_i - \mathbf{x}_j$. If furthermore $\xi_f(\mathbf{x}_i, \mathbf{x}_j)$ only depends on the absolute value $|\mathbf{x}_i - \mathbf{x}_j|$ then f is said to be *isotropic*.

At any time η the fields $\delta_m(\eta, \mathbf{x})$, $\phi(\eta, \mathbf{x})$ and $\vartheta(\eta, \mathbf{x}) \equiv \nabla \mathbf{v}(\eta, \mathbf{x})$ can be considered random fields with a mean value of $\bar{\delta}_m = \bar{\phi} = \bar{\vartheta} = 0$. The 2-point correlation functions of these field are then given by

$$\xi_\phi(\Delta \mathbf{x}, \eta) \equiv \langle \phi(\eta, \mathbf{x}) \phi(\eta, \mathbf{x} + \Delta \mathbf{x}) \rangle \quad (6.5)$$

$$\xi_\delta(\Delta \mathbf{x}, \eta) \equiv \langle \delta(\eta, \mathbf{x}) \delta(\eta, \mathbf{x} + \Delta \mathbf{x}) \rangle \quad (6.6)$$

$$\xi_\vartheta(\Delta \mathbf{x}, \eta) \equiv \langle \vartheta(\eta, \mathbf{x}) \vartheta(\eta, \mathbf{x} + \Delta \mathbf{x}) \rangle, \quad (6.7)$$

where the expectation values $\langle \cdot \rangle$ in these definitions can either be thought of as ensemble averages over different realizations of the initial conditions or as averages over all locations \mathbf{x} in the universe. The power spectra corresponding to these correlation functions are defined as

$$P_{\phi, \delta, \vartheta}(\mathbf{k}, \eta) = \int d^2x e^{-i\mathbf{k}\mathbf{x}} \xi_{\phi, \delta, \vartheta}(\mathbf{x}, \eta). \quad (6.8)$$

Let η_i be a time shortly after the end of inflation. Then the theory of inflation predicts the following initial conditions for ϕ , δ_m and ϑ (see Bernardeau et al. 2002; Mukhanov 2005; these predictions are also quite insensitive to the particular model of inflation that is assumed, see Mukhanov 2013):

- The fields $\delta_m(\eta_i, \mathbf{x})$, $\phi(\eta_i, \mathbf{x})$ and $\vartheta(\eta_i, \mathbf{x})$ are homogeneous, isotropic Gaussian random fields. (For $\delta_m(\eta_i, \mathbf{x})$ this cannot exactly be true because of the condition $\delta_m \geq -1$. But at time η_i the typical fluctuations in δ_m are $\ll 1$ and $\delta_m(\eta_i, \mathbf{x})$ is described by a Gaussian random field to a high accuracy).
- The power spectrum of the scalar metric perturbation $\phi(\eta_i, \mathbf{x})$ has the form

$$P_\phi(\mathbf{k}, \eta_i) = A_s k^{n_s - 4}, \quad (6.9)$$

where A_s is an (unpredicted) amplitude and n_s is close to but smaller than 1. The variance of $\phi(\mathbf{k}, \eta_i)$ when averaged over volumes of the size $V \sim (1/k)^3$ is approximately given by $k^3 P_\phi(\mathbf{k}, \eta_i)$. Hence, equation 6.9 can be written in the form

$$(\text{Var}[\phi_V] \approx) k^3 P_\phi(\mathbf{k}, \eta_i) = A_s k^{n_s - 1}. \quad (6.10)$$

Since n_s is close to 1 this equation means that the fluctuation amplitude of $\phi(\mathbf{k}, \eta_i)$ when averaged over different volumes is *scale independent* (i.e. independent of the size of these volumes).

The scalar potential ϕ and the matter density contrast δ_m are related through the Poisson equation (6.1). In Fourier space this equation reads

$$4\pi a^2 \bar{\rho}_m \tilde{\delta}_m(\mathbf{k}) = -k^2 \tilde{\phi}(\mathbf{k}) . \quad (6.11)$$

Equation 6.9 then tells us that shortly after the end of inflation the power spectrum of δ_m has had the shape

$$P_\delta(\mathbf{k}, \eta_i) = A_\delta k^{n_s} , \quad (6.12)$$

where the amplitude A_δ is still unpredicted and has to be included as a free parameter into analyses of measurement of the matter power spectrum.

In this chapter I am only considering a universe filled with pressure less matter. This is however not an accurate description for the radiation dominated phase of the universe which starts at the end of inflation η_i and lasts roughly until the time of recombination η_{rec} . And even though this phase is short compared to the matter dominated era it has a significant impact on the shape of the matter density power spectrum P_δ . Fluctuation in the density of the radiation plasma evolve as sound waves during the radiation dominated phase (cf. Mukhanov, 2005; Padmanabhan, 2010a). This leads to a rapidly oscillating amplitude of the gravitational potential which suppresses the growth of matter density fluctuations at small scales (even though at least the dark matter is not directly coupled to the radiation). As a consequence the matter power spectrum at η_{rec} will have a shape that is significantly different from equation 6.12 (see, e.g. , Eisenstein & Hu 1998 for an approximate treatment of this effect or Lewis et al. 2000; Refregier et al. 2017 for an exact numerical treatment).

6.2 Linear evolution

As a first step towards solving the equations of motion for the fields δ_m , ϕ and \mathbf{v} (resp. ϑ) let us consider δ_m and \mathbf{v} to be small enough, such that we can neglect all terms in the Euler and continuity equation that are of second order in the perturbations. This way the continuity equation becomes

$$\frac{\partial \delta_m}{\partial \eta} + \vartheta = 0 \quad (6.13)$$

and the Euler equation becomes

$$\frac{\partial \mathbf{v}}{\partial \eta} + \mathcal{H} \mathbf{v} = -\nabla_{\text{co}} \varphi . \quad (6.14)$$

Taking the divergence on both sides, the latter equation becomes an equation involving only δ and ϑ ,

$$\frac{\partial \vartheta}{\partial \eta} + \mathcal{H} \vartheta + 4\pi G \bar{\rho}_m \delta_m = 0 . \quad (6.15)$$

Inserting the linearized continuity equation we get an equation of δ_m alone as

$$\frac{\partial^2 \delta_m}{\partial \eta^2} + \mathcal{H} \frac{\partial \delta_m}{\partial \eta} - 4\pi G \bar{\rho}_m \delta_m = 0 . \quad (6.16)$$

This equation has the solution

$$\delta_{L,m}(\eta, \mathbf{x}) = \frac{D(\eta)}{D(\eta_i)} \delta_m(\eta_i, \mathbf{x}) \quad (6.17)$$

where D solves equation 6.16 and is called the *linear growth factor* and where the subscript L stands for *linear*.

The theory of inflation only predicts the statistical properties of the initial density contrast $\delta_m(\eta_i, \mathbf{x})$ and the other perturbations but not their exact configuration. To confront these predictions with cosmological observations of the large scale structure of the universe (cf. part II) we hence have to study statistical probes of the cosmic fields, such as their power spectrum or 2-point correlation function. Since both the power spectrum P_δ and the 2-point correlation function ξ_δ are second order statistics of δ_m , their evolution in the linear approximation is given by

$$P_{L,\delta}(\mathbf{k}, \eta) \equiv \left(\frac{D(\eta)}{D(\eta_i)} \right)^2 P_\delta(\mathbf{k}, \eta_i) \quad (6.18)$$

$$\xi_{L,\delta}(\Delta \mathbf{x}, \eta) \equiv \left(\frac{D(\eta)}{D(\eta_i)} \right)^2 \xi_\delta(\Delta \mathbf{x}, \eta_i) . \quad (6.19)$$

The linearized power spectrum of ϑ is then given in terms of P_L as

$$P_{L,\vartheta}(\mathbf{k}, \eta) = \left(\frac{\partial \ln D(\eta)}{\partial \eta} \right)^2 P_{L,\delta}(\mathbf{k}, \eta) \quad (6.20)$$

as can be seen from the linearized continuity equation 6.13. The factor $\partial \ln D(\eta) / \partial \eta$ is also called the *linear growth rate* of structures.

Equation 6.16 can be solved numerically in a general Λ CDM universe. Since it is a second order differential equation it requires 2 boundary conditions in order to sort out a specific solution. The amplitude of $D(\eta)$ is completely degenerate with the unknown amplitude A_δ of the initial power spectrum. We can hence absorb our freedom of choice in the normalization of D completely into the free parameter A_δ and set

$$D(\eta_0) = 1 , \quad (6.21)$$

where η_0 is the value of the conformal time today. The second boundary condition can, e.g., be given by fixing the value of the time derivative D' at some moment in time. An easier way is however to note that equation 6.16 has a growing solution and a decaying solution (Bernardeau et al., 2002). Hence, if we start the numerical integration of the equation at a very early time then for any initial conditions the decaying mode will be negligible at the

late (and physically relevant) times of the universe such that we effectively only need to specify one initial condition (i.e. the amplitude). This procedure is, e.g. , adopted in the code **nicaea** by Kilbinger et al. (2009) (see also <https://www.cosmostat.org/software/nicaea>). In this work I take a different approach: At the time of recombination η_{rec} the universe was matter dominated (as opposed to dark energy dominated) and the growing mode solution in a matter dominated universe is known analytically (Bernardeau et al., 2002). Hence I set $D'(\eta_{\text{rec}})$ to match this analytic solution. (Since η_{rec} is still at a very early stage of structure formation, this approach is however almost identical to the one used in **nicaea**.)

To complete our analysis of the linear evolution of the matter density power spectrum we still need to incorporate its initial amplitude as a free parameter into the Λ CDM model. This can directly be done using the parameter A_δ introduced in the previous section. It is however customary to use a different approach. Let us first apply a smoothing to the linear density contrast $\delta_{L,m}$ with a spherical aperture of radius R , i.e. we define

$$\delta_{L,R}(\eta, \mathbf{x}) \equiv \frac{3}{4\pi R^3} \int_{|\mathbf{y}-\mathbf{x}| < R} d^3y \delta_{L,m}(\eta, \mathbf{y}) . \quad (6.22)$$

The variance $\sigma_{L,R}^2(\eta) \equiv \langle \delta_{L,R}(\eta, \mathbf{x})^2 \rangle$ is given by

$$\sigma_{L,R}^2(\eta) = \left(\frac{3}{4\pi R^3} \right)^2 \int_{|\mathbf{x}| < R} d^3x \int_{|\mathbf{y}| < R} d^3y \xi_{L,\delta}(\mathbf{y} - \mathbf{x}, \eta) \quad (6.23)$$

or in a simpler way through the power spectrum by

$$\sigma_{L,R}^2(\eta) = \int \frac{d^3k}{(2\pi)^3} \hat{W}_R^2(\mathbf{k}) P_{L,\delta}(\mathbf{k}, \eta) , \quad (6.24)$$

where

$$\hat{W}_R(\mathbf{k}) = \frac{3}{(R|\mathbf{k}|)^3} \{ R|\mathbf{k}| \cos(R|\mathbf{k}|) - \sin(R|\mathbf{k}|) \} \quad (6.25)$$

is the Fourier transform of the aperture used for the smoothing. It is now common to introduce the parameter

$$\sigma_8 \equiv \sigma_{L,8\text{Mpc}}^2(\eta_0) , \quad (6.26)$$

i.e. σ_8^2 is the variance of the linear density contrast today within spheres of 8Mpc radius. Specifying the value of σ_8 completely fixes the initial power spectrum amplitude A_δ . Hence, σ_8 can be used to parametrize our ignorance of the amplitude of the initial density fluctuations. This is frequently done in studies of the late time large scale structure of the universe and I am also employing this parametrization in this work. (Note that the specific choice of $R = 8\text{Mpc}$ is simply for convenience, since σ_8 turns out to be of $\mathcal{O}(1)$ in our universe.)

6.3 Lagrangian coordinates and symmetric initial conditions

As a first step beyond the linear approximation we want to study the evolution of perturbations that exhibit a spherical, cylindrical or planar symmetry. As we will see in chapter 7, these symmetric collapse solutions are closely related to the leading order of the perturbative solutions that will be introduced in the next section (see also Bernardeau, 1994; Fosalba & Gaztanaga, 1998; Valageas, 2002a).

To study these symmetric solutions it is convenient to change spatial coordinates from the co-moving coordinates \mathbf{x} to the so-called *Lagrangian coordinates* \mathbf{q} that are just labeling the initial positions of all matter elements. The time derivative in at fixed co-moving and fixed Lagrangian coordinates are related by

$$\frac{\partial}{\partial \eta} \bigg|_{\mathbf{q}=\text{const.}} = \frac{\partial}{\partial \eta} \bigg|_{\mathbf{x}=\text{const.}} + \mathbf{v} \cdot \nabla_{\mathbf{x}} . \quad (6.27)$$

Note that there is potential for confusion here: The co-moving coordinates \mathbf{x} are already 'Lagrangian' in the sense that they follow the background Hubble flow (cf. section 4.3.1). The coordinates \mathbf{q} are now following the complete, disturbed flow \mathbf{v} . In the new coordinates the continuity equation 6.2 becomes

$$\begin{aligned} 0 &= (1 + \delta_m) \vartheta + \frac{\partial \delta_m}{\partial \eta} \bigg|_{\mathbf{q}=\text{const.}} \\ \Rightarrow \vartheta &= -\frac{1}{1 + \delta_m} \frac{\partial \delta_m}{\partial \eta} \bigg|_{\mathbf{q}=\text{const.}} , \end{aligned} \quad (6.28)$$

where $\vartheta = \nabla_{\mathbf{x}} \cdot \mathbf{v}$ is still the divergence of the velocity field in the old coordinates \mathbf{x} .

6.3.1 Spherical collapse

We will start by studying spherically symmetric perturbations around $\mathbf{q} = \mathbf{x} = 0$. This means that at fixed time η both δ_m and ϑ are functions of $|\mathbf{q}|$ resp. $|\mathbf{x}|$ only. Let us consider the spherical average $\delta_{R,\text{Lag.}}$ of δ_m within a fixed Lagrangian radius R around $\mathbf{q} = 0$, i.e.

$$\delta_{R,\text{Lag.}}(\eta) \equiv \frac{3}{4\pi R^3} \int_{|\mathbf{q}| < R} d^3q \delta_m(\eta, \mathbf{q}) . \quad (6.29)$$

Since the density field is perfectly isotropic we know that the evolution of $\delta_{R,\text{Lag.}}(\eta, \mathbf{q})$ only depends on the total mass enclosed in the radius R and not in the exact profile of $\delta_m(\eta, \mathbf{q})$. Hence, we can without loss of generality assume that

$$\delta_m(\eta, \mathbf{q}) \equiv \delta_{R,\text{Lag.}}(\eta) \quad \text{for } |\mathbf{q}| < R . \quad (6.30)$$

From the continuity equation in Lagrangian coordinates 6.28 we can then conclude that also $\vartheta(\eta, \mathbf{q})$ is spatially constant within the Lagrangian radius R . As a consequence we must have

$$\mathbf{v}(\eta, \mathbf{q}) = \frac{\vartheta(\eta, \mathbf{0})}{3} \mathbf{x}(\mathbf{q}) \quad \text{for } |\mathbf{q}| < R , \quad (6.31)$$

where the coordinates \mathbf{x} have been expressed as functions of the coordinates \mathbf{q} . For the spatial derivatives of \mathbf{v} this means

$$\frac{\partial v^i}{\partial x^j} = \frac{\vartheta}{3} \delta_i^j \quad \text{for } |\mathbf{q}(\mathbf{x})| < R . \quad (6.32)$$

We can use this to simplify the Euler equation. Taking the divergence of 6.3 wrt. the old coordinates \mathbf{x} and using also the Poisson equation 6.1 one can show that the Euler equation becomes

$$\begin{aligned} 0 &= \frac{\partial \vartheta}{\partial \eta} \Big|_{\mathbf{x}=\text{const.}} + (\mathbf{v} \nabla_{\mathbf{x}}) \vartheta + \sum_{i,j} \frac{\partial v^i}{\partial x^j} \frac{\partial v^j}{\partial x^i} + \mathcal{H} \vartheta + 4\pi a^2 \bar{\rho}_m \delta_m \\ &= \frac{\partial \vartheta}{\partial \eta} \Big|_{\mathbf{q}=\text{const.}} + \sum_{i,j} \frac{\partial v^i}{\partial x^j} \frac{\partial v^j}{\partial x^i} + \mathcal{H} \vartheta + 4\pi a^2 \bar{\rho}_m \delta_m \\ &= \frac{\partial \vartheta}{\partial \eta} \Big|_{\mathbf{q}=\text{const.}} + \frac{\vartheta^2}{3} + \mathcal{H} \vartheta + 4\pi a^2 \bar{\rho}_m \delta_m \quad \text{for } |\mathbf{q}| < R . \end{aligned} \quad (6.33)$$

Inserting the continuity equation 6.28 for ϑ this becomes an equation for δ_m only:

$$\frac{\partial^2 \delta_m}{\partial \eta^2} + \mathcal{H} \frac{\partial \delta_m}{\partial \eta} - \frac{4}{3(1 + \delta_m)} \left(\frac{\partial \delta_m}{\partial \eta} \right)^2 + 4\pi a^2 \bar{\rho}_m \delta_m (1 + \delta_m) = 0 \quad \text{for } |\mathbf{q}| < R , \quad (6.34)$$

where I have now dropped the indicator $|_{\mathbf{q}=\text{const.}}$ and assume that all time derivatives are taken in Lagrangian coordinates. Since we assumed $\delta_m \equiv \delta_{R,\text{Lag.}}$ within $|\mathbf{q}| < R$ this becomes a differential equation for $\delta_{R,\text{Lag.}}$:

$$\frac{\partial^2 \delta_{R,\text{Lag.}}}{\partial \eta^2} + \mathcal{H} \frac{\partial \delta_{R,\text{Lag.}}}{\partial \eta} - \frac{4}{3(1 + \delta_{R,\text{Lag.}})} \left(\frac{\partial \delta_{R,\text{Lag.}}}{\partial \eta} \right)^2 + 4\pi a^2 \bar{\rho}_m \delta_{R,\text{Lag.}} (1 + \delta_{R,\text{Lag.}}) = 0 . \quad (6.35)$$

Especially, as explained above, this equation is independent of the exact profile of δ_m and does require the assumption $\delta_m \equiv \delta_{R,\text{Lag.}}$ within $|\mathbf{q}| < R$ anymore. In the non-relativistic limit equation 6.35 hence gives an exact description of *spherical collapse* (at least in the absence of shell crossing, cf. Bernardeau et al., 2002). Alternative derivations of it can, e.g. , be found in Fosalba & Gaztanaga (1998); Mukhanov (2005). Given the parameters describing a Λ CDM background universe it is straight forward to integrate this equation numerically.

6.3.2 Cylindrical and planar collapse

The above results can be generalized to the collapse of cylindrically symmetric perturbations (Bernardeau & Valageas, 2000; Uhlemann et al., 2018b). If the density field is cylindrically symmetric around the x^3 axis then within a fixed Lagrangian radius R around this axis we can assume

$$\delta_m(\eta, \mathbf{q}) \equiv \delta_{R,\text{Lag.}}^{\text{2D}}(\eta) \quad \text{for } |(q^1)^2 + (q^2)^2| < R , \quad (6.36)$$

where $\delta_{R,\text{Lag.}}^{\text{2D}}$ is now the average of δ_m over an infinitely long cylinder. Within the cylinder radius we then have

$$\begin{aligned} v^1 &= \frac{\vartheta}{2} x^1(\mathbf{q}) \\ v^2 &= \frac{\vartheta}{2} x^2(\mathbf{q}) \\ v^3 &= 0 . \end{aligned} \quad (6.37)$$

Along lines similar to the spherical case one then shows that

$$\delta_{R,\text{Lag.}}^{\text{2D}}'' + \mathcal{H}\delta_{R,\text{Lag.}}^{\text{2D}'} - \frac{3}{2} \frac{(\delta_{R,\text{Lag.}}^{\text{2D}}')^2}{(1 + \delta_{R,\text{Lag.}}^{\text{2D}})} + 4\pi a^2 \bar{\rho}_m \delta_{R,\text{Lag.}}^{\text{2D}} (1 + \delta_{R,\text{Lag.}}^{\text{2D}}) = 0 , \quad (6.38)$$

where ' denotes derivative wrt. conformal time.

Similarly, if the density field and the velocity field have a planar symmetry (e.g. perpendicular to the $x^2 - x^3$ -axis) then the density contrast in a fixed Lagrangian range around the plane of symmetry evolves as

$$\delta_{R,\text{Lag.}}^{\text{1D}}'' + \mathcal{H}\delta_{R,\text{Lag.}}^{\text{1D}'} - 2 \frac{(\delta_{R,\text{Lag.}}^{\text{1D}}')^2}{(1 + \delta_{R,\text{Lag.}}^{\text{1D}})} + 4\pi a^2 \bar{\rho}_m \delta_{R,\text{Lag.}}^{\text{1D}} (1 + \delta_{R,\text{Lag.}}^{\text{1D}}) = 0 . \quad (6.39)$$

(See, e.g. , Mukhanov 2005 for a derivation of this case.) More generally we can conclude that

$$\delta_{R,\text{Lag.}}^{\text{ND}}'' + \mathcal{H}\delta_{R,\text{Lag.}}^{\text{ND}'} - \frac{N+1}{N} \frac{(\delta_{R,\text{Lag.}}^{\text{ND}}')^2}{(1 + \delta_{R,\text{Lag.}}^{\text{ND}})} + 4\pi a^2 \bar{\rho}_m \delta_{R,\text{Lag.}}^{\text{ND}} (1 + \delta_{R,\text{Lag.}}^{\text{ND}}) = 0 , \quad (6.40)$$

where $N = 3$ for spherical collapse, $N = 2$ for cylindrical collapse and $N = 1$ for planar collapse.

6.4 Non-linear perturbation theory

I will now derive the standard perturbative scheme to solve the equations 6.1 to 6.3 in increasingly higher orders of the linear density contrast $\delta_{L,m}$, following closely the presentation in Bernardeau et al. (2002). To do so, we return to the co-moving coordinates \mathbf{x}

(the ones that follow only the Hubble flow but not the individual dust elements) and introduce the corresponding Fourier space coordinates as \mathbf{k} . As explained in section section 5.2, we only consider scalar perturbations in this work. This means that \mathbf{v} is completely determined through its divergence $\vartheta = \nabla \cdot \mathbf{v}$. In Fourier space this reads

$$\begin{aligned}\tilde{\vartheta} &= -i\mathbf{k}\tilde{\mathbf{v}} \\ \Rightarrow \tilde{\mathbf{v}} &= i\frac{\mathbf{k}}{k^2}\tilde{\vartheta}.\end{aligned}\quad (6.41)$$

where $\tilde{\mathbf{v}}$ and $\tilde{\vartheta}$ are the Fourier space versions of the perturbations and $k = |\mathbf{k}|$. Using Equation 6.41 we can write the Fourier space version of the continuity equation as

$$\begin{aligned}\tilde{\delta}'_m(\mathbf{k}) + \tilde{\vartheta}(\mathbf{k}) &= - \int d^3k_1 \tilde{\delta}_m(\mathbf{k} - \mathbf{k}_1)\tilde{\vartheta}(\mathbf{k}_1) \left[1 + \frac{(\mathbf{k} - \mathbf{k}_1)\mathbf{k}_1}{k_1^2} \right] \\ &= - \int d^3k_1 d^3k_2 \delta^D(\mathbf{k} - \mathbf{k}_{12})\tilde{\delta}_m(\mathbf{k}_2)\tilde{\vartheta}(\mathbf{k}_1) \left[1 + \frac{\mathbf{k}_2\mathbf{k}_1}{k_1^2} \right].\end{aligned}\quad (6.42)$$

where δ^D is the Dirac delta distribution and $\mathbf{k}_{12} = \mathbf{k}_1 + \mathbf{k}_2$. (Also, I will not explicitly note the dependence on conformal time η to keep the notation concise.) From the Poisson equation we know that the metric perturbation ϕ and the matter density contrast δ_m are related in Fourier space through

$$\begin{aligned}\tilde{\phi} &= -4\pi a^2 \bar{\rho}_m \frac{\tilde{\delta}_m}{k^2} \\ \Rightarrow \widetilde{\nabla\phi} &= 4\pi a^2 \bar{\rho}_m \frac{i\mathbf{k}}{k^2} \tilde{\delta}_m.\end{aligned}\quad (6.43)$$

Using both 6.41 and 6.43 we can write the Euler equation 6.3 in Fourier space as

$$i\frac{\mathbf{k}}{k^2}\tilde{\vartheta}'(\mathbf{k}) + 4\pi a^2 \bar{\rho}_m \frac{i\mathbf{k}}{k^2} \tilde{\delta}_m(\mathbf{k}) + i\frac{\mathbf{k}}{k^2} \mathcal{H}\tilde{\vartheta}(\mathbf{k}) = -i \int d^3k_1 \tilde{\vartheta}(\mathbf{k} - \mathbf{k}_1)\tilde{\vartheta}(\eta, \mathbf{k}_1) \frac{[(\mathbf{k} - \mathbf{k}_1)\mathbf{k}_1]}{|\mathbf{k} - \mathbf{k}_1|^2 k_1^2} \mathbf{k}_1. \quad (6.44)$$

Multiplying this equation by $-i\mathbf{k}$ we get

$$\begin{aligned}\tilde{\vartheta}'(\mathbf{k}) + 4\pi a^2 \bar{\rho}_m \tilde{\delta}_m(\mathbf{k}) + \mathcal{H}\tilde{\vartheta}(\mathbf{k}) &= - \int d^3k_1 \tilde{\vartheta}(\mathbf{k} - \mathbf{k}_1)\tilde{\vartheta}(\eta, \mathbf{k}_1) \frac{[(\mathbf{k} - \mathbf{k}_1)\mathbf{k}_1]}{|\mathbf{k} - \mathbf{k}_1|^2 k_1^2} \mathbf{k}\mathbf{k}_1 \\ &= - \int d^3k_1 d^3k_2 \delta^D(\mathbf{k} - \mathbf{k}_{12})\tilde{\vartheta}(\mathbf{k}_2)\tilde{\vartheta}(\eta, \mathbf{k}_1) \frac{[\mathbf{k}_2\mathbf{k}_1]}{k_2^2 k_1^2} \mathbf{k}_{12}\mathbf{k}_1 \\ &= -\frac{1}{2} \int d^3k_1 d^3k_2 \delta^D(\mathbf{k} - \mathbf{k}_{12})\tilde{\vartheta}(\mathbf{k}_2)\tilde{\vartheta}(\eta, \mathbf{k}_1) \frac{[\mathbf{k}_2\mathbf{k}_1]}{k_2^2 k_1^2} k_{12}^2,\end{aligned}\quad (6.45)$$

where in the last line we have used the freedom to relabel the integration variables $\mathbf{k}_1 \leftrightarrow \mathbf{k}_2$. Introducing the kernel functions

$$\begin{aligned}\alpha(\mathbf{k}_1, \mathbf{k}_2) &= 1 + \frac{\mathbf{k}_1 \cdot \mathbf{k}_2}{k_1^2} \\ \beta(\mathbf{k}_1, \mathbf{k}_2) &= \frac{k_{12}^2}{2} \frac{[\mathbf{k}_2 \cdot \mathbf{k}_1]}{k_2^2 k_1^2}\end{aligned}\quad (6.46)$$

we can write the continuity and Euler equation in the more concise form (cf. equations 37 & 38 of Bernardeau et al., 2002)

$$\begin{aligned}\tilde{\delta}'_m(\mathbf{k}) + \tilde{\vartheta}(\mathbf{k}) &= - \int d^3 k_1 d^3 k_2 \delta^D(\mathbf{k} - \mathbf{k}_{12}) \tilde{\delta}_m(\mathbf{k}_2) \tilde{\vartheta}(\mathbf{k}_1) \alpha(\mathbf{k}_1, \mathbf{k}_2) \\ \tilde{\vartheta}'(\mathbf{k}) + 4\pi a^2 \bar{\rho}_m \tilde{\delta}_m(\mathbf{k}) + \mathcal{H} \tilde{\vartheta}(\mathbf{k}) &= - \int d^3 k_1 d^3 k_2 \delta^D(\mathbf{k} - \mathbf{k}_{12}) \tilde{\vartheta}(\mathbf{k}_2) \tilde{\vartheta}(\eta, \mathbf{k}_1) \beta(\mathbf{k}_1, \mathbf{k}_2).\end{aligned}\quad (6.47)$$

To solve these equations we assume that $\tilde{\delta}_m$ and $\tilde{\vartheta}$ can be written in the form

$$\begin{aligned}\tilde{\delta}_m(\eta, \mathbf{k}) &= \sum_{n=1}^{\infty} \delta_m^{(n)}(\eta, \mathbf{k}) \\ \tilde{\vartheta}(\eta, \mathbf{k}) &= \sum_{n=1}^{\infty} \vartheta^{(n)}(\eta, \mathbf{k}),\end{aligned}\quad (6.48)$$

where the terms $\delta_m^{(n)}, \vartheta^{(n)}$ are of the order $\mathcal{O}(\tilde{\delta}_{L,m}^n)$, i.e. of n th order in the linear density contrast. With this ansatz 6.47 can be solved order-by-order by inserting the series' 6.48 and equating only those terms on the left and right hand-side of 6.47 that are of the same order in $\tilde{\delta}_{L,m}$. At first order this reads

$$\begin{aligned}\tilde{\delta}^{(1)'}(\mathbf{k}) + \tilde{\vartheta}^{(1)}(\mathbf{k}) &= 0 \\ \tilde{\vartheta}^{(1)'}(\mathbf{k}) + 4\pi a^2 \bar{\rho}_m \tilde{\delta}^{(1)}(\mathbf{k}) + \mathcal{H} \tilde{\vartheta}^{(1)}(\mathbf{k}) &= 0,\end{aligned}\quad (6.49)$$

i.e. the first order gives exactly the linearized equations that were discussed in section 6.2 and we get

$$\tilde{\delta}^{(1)} = \tilde{\delta}_{L,m}, \quad \tilde{\vartheta}^{(1)} = \tilde{\vartheta}_L. \quad (6.50)$$

In part IV we will also need the second order perturbations $\tilde{\delta}^{(2)}$ and $\tilde{\vartheta}^{(2)}$ and I discuss those in detail in section 16.B (which is part of the paper Friedrich et al. 2018). For a detailed discussion of the complete perturbative series see Bernardeau et al. (2002) and references therein (especially Fry, 1984, 1985).

6.5 N-point functions

Consider a set of random variables f_1, \dots, f_N that are possibly correlated. The *connected part* of their joint moment $\langle f_1 f_2 \dots f_N \rangle$ is defined as (cf. sections 3.2 & 3.3 of Bernardeau et al., 2002)

$$\langle f_1 f_2 \dots f_N \rangle_c = \langle f_1 f_2 \dots f_N \rangle - \sum_{\substack{\text{partitions } \mathcal{P} \text{ of } \{1, \dots, N\} \\ \text{into subsets } s}} \prod_{s \in \mathcal{P}} \langle f_{s_1} \dots f_{s_{\#s}} \rangle . \quad (6.51)$$

For $N = 1, 2, 3$ this definition, e.g., gives

$$\begin{aligned} \langle f_1 \rangle_c &= \langle f_1 \rangle \\ \langle f_1 f_2 \rangle_c &= \langle f_1 f_2 \rangle - \langle f_1 \rangle \langle f_2 \rangle \\ \langle f_1 f_2 f_3 \rangle_c &= \langle f_1 f_2 f_3 \rangle - \langle f_1 f_2 \rangle \langle f_3 \rangle - \langle f_2 f_3 \rangle \langle f_1 \rangle - \langle f_3 f_1 \rangle \langle f_2 \rangle - \langle f_1 \rangle \langle f_2 \rangle \langle f_3 \rangle . \end{aligned} \quad (6.52)$$

This definition is chosen exactly such that $\langle f_1 f_2 \dots f_N \rangle_c = 0$ for $N \geq 3$ whenever f_1, \dots, f_N have a joint Gaussian distribution.

Now if $\mathbf{x}_1, \dots, \mathbf{x}_N$ are a set of locations in the universe, then the *N-point correlation function* $\xi_{\delta}^{(N)}(\mathbf{x}_1, \dots, \mathbf{x}_N)$ of the matter density contrast field δ_m is defined as

$$\xi_{\delta}^{(N)}(\eta, \mathbf{x}_1, \dots, \mathbf{x}_N) = \langle \delta_m(\eta, \mathbf{x}_1) \delta_m(\eta, \mathbf{x}_2) \dots \delta_m(\eta, \mathbf{x}_N) \rangle_c . \quad (6.53)$$

For $N = 2$ this is exactly the 2-point correlation function. We can also consider connected moments in Fourier space for Fourier modes $\mathbf{k}_1, \dots, \mathbf{k}_N$. Because of the statistical homogeneity of the density field only modes with $\mathbf{k}_1 + \dots + \mathbf{k}_N = 0$ can be correlated (see section 3.2 of Bernardeau et al., 2002) which results in the factor $\delta^D(\mathbf{k}_1 + \dots + \mathbf{k}_N)$ in the expression

$$\langle \tilde{\delta}_m(\eta, \mathbf{k}_1) \tilde{\delta}_m(\eta, \mathbf{k}_2) \dots \tilde{\delta}_m(\eta, \mathbf{k}_N) \rangle_c = (2\pi)^3 \delta^D(\mathbf{k}_1 + \dots + \mathbf{k}_N) B_{\delta}^{(N)}(\eta, \mathbf{k}_1, \dots, \mathbf{k}_N) . \quad (6.54)$$

The functions $B_{\delta}^{(N)}(\eta, \mathbf{k}_1, \dots, \mathbf{k}_N)$ are the N-point spectra of δ_m and we have $B_{\delta}^{(2)} = P_{\delta}$.

Since the initial density contrast $\delta_m(\eta, \mathbf{x})$ is a Gaussian random field the linear density contrast $\delta_{L,m}(\eta, \mathbf{x})$ is also a Gaussian random field. This means that all N-point correlations of $\delta_{L,m}$ with $N \geq 3$ vanish. Together with the perturbative series introduced in the previous section this has the useful consequence, that all N-point correlation functions of the total, non-linear density contrast can be expressed as a series

$$B_{\delta}^{(N)}(\eta, \mathbf{k}_1, \dots, \mathbf{k}_N) = \sum_{n=n_{\min}}^{\infty} B_{\delta}^{(N)(n)}(\eta, \mathbf{k}_1, \dots, \mathbf{k}_N) , \quad (6.55)$$

where $B_{\delta}^{(N)(n)}$ is of the order $(P_{L,\delta})^n$. Using the perturbative series of the previous section and the fact that $\delta_{L,m}$ is Gaussian one can also show that the lowest non-zero order n_{\min} in equation 6.55 is given by

$$n_{\min} = N - 1 . \quad (6.56)$$

I demonstrate this in detail for $N = 3$ in section 16.B (which is part of the paper Friedrich et al. 2018) and a general derivation can, e.g. , be found in Fry (1984, 1985); Bernardeau et al. (2002). The order n_{\min} is called the *tree level* of perturbation theory, the next to leading order is called the *one-loop level*, the second to leading order is called the *2-loop level* etc. .

Chapter 7

The cosmic density PDF

The N-point correlation functions ξ_δ^N vanish for the initial density contrast field $\delta_m(\eta_i, \mathbf{x})$ because it is a Gaussian random field (cf. the discussion of connected moments in section 6.5). As a consequence, all cosmological information is initially encoded in the 2-point statistics of $\delta_m(\eta_i, \mathbf{x})$ (i.e. in its 2-point correlation function or equivalently in its power spectrum). As the density field evolves the ξ_δ^N become nonzero and the power spectrum becomes an incomplete description of the cosmic density fluctuations. In other words, the non-linear evolution of the density field migrates cosmological information from the 2-point correlation function to higher order statistics of the density field. When we are using the late-time large scale structure of the Universe to test cosmological models it is hence desirable to investigate probes that are sensitive to higher order statistical properties of the density field.

One such probe that is important for this work is the probability density function (PDF) of density fluctuations within a certain radius R . I.e. defining the smoothed density contrast field

$$\delta_R(\eta, \mathbf{x}) \equiv \frac{3}{4\pi R^3} \int_{|\mathbf{y}-\mathbf{x}| < R} d^3y \delta_m(\eta, \mathbf{y}) \quad (7.1)$$

one goal of my work is to compare observations and theoretical predictions of the PDF $p(\delta_R)$. As we will see in section 7.1, this function is sensitive to all higher order correlation functions of the density field. In section 7.2 I will review a way of theoretically predicting $p(\delta_R)$ that was put forward by Valageas (2002a). This modeling approach can not directly be applied to data since real observations are only sensitive projections of the density field along our backward light-cone (cf. chapter 9). But it is possible to extend the 3D frame work to projected versions of the density field (see, e.g. , Bernardreau & Valageas 2000 as well as chapter 16 of this work). A particular benefit of studying the density PDF is that it can be modeled accurately even in the absence of an exact non-linear model for the higher order correlation function. I will review how to do so in section 7.3.

7.1 From the cumulant generating function to the PDF

Because of the homogeneity of the universe the PDF of $\delta_R(\eta, \mathbf{x})$ does not depend on \mathbf{x} and I will hence drop the dependence on the spatial coordinates in the following. The *moment generating function* of δ_R is defined as

$$\psi_R(y, \eta) = \sum_{n=0}^{\infty} \frac{\langle \delta_R(\eta)^n \rangle}{n!} y^n , \quad (7.2)$$

where the expectation value $\langle \cdot \rangle$ can either be understood as taking the ensemble average over different realizations of the initial conditions or as averaging over the spatial coordinates \mathbf{x} . Using the connected moments instead of the plain moments (cf. section 6.5) gives the definition of the *cumulant generating function* (CGF) as

$$\varphi_R(y, \eta) = \sum_{n=0}^{\infty} \frac{\langle \delta_R(\eta)^n \rangle_c}{n!} y^n , \quad (7.3)$$

It can be shown that the two functions are related through (Bernardeau et al., 2002, 2014)

$$e^{\varphi_R(y, \eta)} = \psi_R(y, \eta) . \quad (7.4)$$

Using this we can express the CGF through the expectation value

$$e^{\varphi_R(y, \eta)} = \int d\delta_R p(\delta_R | \eta) e^{y\delta_R} , \quad (7.5)$$

where $p(\delta_R | \eta)$ is the PDF of δ_R at time η . Equation 7.5 is simply the Laplace transform of the PDF $p(\delta_R | \eta)$ and as long as the power series defining the CGF converges we can invert the relation between CGF and PDF via the inverse Laplace transform (Bernardeau & Valageas, 2000; Valageas, 2002a; Bernardeau et al., 2014). This is given by

$$p(\delta_R | \eta) = \int_{-\infty}^{\infty} \frac{dy}{2\pi} e^{-iy\delta_R + \varphi_R(iy)} . \quad (7.6)$$

A brief description of how to evaluate this integral numerically is given in appendix B of Bernardeau et al. (2014).

The connected moments (or *cumulants*) $\langle \delta_R(\eta)^n \rangle_c$ are related to the N-point functions of the density contrast field via

$$\begin{aligned} \langle \delta_R(\eta, \mathbf{0})^n \rangle_c &= \left(\frac{3}{4\pi R^3} \right)^n \int_{|\mathbf{x}_1| < R} d^3 x_1 \dots \int_{|\mathbf{x}_n| < R} d^3 x_n \langle \delta_m(\eta, \mathbf{x}_1) \dots \delta_m(\eta, \mathbf{x}_n) \rangle_c \\ &= \left(\frac{3}{4\pi R^3} \right)^n \int_{|\mathbf{x}_1| < R} d^3 x_1 \dots \int_{|\mathbf{x}_n| < R} d^3 x_n \xi_{\delta}^{(n)}(\eta, \mathbf{x}_1, \dots, \mathbf{x}_n) . \end{aligned} \quad (7.7)$$

Equation 7.6 then tells us that $p(\delta_R | \eta)$ depends on the N-point correlation functions of the density field at *all orders*.

7.2 Path integral approach for the CGF

Following Valageas (2002a) I will now review a way to predict the cumulant generating function $\varphi_R(y)$ without having to compute all cumulants $\langle \delta_R(\eta)^n \rangle_c$ individually.

Let $\delta_{L,m}$ be the linearly extrapolated initial conditions of the density field, i.e. $\delta_{L,m}(\eta, \mathbf{x}) = D(\eta)\delta_m(\eta, \mathbf{x})$. This is a Gaussian random field. Since δ_R is completely determined by the linear field $\delta_{L,m}$ we can express it as a functional of the form $\delta_R[\delta_{L,m}]$. The expectation value in 7.5 can then be expressed as a Gaussian path integral over all possible configurations of $\delta_{L,m}$, i.e.

$$e^{\phi_R(y,\eta)} = |2\pi\Delta_L|^{-\frac{1}{2}} \int \mathcal{D}\delta_{L,m} \exp \left\{ y\delta_R[\delta_{L,m}] - \frac{1}{2}\delta_{L,m}\Delta_L^{-1}\delta_{L,m} \right\}. \quad (7.8)$$

Here we introduced the notation

$$\delta_{L,m}\Delta_L^{-1}\delta_{L,m} \equiv \int d^3x_1 d^3x_2 \delta_{L,m}(\eta, \mathbf{x}_1)\delta_{L,m}(\eta, \mathbf{x}_2)\Delta_L^{-1}(\mathbf{x}_1, \mathbf{x}_2|\eta) \quad (7.9)$$

and the (infinite dimensional) matrix Δ_L is given in terms of the linear power spectrum as

$$\Delta_L(\mathbf{x}_1, \mathbf{x}_2|\eta) = \xi_\delta(\eta, \mathbf{x}_1, \mathbf{x}_2). \quad (7.10)$$

Let us denote the exponent in 7.8 by $-S_y$. Then this path integral can be approximated using Laplace's method as

$$e^{\phi_R(y,\eta)} \approx |2\pi\Delta_L|^{-\frac{1}{2}} |M/(2\pi)|^{-\frac{1}{2}} e^{-S_y[\delta_{L,m}^*]} \quad (7.11)$$

where $\delta_{L,m}^*$ is the global minimum of the functional $S_y[\cdot]$ and the matrix M is given by

$$M(\mathbf{x}_1, \mathbf{x}_2) = \frac{\delta^2 S_y}{\delta[\delta_L(\mathbf{x}_1)]\delta[\delta_L(\mathbf{x}_2)]} \Big|_{\delta_{L,m}^*} \quad (7.12)$$

$$= -y \frac{\delta^2 \delta_R}{\delta[\delta_L(\mathbf{x}_1)]\delta[\delta_L(\mathbf{x}_2)]} \Big|_{\delta_L^*} + \Delta_L^{-1}(\mathbf{x}_1, \mathbf{x}_2). \quad (7.13)$$

Define the matrix A_y through

$$\begin{aligned} \delta_D(\mathbf{x}_1 - \mathbf{x}_2) + A_y(\mathbf{x}_1, \mathbf{x}_2) &= \Delta_L^{1/2} M \Delta_L^{1/2}(\mathbf{x}_1, \mathbf{x}_2) \\ &\equiv \int d^3y_1 d^3y_2 \Delta_L^{1/2}(\mathbf{x}_1, \mathbf{y}_1) M(\mathbf{y}_1, \mathbf{y}_2) \Delta_L^{1/2}(\mathbf{y}_2, \mathbf{x}_2). \end{aligned} \quad (7.14)$$

Then the above approximation for the cumulant generating function becomes

$$\begin{aligned} \phi_R(y, \eta) &\approx -S_y[\delta_{L,m}^*] - \frac{1}{2} \ln |\Delta_L M| \\ &= -S_y[\delta_L^*] - \frac{1}{2} \text{Tr} \ln(\delta_D + A_y). \end{aligned} \quad (7.15)$$

It has been shown by Valageas (2002a,b) that the first term in the above approximation contains the tree level perturbative results for the connected moments of δ_R (i.e. when using only the tree level contribution the N-point functions when evaluating equation 7.7; cf. section 6.5) while expanding the second term in powers of y is equivalent to the 1-loop results.

Valageas (2002a) has shown that the minimum $\delta_{L,m}^*$ of the action S_y is spherically symmetric. The functional $\delta_R[\delta_{L,m}^*]$ can then be computed exactly by solving the spherical collapse equation (cf. section 6.3). This means that it is possible to compute the tree level contribution $\phi_R(y, \eta) \approx -S_y[\delta_L^*]$ numerically.

7.3 Non-linear rescaling of the CGF

The accuracy of the tree level approximation $\phi_R(y, \eta) \approx -S_y[\delta_L^*]$ can be significantly boosted with a rescaling technique that has, e.g. , been employed by Bernardeau & Valageas (2000); Valageas (2002a); Bernardeau et al. (2014, 2015). For this rescaling one defines the coefficients

$$S_n \equiv \frac{\langle \delta_R(\eta)^n \rangle_c}{\{\langle \delta_R(\eta)^2 \rangle_c\}^{n-1}} \quad (7.16)$$

and re-expresses the cumulant generating function as

$$\begin{aligned} \varphi_R(y, \eta) &= \sum_{n=0}^{\infty} \frac{S_n}{n!} \left\{ \langle \delta_R(\eta)^2 \rangle_c \right\}^{n-1} y^n \\ &= \frac{1}{\langle \delta_R(\eta)^2 \rangle_c} \sum_{n=0}^{\infty} \frac{S_n}{n!} \left\{ y \langle \delta_R(\eta)^2 \rangle_c \right\}^n. \end{aligned} \quad (7.17)$$

The coefficients S_n have a very weak dependence on time η and smoothing radius R that is well described by the tree level of perturbation theory for values of $R \geq 10\text{Mpc}$ (see figure 27 of Bernardeau et al. 2002). This means that even for the highly evolved density field we may use the leading order of perturbation theory to approximate the S_n :

$$\frac{\langle \delta_R(\eta)^n \rangle_c^{\text{exact}}}{\{\langle \delta_R(\eta)^2 \rangle_c^{\text{exact}}\}^{n-1}} = S_n^{\text{exact}} \approx S_n^{\text{tree}} = \frac{\langle \delta_R(\eta)^n \rangle_c^{\text{tree}}}{\{\langle \delta_R(\eta)^2 \rangle_c^{\text{tree}}\}^{n-1}}. \quad (7.18)$$

This can be used to improve the tree level approximation φ_R^{tree} of the CGF through the rescaling

$$\varphi_R^{\text{exact}}(y, \eta) \approx \frac{\langle \delta_R(\eta)^2 \rangle_c^{\text{tree}}}{\langle \delta_R(\eta)^2 \rangle_c^{\text{exact}}} \varphi_R^{\text{tree}} \left(y \frac{\langle \delta_R(\eta)^2 \rangle_c^{\text{exact}}}{\langle \delta_R(\eta)^2 \rangle_c^{\text{tree}}}, \eta \right). \quad (7.19)$$

In Bernardeau & Valageas (2000); Valageas (2002a); Bernardeau et al. (2014, 2015) the variance $\langle \delta_R(\eta)^2 \rangle_c^{\text{exact}}$ of the evolved density contrast is left a free parameter that is to be fit to observations. In this work, I instead use a fitting formula for the full non-linear power spectrum (obtained from N-body simulations by Smith et al., 2003; Takahashi et al., 2012) to derive predictions for $\langle \delta_R(\eta)^2 \rangle_c^{\text{exact}}$.

As me and my collaborators demonstrated in Friedrich et al. (2018), this approximation becomes quite accurate if 2D projections of the density field are considered instead of the 3D density field studied in this chapter (see chapter 16 of this work).

Part II

Background: Observations of the large scale Structure

Chapter 8

Distances and redshift estimation

8.1 Distance measures

In a Friedmann universe as described by the metric 5.7 there is no unambiguous concept of distance between two points in space time. Instead several distance measures can be defined (Hogg, 1999; Bartelmann & Schneider, 2001; Mukhanov, 2005). By determining some of those distance measures as a function of redshift (see below) for a set of astronomical objects one can study the geometry of the universe as well as its expansion history. In the following we summarize the most important distance measures. The different distance measure are also compared in figure 8.1 for a fiducial Λ CDM universe.

- Redshift:

Consider a photon from a distant light source (e.g. a distant galaxy) that reaches a telescope on earth with a wavelength λ_{obs} in the rest frame of the telescope. The redshift z of this photon describes a relative shift between that wavelength and the wavelength λ_{emit} that the photon had in the rest frame of the light source when it was emitted, i.e.

$$z := \frac{\lambda_{\text{obs}} - \lambda_{\text{emit}}}{\lambda_{\text{emit}}} . \quad (8.1)$$

This shift is found to be positive for most galaxies and it increases with the traveling time of the photons. This was the first observational hint for the expansion of the universe (Hubble, 1929).

In a perfectly homogeneous Friedmann universe the redshift is solely caused by the expansion of the universe, i.e. the expanding space stretches the wavelength of a photon while it is traveling towards us. In this case, the redshift can be expressed in terms of the scale factor of the Friedmann metric as

$$1 + z = \frac{a(t_{\text{obs}})}{a(t_{\text{emit}})} , \quad (8.2)$$

where t_{obs} is the time of observation and t_{emit} is the time of emission of the light. In the real universe, an additional contribution to the redshift of galaxies and other

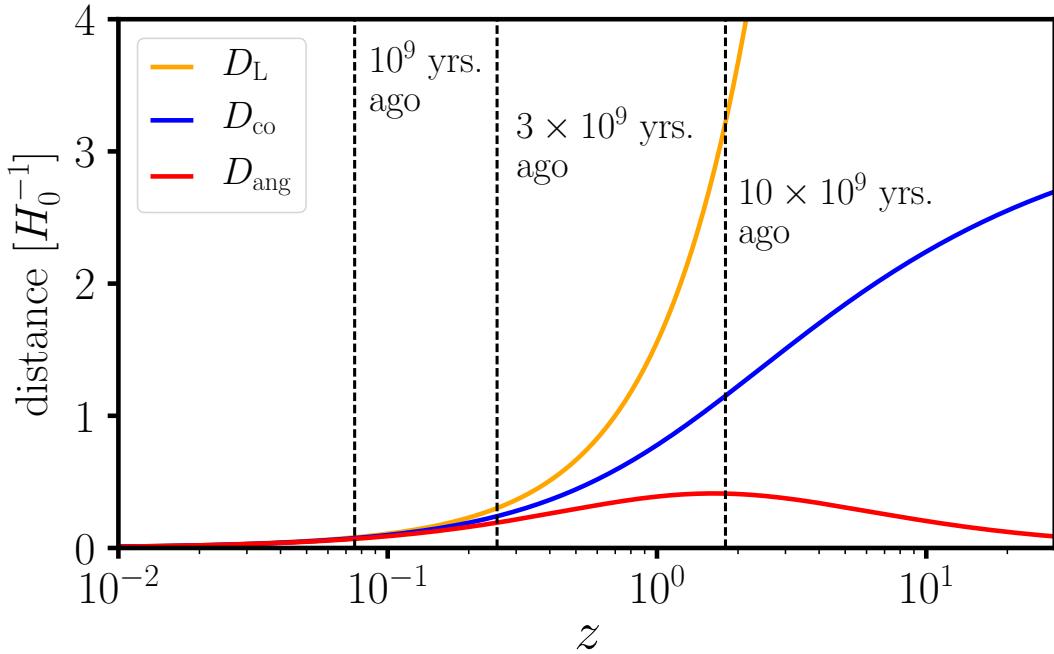


Figure 8.1: Comparing different distance measures as a function of redshift in a universe with $\Omega_{m,0} = 0.286$, $\Omega_{\Lambda,0} \approx 0.714$ and a negligible (but realistic) value of $\Omega_{r,0}$. All distances are measured in units of $1/H_0 \approx 4300$ Mpc. Vertical dashed lines indicate the cosmological time corresponding to different points along the redshift axis. Both luminosity distance (D_L , orange line) and co-moving distance (D_{co} , blue line) increase monotonically with redshift. The angular diameter distance however (D_{ang} , red line) decreases with increasing redshift for $z \gtrsim 2.0$. This means that above this redshift the angular sizes of objects on the sky would increase if they would be moved to higher redshifts. It should be noted that only D_{ang} and D_L are actually observable.

distant light sources comes from the peculiar motions of those sources wrt. the rest frame of the earth. If the velocity \mathbf{v} of the source relative to the Hubble flow is $\ll 1$ and if we assume the observer to rest in the Hubble flow, then the total redshift is given by

$$1 + z \approx \frac{a(t_{\text{obs}})}{a(t_{\text{emit}})} + v_{\parallel} , \quad (8.3)$$

where v_{\parallel} is the component of \mathbf{v} parallel to the line of sight of the observer.

- Co-moving distance:

If at time t_{obs} we observe an event at time t_{emit} , then the co-moving distance to this event is simply given by the conformal time between event and observation, i.e.

$$D_{\text{co}}(t_{\text{emit}}, t_{\text{obs}}) = \eta_{\text{obs}} - \eta_{\text{emit}} = \int_{t_{\text{emit}}}^{t_{\text{obs}}} \frac{dt}{a(t)} . \quad (8.4)$$

In contrast to the redshift of an object, the co-moving distance is not a direct observable but depends on our assumptions about cosmological parameters. This can be seen most directly by expressing D_{co} as a function of redshift. Differentiating equation 8.2 wrt. t one can see that

$$\frac{dz}{dt}(t) = -\frac{a(t_{\text{obs}})}{a(t)^2} \dot{a}(t) = -\frac{H(t)}{a(t)} , \quad (8.5)$$

where we have assumed that the peculiar velocity of the observed object is negligible. Using the fact that $z(t_{\text{obs}}) = 0$ this gives

$$D_{\text{co}}(z_{\text{emit}}) = \int_0^{z_{\text{emit}}} \frac{dz}{H(z)} . \quad (8.6)$$

Assuming that the universe is adequately described by the Λ CDM model (cf. section 5.1.6), both $a(t)$ and $H(t)$ depend on the exact values of the parameters Ω_m , Ω_r , Ω_{Λ} and H_0 that are used to integrate equations 5.34.

- Angular diameter distance:

Suppose the object that emits light towards us at time t_{emit} has a surface cross section perpendicular to the line of sight of A and that the solid angle of this object on the sky is $\delta\omega$. Then in a euclidean spacetime the distance to that object would be given by

$$D \approx \sqrt{\frac{A}{\delta\omega}} . \quad (8.7)$$

In analogy to this we define the ratio $\sqrt{A/\delta\omega}$ of an object in a Friedmann universe to be the angular diameter distance. It is given in terms of the scale factor a by (cf. Hogg, 1999)

$$D_{\text{ang}}(t_{\text{emit}}, t_{\text{obs}}) = \frac{a(t_{\text{emit}})}{a(t_{\text{obs}})} \cdot \Phi_k(D_{\text{co}}) , \quad (8.8)$$

where the function Φ_k has been defined in equations 5.8 and depends on the 3D-curvature of the universe (i.e. it describes the impact of the geometry of the universe on the divergence or convergence of light beams).

Using equations 8.2 and 8.6 one can express D_{ang} as a function of redshift. If the physical size of an object as well as its redshift are known then one can measure D_{ang} at that redshift. Measuring this for several values of z then yields a reconstruction of the function $D_{\text{ang}}(z)$. Just as we discussed for D_{co} , the redshift dependence of $D_{\text{ang}}(z)$ is determined by the exact values of the parameters of the Λ CDM model. Hence, a measurement of $D_{\text{ang}}(z)$ can be used to determine these parameters or more generally, to test whether any parameters of that model give a good representation of the observed universe. Such tests of the Λ CDM model based on measurements of the angular diameter distance have indeed been carried out, see, e.g. , Wang et al. (2017). This analysis is based on the Baryonic Acoustic Oscillation feature (BAO feature) in the 2-point correlation function of the galaxy distribution, whose physical size is known. (Any such feature or object with a known physical size is called a *standard ruler*. A standard ruler alternative to the BAO feature is, e.g. , given by the *linear point* of the correlation function, cf. Anselmi et al..)

- Luminosity distance:

If we receive a total flux of photon energy F from the object emitting at time t_{emit} then in a Euclidean space this flux will be related to the luminosity L of the object by

$$F = \frac{L}{4\pi D^2} , \quad (8.9)$$

where D is the distance to the object. In analogy to that the luminosity distance in a Friedmann universe is defined as

$$D_L = \left(\frac{L}{4\pi F} \right)^{\frac{1}{2}} . \quad (8.10)$$

For the metric 5.6 one can show that D_L is given by (cf. Hogg, 1999)

$$D_L(t_{\text{emit}}, t_{\text{obs}}) = D_{\text{ang}}(t_{\text{emit}}, t_{\text{obs}}) \cdot \left(\frac{a(t_{\text{obs}})}{a(t_{\text{emit}})} \right)^2 . \quad (8.11)$$

If we determine the redshift of objects with a known luminosity (so called *standard candles*) then this can be used to measure D_L as a function of z . As discussed for the angular diameter distance, a measurement of $D_L(z)$ can be used to test the Λ CDM model and to determine its parameters. This has been done for supernovae of type Ia, e.g. , by Riess et al. (1998). This study revealed for the first time the existence of a significant dark energy component in the total energy density of the universe.

8.2 Redshift estimation

The redshift of a luminous object can, e.g. , be determined by identifying emission or absorption lines in the spectral energy distribution (SED) of the light emitted by that object and comparing the location of these lines the location expected in the rest frame of the object (see, e.g. , Bolton et al., 2012, where this is described in detail for the Sloan Digital Sky Survey). Such a spectroscopic redshift determination is very time consuming which limits the amount of objects for which such redshift measurements have been obtained.

A less expensive method to determine the redshift values of, e.g. , a sample of galaxies is the so-called *photometric redshift estimation* (photo z estimation). To obtain photo z values for an individual galaxy one measures the flux of the light from that galaxy in different spectral filters. The flux values in the different filters can that, e.g., be compared to the convolution of these filters with model SEDs for fit for the redshift of these SEDs (Bender et al., 2001; Greisel et al., 2013). Alternatively one can determine the relation between redshift and photometric fluxes empirically using a small training sample for which spectroscopic redshift measurements are available (e.g. using machine learning techniques, see Collister & Lahav, 2004; Rau et al., 2015) or one can use a combined approach based on both model SED fitting and machine learning (Hoyle et al., 2015).

Photometric redshift estimation can be considered a low-resolution version of spectroscopic redshift estimation. But while photo z estimates for individual galaxies can be unreliable it is nevertheless possible to determine the overall redshift distribution of an ensemble of object accurately enough to allow studies of the line-of-sight density field (Rau et al., 2015; Hoyle et al., 2018, see also chapter 9).

Chapter 9

2D power spectra and correlation functions

9.1 2-point statistics on the sky

Consider the 3D density contrast field $\delta(\eta, \mathbf{x})$ that was defined in section 4.3, where η and \mathbf{x} are the coordinates of the conformal Friedmann metric. In section 6.1 we already defined the 2-point correlation function and the power spectrum of this field as

$$\xi_\delta(\mathbf{x}, \eta) \equiv \langle \delta(\eta, \mathbf{x}_0) \delta(\eta, \mathbf{x}_0 + \mathbf{x}) \rangle \quad (9.1)$$

and

$$P_\delta(\mathbf{k}, \eta) = \int d^2x e^{-i\mathbf{k}\mathbf{x}} \xi_\delta(\mathbf{x}, \eta) . \quad (9.2)$$

Here, the expectation value $\langle \cdot \rangle$ can be understood either as an ensemble average over different realizations of the initial conditions or as an average over all locations \mathbf{x}_0 in the universe. Because of the isotropy of the initial conditions we have $\xi_\delta(\mathbf{x}, \eta) = \xi_\delta(x, \eta)$ and $P_\delta(\mathbf{k}, \eta) = P_\delta(k, \eta)$, i.e. both functions only depend on the absolute value of their spatial arguments.

The functions ξ_δ and P_δ encode cosmological information. Given an observation of the density contrast $\delta(\eta, \mathbf{x})$ we could measure these 2-point statistics - i.e. by evaluating 9.1 as a spatial average - and use those measurements to test the validity of cosmological models. There are however several reasons, why neither ξ_δ nor P_δ - at least in the above form - are observable:

1. We only observe our past light cone, i.e. we only have access to those values of $\delta(\eta, \mathbf{x})$ for which η and \mathbf{x} are located on past light-like geodesics converging at our position.
2. We usually don't have direct information about the total matter density field. Instead, we, e.g. , observe the galaxy density field or the cosmic shear field (cf. chapter 10) that are only tracing the total matter density field in a statistical way. Inferring δ from such observations is subject to noise (e.g. shot-noise of the galaxy distribution

or shape-noise of the cosmic shear field) and potentially biases (e.g. galaxy clustering bias or intrinsic alignments in the shear field).

3. Even along the light cone, photometric surveys such as the Dark Energy Survey (cf. chapter 11) have no exact information about the location of observed objects. Instead, only a distribution of tracers along the redshift coordinate can be inferred (cf. section 8.2).

We will address point 2 above briefly in sections 9.2 and 10.4.5 and in more detail in the papers Friedrich, Gruen et al., 2018; Gruen, Friedrich et al., 2018 that are presented in chapters 16 and 17.

Points 1 and 3 state that our actual observables are 2D projections of the density contrast onto the celestial sphere. Such a line-of-sight projection can, e.g. , be expressed as

$$\delta_{m,2D}(\hat{\mathbf{n}}) = \int_0^\infty dw q(w) \delta_m(\eta_0 - w, w\hat{\mathbf{n}}) . \quad (9.3)$$

Here, w is the co-moving distance along the light cone (cf. section 8.1), η_0 is the value of the conformal time today and $q(w)$ is a particular projection kernel that describes how an ensemble of tracer objects is distributed along the line-of-sight. $\delta_{m,2D}(\hat{\mathbf{n}})$ is then a scalar function on the sky and $\hat{\mathbf{n}}$ is a unit vector specifying a particular position on the sky.

The 3D correlation function ξ_δ is no direct observable but we can attempt to measure the projected correlation function

$$\begin{aligned} \xi_{2D}(\triangleleft\hat{\mathbf{n}}_1\hat{\mathbf{n}}_2) &= \langle \delta_{m,2D}(\hat{\mathbf{n}}_1) \delta_{m,2D}(\hat{\mathbf{n}}_2) \rangle \\ &= \int_0^\infty dw_1 \int_0^\infty dw_2 q(w_1) q(w_2) \langle \delta_m(\eta_0 - w_1, w_1\hat{\mathbf{n}}_1) \delta_m(\eta_0 - w_2, w_2\hat{\mathbf{n}}_2) \rangle . \end{aligned} \quad (9.4)$$

Because of the isotropy of the universe this only depends on the angle between the unit vectors $\hat{\mathbf{n}}_1$ and $\hat{\mathbf{n}}_2$. The expectation value in the 2nd line of 9.4 is the correlation function of 2 points that are located in different shells of co-moving distance. In the late time, non-linear evolution of the density field this function is hard to predict theoretically, so we would like to reduce equation 9.4 to an expression that only involves the single-time correlation function $\xi_\delta(\eta, \cdot)$. This can be done with the so-called *Limber approximation* (Limber, 1953). A detailed derivation of this can be found in Bartelmann & Schneider (2001). I want to repeat their arguments here because I would like the reader to be aware of the exact meaning of this approximation.

The Limber approximation assumes that the inter-shell-correlation $\langle \delta_m(\eta_0 - w_1, w_1\hat{\mathbf{n}}_1) \delta_m(\eta_0 - w_2, w_2\hat{\mathbf{n}}_2) \rangle$ approaches 0 much faster for $\Delta w = w_2 - w_1 \rightarrow \pm\infty$ than the average variations in the functions $q(w_1)$ and $q(w_2)$. Writing $w = w_1$ and $\Delta w = w_2 - w_1$ Limber's

approximation reads

$$\begin{aligned}\xi_{2D}(\langle \hat{\mathbf{n}}_1 \hat{\mathbf{n}}_2 \rangle) &\approx \int_0^\infty dw q(w)^2 \int_{-\infty}^\infty d\Delta w \langle \delta_m(\eta_0 - w, w\hat{\mathbf{n}}_1) \delta_m(\eta_0 - w, (w + \Delta w)\hat{\mathbf{n}}_2) \rangle \\ &= \int_0^\infty dw q(w)^2 \int_{-\infty}^\infty d\Delta w \xi_\delta(\eta_0 - w, |w\hat{\mathbf{n}}_1 - (w + \Delta w)\hat{\mathbf{n}}_2|) .\end{aligned}\quad (9.5)$$

To simplify this further we need two additional assumptions. First, we assume that the integrand $\xi_\delta(\eta_0 - w, |w\hat{\mathbf{n}}_1 - (w + \Delta w)\hat{\mathbf{n}}_2|)$ is negligible for any values other than $\Delta w \ll \Phi_k(w)$ where $\Phi_k(w)$ is the function defined in equation 5.8 that depends on the spatial curvature of the Friedmann universe. This assumption can still be considered a part of the Limber approximation, since typical photometric galaxy survey the width of the projection kernel $q(w)$ is of the order of w itself (and in a flat universe $w = \Phi_k(w)$). Secondly, we assume that $|\langle \hat{\mathbf{n}}_1 \hat{\mathbf{n}}_2 \rangle| \ll 1$ i.e. we employ a small angle approximation. This is NOT part of the Limber approximation, though the two are often used at the same time.

With the above assumptions we can approximate

$$|w\hat{\mathbf{n}}_1 - (w + \Delta w)\hat{\mathbf{n}}_2|^2 \approx \Phi_k(w)^2 (\langle \hat{\mathbf{n}}_1 \hat{\mathbf{n}}_2 \rangle)^2 + \Delta w^2 \quad (9.6)$$

and the 2D correlation function becomes

$$\xi_{2D}(\langle \hat{\mathbf{n}}_1 \hat{\mathbf{n}}_2 \rangle) \approx \int_0^\infty dw q(w)^2 \int_{-\infty}^\infty d\Delta w \xi_\delta(\eta_0 - w, \sqrt{\Phi_k(w)^2 (\langle \hat{\mathbf{n}}_1 \hat{\mathbf{n}}_2 \rangle)^2 + \Delta w^2}) .\quad (9.7)$$

Since this assumes small angles we can approximate the sky by its tangential plane at some region of interest. The angle $\langle \hat{\mathbf{n}}_1 \hat{\mathbf{n}}_2 \rangle$ can then be represented by some vector $\boldsymbol{\theta}$ in this plane. This gives

$$\xi_{2D}(\boldsymbol{\theta}) \approx \int_0^\infty dw q(w)^2 \int_{-\infty}^\infty d\Delta w \xi_\delta(\eta_0 - w, \sqrt{\Phi_k(w)^2 \boldsymbol{\theta}^2 + \Delta w^2}) .\quad (9.8)$$

Even if a prediction for the 3D correlation function ξ_δ is available, this expression is still very tedious to evaluate numerically. We will however see that at the level of the power spectrum - which contains the same cosmological information as the 2-point function - the Limber equation is much more convenient.

Using the fact that ξ_δ and P_δ are related by a Fourier transform, we can write

$$\begin{aligned}\xi_{2D}(\boldsymbol{\theta}) &\approx \int_0^\infty dw q(w)^2 \int_{-\infty}^\infty d\Delta w \xi_\delta(\eta_0 - w, \sqrt{\Phi_k(w)^2 \boldsymbol{\theta}^2 + \Delta w^2}) \\ &= \int_0^\infty dw q(w)^2 \int_{-\infty}^\infty d\Delta w \int \frac{d^2 k_\perp dk_\parallel}{(2\pi)^3} * \\ &\quad * P_\delta \left(\eta_0 - w, \sqrt{k_\perp^2 + k_\parallel^2} \right) \exp(i\mathbf{k}_\perp \boldsymbol{\theta} \Phi_k(w) + ik_\parallel \Delta w) .\end{aligned}\quad (9.9)$$

In this expression the integration over $d\Delta w$ can be carried out to give a factor of $2\pi\delta_D(k_{\parallel})$, where δ_D is again the Dirac distribution. We hence get to

$$\begin{aligned}\xi_{2D}(\boldsymbol{\theta}) &\approx \int_0^\infty dw q(w)^2 \int \frac{d^2 k_{\perp}}{(2\pi)^2} P_{\delta}(\eta_0 - w, k_{\perp}) \exp(i\mathbf{k}_{\perp}\boldsymbol{\theta}\Phi_k(w)) \\ &= \int_0^\infty dw \frac{q(w)^2}{\Phi_k(w)^2} \int \frac{d^2 \ell}{(2\pi)^2} P_{\delta}\left(\eta_0 - w, \frac{\ell}{\Phi_k(w)}\right) \exp(i\ell\boldsymbol{\theta}) \\ &= \int \frac{d^2 \ell}{(2\pi)^2} \exp(i\ell\boldsymbol{\theta}) \int_0^\infty dw \frac{q(w)^2}{\Phi_k(w)^2} P_{\delta}\left(\eta_0 - w, \frac{\ell}{\Phi_k(w)}\right).\end{aligned}\quad (9.10)$$

Defining P_{2D} as the Fourier transform of ξ_{2D} we can see from the last line of 9.10 that

$$P_{2D}(\ell) = \int_0^\infty dw \frac{q(w)^2}{\Phi_k(w)^2} P_{\delta}\left(\eta_0 - w, \frac{\ell}{\Phi_k(w)}\right).\quad (9.11)$$

This is the power spectrum of the projected field δ_{2D} using both the Limber and the flat sky approximation. On the scales that are probed by galaxy surveys such as the Dark Energy Survey (DES Collaboration et al., 2017) or the Kilo Degree Survey (van Uitert et al., 2018) it can be shown that this approximation quite accurate (Lemos et al., 2017; Kitching et al., 2017), though this statement will likely change with upcoming, more precise data. Kitching et al. (2017) showed that the flat sky approximation can be partly corrected by setting

$$P_{2D}^{\text{spherical}}(\ell) = \frac{(\ell + 2)(\ell + 1)\ell(\ell - 1)}{\left(\ell + \frac{1}{2}\right)^4} P_{2D}^{\text{flat}}\left(\ell + \frac{1}{2}\right).\quad (9.12)$$

9.2 Galaxy clustering 2-point function and galaxy bias

One situation in which a 2D projection of the density field is observed is the analysis of the galaxy density field on the sky. In photometric surveys (i.e. surveys that don't yield exact redshift information about observed objects, cf. section 8.2) one can, e.g. , observe a sample of galaxies which a certain distribution of redshift values $n_g(z)$. For low values of z the distribution n_g will typically be an increasing function of redshift because the volume of any shell dz increases roughly quadratically with the distance from the observer. Above some turnaround redshift, $n_g(z)$ will then start to decrease with increasing z because the increasingly faint light of distant galaxies falls below a limited detector sensitivity. If the survey depth is limited a cut in the apparent brightness of objects then these two effects are sometimes put into an analytical form for $n_g(z)$ as (Brainerd et al., 1996; Hilbert et al., 2011)

$$n_g(z) = \frac{3z^2}{2z_0^3} \exp\left[-\left(\frac{z}{z_0}\right)^{3/2}\right],\quad (9.13)$$

where z_0 is a free parameter that determines the median redshift of the distribution via $z_0 = z_{\text{median}}/1.412$. Though this analytic expression for n_g provides an intuition for the

typical shape of observed redshift distributions, in real analyses of photometric surveys the redshift distribution is usually estimated from the magnitudes of the observed objects in different chromatic filters (see section 8.2). Also, one usually attempts to split an observed galaxy sample into subsamples with different redshift distributions, e.g. , to study the redshift evolution of 2-point statistics (Elvin-Poole et al., 2017; van Uitert et al., 2018; DES Collaboration et al., 2017).

Given a particular redshift distribution $n_g(z)$ we can determine the corresponding projection kernel along co-moving distance as

$$\begin{aligned} dw q_g(w) &= dz n_g(z) \\ \Rightarrow q_g(w) &= n_g(z(w)) \frac{dz}{dw} \\ &= n_g(z(w)) H(w) . \end{aligned} \quad (9.14)$$

This means that the galaxy density contrast we observe on the sky is given in terms of the 3D galaxy density contrast by

$$\delta_{g,2D}(\hat{\mathbf{n}}) = \int_0^\infty dw n_g(z(w)) H(w) \delta_g(\eta_0 - w, w\hat{\mathbf{n}}) . \quad (9.15)$$

The 2-point correlation function of $\delta_{g,2D}$ is usually denoted with $w(\theta)$, i.e.

$$w(\theta) = \langle \delta_{g,2D}(\boldsymbol{\alpha}) \delta_{g,2D}(\boldsymbol{\alpha} + \boldsymbol{\theta}) \rangle , \quad (9.16)$$

where we have assumed flat coordinates $\boldsymbol{\alpha}$ on the sky (instead of $\hat{\mathbf{n}}$) and used that due to the isotropy of the universe $w(\theta)$ only depends on $\theta = |\boldsymbol{\theta}|$.

Unfortunately, the galaxy density contrast δ_g is not identical to the total matter density contrast δ_m . This is obvious from the fact the galaxies are located at discrete positions and don't form a continuous density field. But even when the galaxy density field is smoothed over scales that are significantly larger than the average galaxy separation, there is still a mismatch between the total density contrast of matter and that of the galaxies. One reason for this is that the process of galaxy formation depends in a complicated, non-linear way on the average matter density in a particular region. Very massive galaxies will, e.g. , preferentially form in dense environments and their formation efficiency is a super-linear function of that environment density. Such galaxies can hence be thought of as amplifying the universe's matter density contrast.

On large scales, the mismatch between galaxy density contrast and matter density contrast can be approximated as a constant multiplicative factor between the two (see chapter 16), i.e.

$$\delta_g \approx b \cdot \delta_m , \quad (9.17)$$

where b is called the (linear) *galaxy bias*. Following the above considerations about massive galaxies, it can be expected, that more massive galaxies have a higher galaxy bias. A simple, heuristic explanation of how this bias arises has, e.g. , been given by Kaiser (1984, the *Kaiser bias model*). We describe this model in figure 9.1 . The simply linear bias

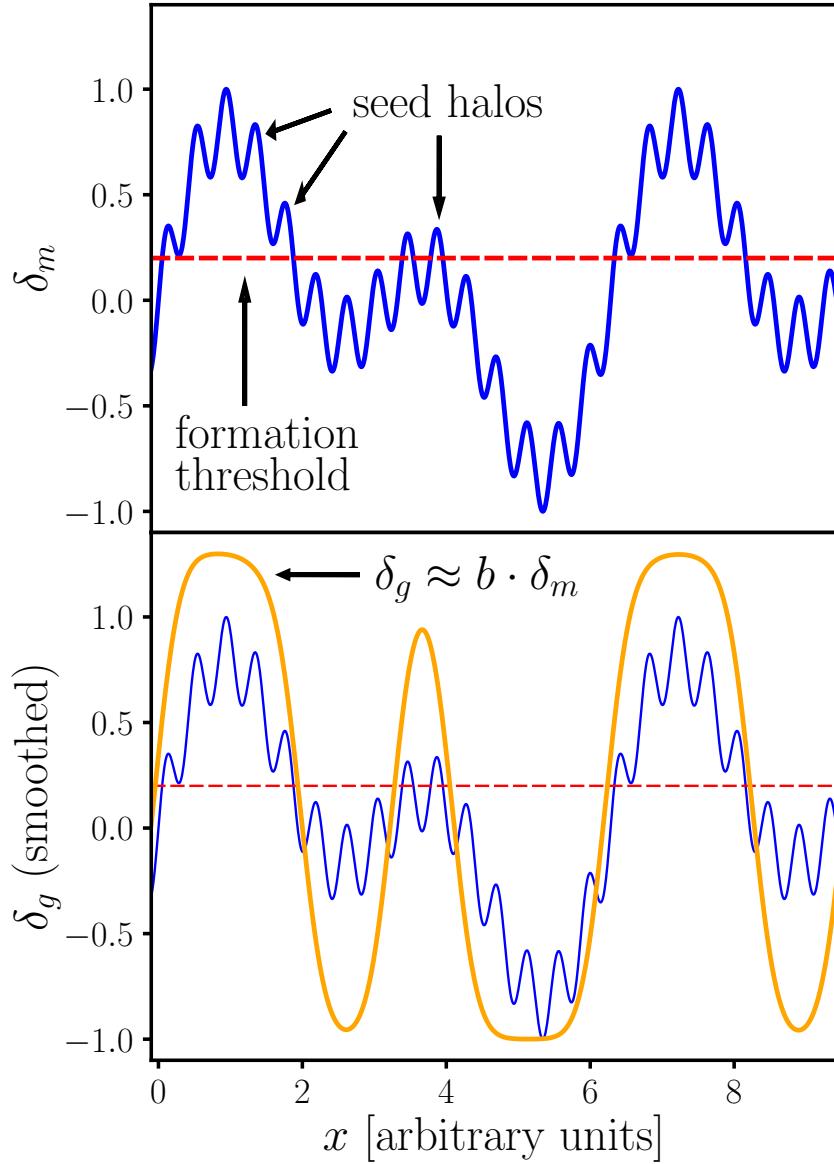


Figure 9.1: An illustration of the so-called Kaiser bias: Galaxies are assumed to form in small density peaks (the peaks in the blue line of the upper panel) when the amplitude of those peaks exceed a certain threshold (the red dashed line in the upper panel). Peaks that are located in large scale overdensities are more likely to exceed this threshold than peaks located in large scale underdensities. This leads to an enhancement of the galaxy density contrast δ_g (orange line in the lower panel) wrt. total matter density contrast δ_m . Averaging both the density and galaxy density field over large enough scales, this enhancement can be described by a multiplicative factor - the linear bias factor b .

assumption of equation 9.17 is only valid on large smoothing scales. On small scales ($\lesssim 10\text{Mpc}$) it must be replaced by non-linear and also stochastic relations between δ_g and δ_m (Dekel & Lahav, 1999; Seljak & Warren, 2004; Cacciato et al., 2012; Uhlemann et al., 2018a; Friedrich et al., 2018). We investigate this point further in chapters 16 and 17 where we present the papers Friedrich, Gruen et al., 2018 and Gruen, Friedrich et al., 2018.

One consequence of galaxy bias is that the 2-point correlation function of δ_g cannot directly be compared to cosmological theory. If on scales where the linear bias assumption is valid, the correlation functions of δ_g and δ_m are still only related via

$$\xi_g \approx b^2 \xi_m \quad (9.18)$$

such that there is no cosmological information in the amplitude of the galaxy correlation function (at least in the absence of a fundamental theory of galaxy formation that predicts the value of b). In addition to that, the bias of a particular galaxy sample might even evolve with redshift (Clerkin et al., 2015) which can obscure observations of the growth of structure from measurements of ξ_g at different redshifts.

Nevertheless, characteristic features in the slope of the 2-point correlation function, that are independent of the amplitude, can be used to determine the expansion history of the universe and to test cosmological models (see the discussion about angular diameter distances in section 8.1). A probe of the 2-point correlation function that is not influenced by galaxy bias are the so-called cosmic shear correlation functions. They are based on the effect of gravitational lensing, which we describe in the next chapter (chapter 10).

Chapter 10

Gravitational lensing

In this chapter we briefly review gravitational lensing, i.e. the deflection of light in a curved spacetime. In section 10.1 we consider the deflection of light paths in a static spacetime using the thin lens approximation. Section 10.4 generalizes this to *cosmic shear*, i.e. to gravitational lensing by the large scale density field in a perturbed Friedmann universe. Our main argument in deriving an expression for the lensing convergence of cosmic shear is a rather heuristic one and we refer the reader to the reviews of Schneider et al. (2002) and Bartelmann & Schneider (2001) and to the paper by Seitz et al. (1994) for a more detailed derivation.

10.1 Gravitational light deflection

Gravitational lensing by a point mass can be described with the (exterior) Schwarzschild metric. The line element of this metric takes the form (Schwarzschild, 1916)

$$ds^2 = \left(1 - \frac{2m}{r}\right) dt^2 - \frac{dr^2}{1 - \frac{2m}{r}} + r^2 (d\theta^2 + \sin^2 \theta d\phi^2) . \quad (10.1)$$

Here the singularity at $r = 0$ is the location of the point mass m and $r_s = 2m$ is the so-called *Schwarzschild radius*. There are no light-like geodesics that lead from inside to outside of this radius.

For $r \rightarrow \infty$ equation 10.1 becomes the line element of flat Minkowsky space. Hence, at large distances from the point mass the trajectories of light rays become straight lines in the Schwarzschild coordinates. This is however not the case in the vicinity of $r = 0$. There, light-like geodesics are deflected from straight coordinate lines which leads to a deflection angle of any light ray that approaches the point mass from infinity and then departs to infinity again. This deflection of light rays by the gravitational field (or rather: by the curvature of spacetime) is called *gravitational lensing* and the mass causing the deflection is called the (gravitational) lens.

To investigate the effects of gravitational lensing, consider figure 10.1. This figure from Bartelmann & Schneider (2001) displays a situation where a gravitational lens is located

somewhere between an observer and a light source. If the observer faces directly toward the lens, two important planes perpendicular to the line-of-sight can be identified: the so-called *lens plane* at the location of the gravitational lens and the *source plane* at the location of the light source. We will denote the distance between observer and lens plane by D_{ol} , the distance between observer and source plane by D_{os} and the distance between lens plane and source plane by D_{ls} .

If the distances D_{ol} and D_{ls} are $\gg r_s$ then the deflection of a light ray that passes the lens can be approximated as a single scattering event that takes place in the lens plane. This is called the *thin lens approximation*. Let ξ be 2D coordinates in the lens plane and η be 2D coordinates in the source plane such that $\eta = 0 = \xi$ is the line of sight of the observer. If light from a source at $|\eta| \ll D_{ls}$ hits the lens plane at a location ξ with $r_s \ll |\xi| \ll D_{ls}$ then it is deflected by an angle

$$\hat{\alpha}(\xi) = 4m \frac{\xi}{|\xi|^2}. \quad (10.2)$$

(The total angle of deflection is given by the absolute value $|\hat{\alpha}|$ but for later convenience we define the deflection angle as a 2-dimensional vector, cf. equation 10.7 below.)

If the lens is not a point mass but an object of finite size then as long as the spatial extend of this object is $\ll D_{ol}, D_{ls}$ we can still employ the thin lens approximation and equation 10.2 generalizes to

$$\hat{\alpha}(\xi) = 4 \int d^2\xi' \Sigma(\xi') \frac{\xi - \xi'}{|\xi - \xi'|^2}. \quad (10.3)$$

Here Σ is the surface mass density of the lens. It is given in terms of the 3D density distribution ρ of the lens as

$$\Sigma(\xi) = \int dz \rho(\xi, z), \quad (10.4)$$

where $\int dz$ means integration along the observers line of sight (i.e. perpendicular to the lens plane).

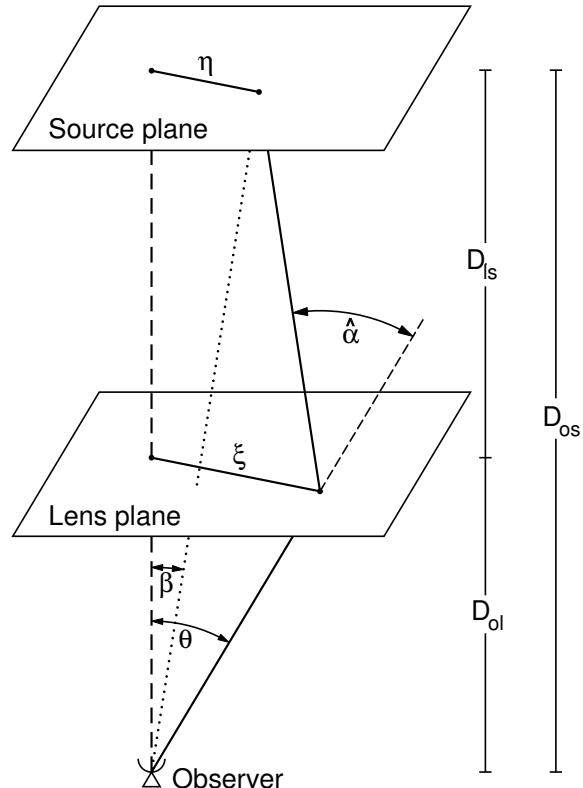


Figure 10.1: Figure from Bartelmann & Schneider (2001), see main text for a description (and note that the notation of the figure was slightly altered compared to the original version of Bartelmann & Schneider).

In order for the light ray originating at $\boldsymbol{\eta}$ in the source plane and hitting the lens plane at $\boldsymbol{\xi}$ to actually reach the observer we must have

$$\frac{\boldsymbol{\eta}}{D_{os}} = \frac{\boldsymbol{\xi}}{D_{ol}} - \frac{D_{ls}}{D_{os}} \hat{\boldsymbol{\alpha}}(\boldsymbol{\theta}) . \quad (10.5)$$

Defining the rescaled deflection angle as

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) := \frac{D_{ls}}{D_{os}} \hat{\boldsymbol{\alpha}}(\boldsymbol{\theta}) \quad (10.6)$$

this can be turned into an equation about angular positions as seen by the observer:

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \boldsymbol{\alpha}(\boldsymbol{\theta}) . \quad (10.7)$$

Here $\boldsymbol{\theta}$ is the observed angular position of the light source and $\boldsymbol{\beta}$ would be its position if the lens was removed. Note that these definitions employ the flat sky approximation, i.e. we assume that all involved angles are small enough to approximate the sky by a 2D plane. Equation 10.7 is called the *lens equation*.

For fixed $\boldsymbol{\beta}$ the lens equation may have several solutions $\boldsymbol{\theta}$. Any light source located at such a position $\boldsymbol{\beta}$ would then be observed at several image locations. Defining the so-called *critical surface mass density*

$$\Sigma_{cr} = \frac{c^2}{4\pi G} \frac{D_{os}}{dD_{ls}} \quad (10.8)$$

one can formulate a sufficient condition for such multiple images to exist: if the surface mass density $\Sigma(\boldsymbol{\xi})$ of a lens decreases faster than $|\boldsymbol{\xi}|^{-2}$ and if $\Sigma(\boldsymbol{\xi}) > \Sigma_{cr}$ somewhere then there exist positions $\boldsymbol{\eta}$ in the source plane for which the lens equation 10.5 (resp. 10.7) has several solutions (cf. section 5.4 of Schneider et al., 1992).

10.2 Lensing potential, convergence and shear

The lens equation defines a mapping $\boldsymbol{\beta}(\boldsymbol{\theta})$ from the lens plane to the source plane. This mapping can be most conveniently studied using a number of quantities that have been constructed by Schneider (1985): the deflection potential, the lensing convergence and the shear field. In the following we will introduce these quantities and discuss their meaning in terms of the lens mapping.

The so called *convergence field* of a thin gravitational lens is defined as

$$\kappa(\boldsymbol{\theta}) := \frac{\Sigma(D_{ol}\boldsymbol{\theta})}{\Sigma_{cr}} . \quad (10.9)$$

The rescaled deflection angle can be expressed in terms of the convergence field as

$$\alpha(\boldsymbol{\theta}) = \frac{1}{\pi} \int d^2\theta' \kappa(\boldsymbol{\theta}') \frac{\boldsymbol{\theta} - \boldsymbol{\theta}'}{|\boldsymbol{\theta} - \boldsymbol{\theta}'|^2} . \quad (10.10)$$

Introducing the *deflection potential*

$$\psi(\boldsymbol{\theta}) = \frac{1}{\pi} \int d^2\theta' \kappa(\boldsymbol{\theta}') \ln(|\boldsymbol{\theta} - \boldsymbol{\theta}'|) \quad (10.11)$$

we see that ψ indeed acts as a potential for the deflection field,

$$\alpha(\boldsymbol{\theta}) = \nabla\psi(\boldsymbol{\theta}) , \quad (10.12)$$

and that κ is the source of this potential,

$$\kappa(\boldsymbol{\theta}) = \frac{1}{2} \Delta\psi(\boldsymbol{\theta}) . \quad (10.13)$$

To see how the lens mapping affects extended light sources, let us approximate $\beta(\boldsymbol{\theta})$ around a location $\boldsymbol{\theta}_0$ in the lens plane as

$$\beta(\boldsymbol{\theta}) \approx \beta(\boldsymbol{\theta}_0) + \mathbf{A}(\boldsymbol{\theta}_0) \cdot (\boldsymbol{\theta} - \boldsymbol{\theta}_0) , \quad (10.14)$$

where $A_{ij} = \partial\beta_i/\partial\theta_j$ is the Jacobi matrix of the mapping. Because the deflection angle can be expressed as the gradient of a potential, the Jacobian is actually a symmetric matrix. This especially means that there is no rotational degree of freedom in the lens mapping! The components of the matrix \mathbf{A} are given by

$$\mathbf{A} = \begin{pmatrix} 1 - \kappa - \gamma_1 & -\gamma_2 \\ -\gamma_2 & 1 - \kappa + \gamma_1 \end{pmatrix} , \quad (10.15)$$

where we have defined

$$\gamma_1 = \frac{1}{2}(\partial_1\partial_1 - \partial_2\partial_2)\psi , \quad \gamma_2 = \partial_1\partial_2\psi . \quad (10.16)$$

These quantities can be combined to the complex number

$$\gamma := \gamma_1 + i\gamma_2 \quad (10.17)$$

which is called the *shear field*. To determine the eigenvalues of \mathbf{A} we solve

$$\begin{aligned} 0 &= (1 - \kappa)^2 + \lambda^2 - 2\lambda(1 - \kappa) - |\gamma|^2 \\ \Rightarrow \lambda_{1,2} &= (1 - \kappa) \pm |\gamma| . \end{aligned} \quad (10.18)$$

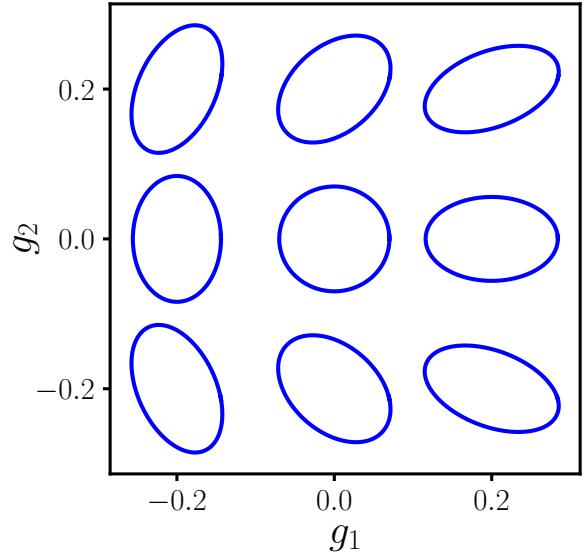


Figure 10.2: The figure shows how spherical light sources are sheared into elliptic images for different values of the reduced shear components g_1 and g_2 (cf. equation 10.21 for their definition). Note that we ignore here the isotropic magnification μ of the images.

The inverse matrix \mathbf{A}^{-1} describes how the light profile of a small (but extended) light source at the position $\beta(\theta_0)$ in the source plane appears around θ_0 in the lens plane. Consider, e.g. , a circular light source that in the absence of lensing would occupy a solid angle $\delta\omega$ on the sky. The mapping by \mathbf{A}^{-1} will produce an image of that light source whose angular area is increased wrt. $\delta\omega$ by a factor of

$$\mu = \det \mathbf{A}^{-1} = \frac{1}{\lambda_1 \lambda_2} = \frac{1}{(1 - \kappa)^2 - |\gamma|^2} . \quad (10.19)$$

This is called the *magnification factor*. In addition to a magnification of the image, gravitational lensing also causes a distortion wrt. the original light profile of the source. The circular source from the above example will, e.g. , be imaged as an ellipse with an axis ratio given by

$$r = \frac{\lambda_2}{\lambda_1} = \frac{1 - \frac{|\gamma|}{(1-\kappa)}}{1 + \frac{|\gamma|}{(1-\kappa)}} . \quad (10.20)$$

Defining the so-called *reduced shear* as

$$g = \frac{\gamma}{1 - \kappa} \quad (10.21)$$

this simplifies to

$$r = \frac{1 - |g|}{1 + |g|} . \quad (10.22)$$

In figure 10.2 we display how a circular light distribution is imaged for different values of the reduced shear components g_1 and g_2 . Our findings about the properties of the lens mapping can be summarized as follows:

- Gravitational lensing introduced a shift between the true and apparent angular positions of light sources. Within the thin lens approximation, this shift is given by the gradient of the deflection potential ψ . The convergence field κ , which is proportional to the surface mass density Σ of the lens, acts as a source for the deflection potential.
- Within the thin lens approximation, gravitational lensing cannot rotate the images of light sources wrt. their original light distribution.
- The images of small but extended light sources are magnified wrt. the original solid angle of the light distribution by a factor of $\mu = 1/((1 - \kappa)^2 - |\gamma|^2)$, where γ is the gravitational shear defined in equation 10.17.
- Small circular light sources are imaged as ellipses with axis ratio $r = (1 - |g|)/(1 + |g|)$ where g is the reduced shear defined in equation 10.21.

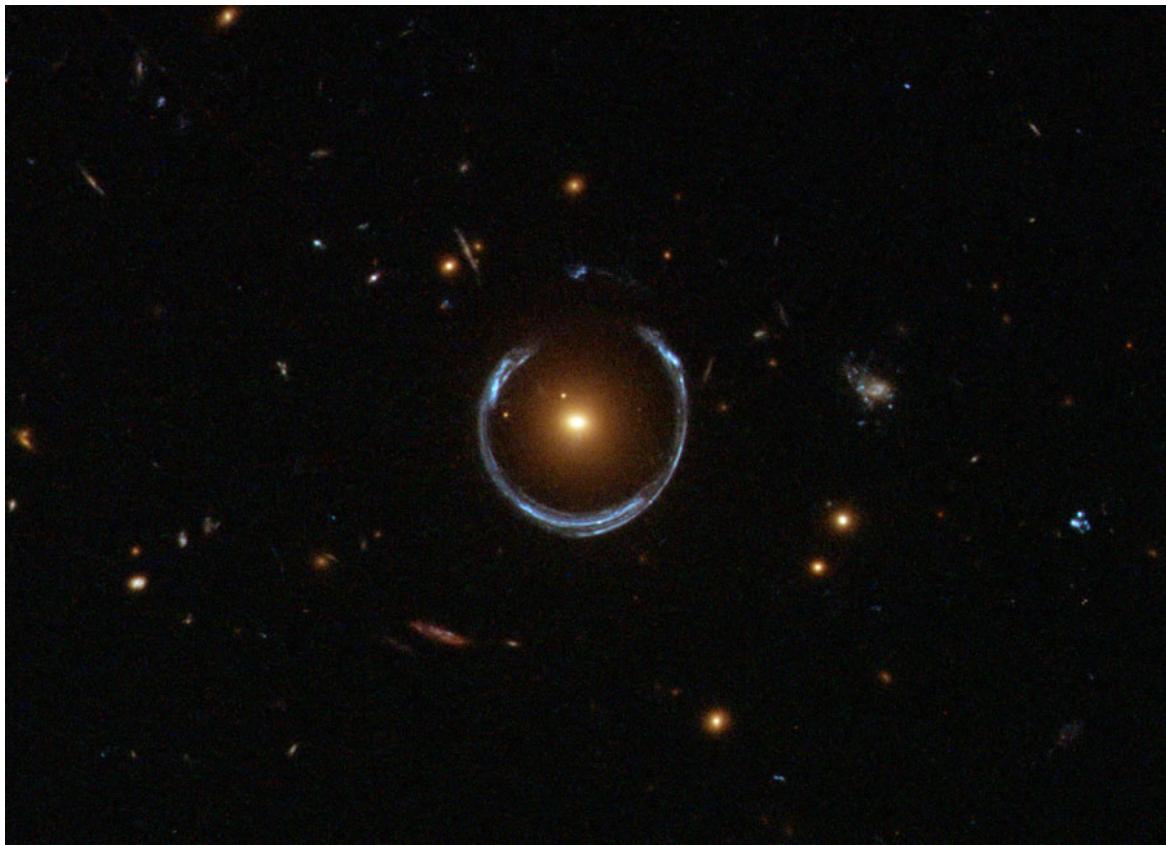


Figure 10.3: The *Horse Shoe Galaxy*, a distant spiral galaxy whose multiple images merge into an almost perfect ring around a massive elliptical foreground galaxy. Such ring-like images are also called Einstein rings. Image credit: Astronomy Picture of the Day, NASA, <https://apod.nasa.gov/apod/ap111221.html> .

10.3 Shapes of background galaxies as tracers of foreground density

As discussed in the previous section, gravitational lensing affects the images of extended light sources in 3 ways: it shifts the image positions (deflection), if magnifies images (magnification) and it distorts the shapes of the images (shear). In the vicinity of massive foreground objects (e.g. massive elliptical galaxies or galaxy clusters) these lensing effects can be quite drastic, i.e. by imaging individual light sources at several distinct image positions or even by distorting light sources into giant luminous arcs (see, e.g. , figure 10.3). The presence of such strong lensing effects indicates the so-called *strong lensing* regime.

Farther away from the gravitational lenses, e.g. , in the outskirts of a galaxy cluster, lensing effects become so weak that they can't be inferred from individual objects. This is, e.g. , the case when the lensing shear of an image is comparable to or smaller than the intrinsic ellipticity of a light source. Assuming that the intrinsic orientations of distant light sources are independent of each other, one can nevertheless infer the large scale shear field by averaging over the ellipticity and orientation of a large sample of light sources. The regime where this is necessary is called the *weak lensing regime* (the reviews of Bartelmann & Schneider, 2001; Schneider, 2005, cover exactly that regime).

In order to average over galaxy ellipticities we need to quantify these ellipticities in a way that can be related to the lensing shear field. One way to do so is to measure the *second order brightness moments* of the light distribution $\mathbf{I}(\boldsymbol{\theta})$ of a galaxy image on the sky (where $\boldsymbol{\theta}$ are again angular coordinates on the sky and $\mathbf{I}(\boldsymbol{\theta})$ is the light intensity measured for the image in different locations; cf. section 2.2 of Schneider 2005, which is the basis for the following derivations). Employing a flat sky approximation we can first determine the image center as

$$\bar{\boldsymbol{\theta}} = \frac{\int d^2\theta \mathbf{I}(\boldsymbol{\theta}) q_l[\mathbf{I}(\boldsymbol{\theta})] \boldsymbol{\theta}}{\int d^2\theta \mathbf{I}(\boldsymbol{\theta}) q_l[\mathbf{I}(\boldsymbol{\theta})]} . \quad (10.23)$$

Here q_l represents a cut-off at low image intensities in order to avoid integration over photon noise. The second order brightness moments are then given by

$$Q_{ij} = \frac{\int d^2\theta \mathbf{I}(\boldsymbol{\theta}) q_l[\mathbf{I}(\boldsymbol{\theta})] (\theta_i - \bar{\theta}_i)(\theta_j - \bar{\theta}_j)}{\int d^2\theta \mathbf{I}(\boldsymbol{\theta}) q_l[\mathbf{I}(\boldsymbol{\theta})]} . \quad (10.24)$$

From the Q_{ij} one can determine an estimate for the ellipticity of the light distribution $\mathbf{I}(\boldsymbol{\theta})$ as

$$\epsilon = \frac{Q_{11} - Q_{22} + 2iQ_{12}}{Q_{11} + Q_{22} + 2(Q_{11}Q_{22} - Q_{12}^2)^{1/2}} . \quad (10.25)$$

As for the reduced lensing shear $g = g_1 + ig_2$, the complex quantity $\epsilon = \epsilon_1 + i\epsilon_2$ describes both the eccentricity and the orientation of the distribution $\mathbf{I}(\boldsymbol{\theta})$ on the sky (cf. figure 10.2). Consider a galaxy for which one has measured an image ellipticity ϵ in the presence of lensing and for which - without lensing - we would have obtained a intrinsic ellipticity

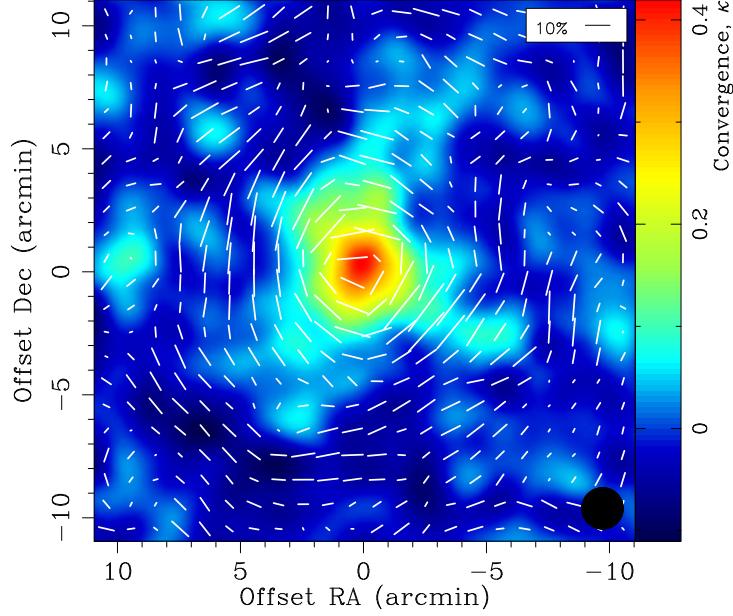


Figure 10.4: Taken from Umetsu (2010). The figure shows the lensing shear field estimated around the cluster Abell 1689 (indicated by the white ticks) as well as the corresponding reconstruction of the convergence field (color coding).

$\epsilon^{\text{int.}}$. Then, in the absence of photon noise in the detector, Seitz & Schneider (1997) have shown that ϵ , $\epsilon^{\text{int.}}$ and the reduced shear g are related by

$$\epsilon = \frac{\epsilon^{\text{int.}} + g}{1 + g^* \epsilon^{\text{int.}}} , \quad (10.26)$$

where $*$ denotes complex conjugation and we also assumed that $|g| < 1$. Seitz & Schneider (1997) also showed that the ensemble average of ϵ over the possible orientations of the intrinsic ellipticity $\epsilon^{\text{int.}}$ is exactly the reduced shear g . However, in real observations there will be detector noise so that equation 10.26 does not hold. Hence, in real reconstructions of the shear field from galaxy shapes one simply accepts the fact that any estimator of the reduced shear is biased and then calibrates this bias. This can, e.g. , be done with the help of image simulations or even by applying artificial shearing to real data (see, e.g. , Kitching et al., 2012; Sheldon & Huff, 2017; Zuntz et al., 2017a, and references therein).

An estimate of the shear field in a region on the sky can, e.g. , be used to reconstruct the large scale mass profiles of individual massive objects (cf. Umetsu, 2010; Gruen et al., 2013, 2014; Rehmann et al., 2018). As an example for that I show a reconstruction of the shear and convergence profile of the galaxy cluster Abell 1689 by Umetsu (2010) in figure 10.4. In contrast to such studies of individual objects, one can also study the large scale statistical properties of the shear field to infer the power spectrum of total matter density fluctuations. I will discuss this technique, called *cosmic shear*, in the following section.

10.4 Cosmic Shear

10.4.1 Lensing by the Cosmos

When we apply gravitational lens theory to lensing by the matter distribution of the whole universe, we are facing two differences compared to the previous section: the thin lens approximation doesn't hold any more and the universe is not static anymore (not even necessarily flat, though we will assume that in this chapter). In the reviews Bartelmann & Schneider (2001) and Schneider (2005) (see also Seitz et al., 1994) it is thoroughly derived, how in that case one can still define an effective convergence κ_{eff} and an effective deflection potential, such that the lensing equation can be put into the same form as in the previous section. I will present only an informal derivation of these results and refer the reader to the mentioned references for a more rigorous argument.

In chapter 5 we saw that the line element of a perturbed Friedmann universe in the presence of only scalar perturbations can be given the form

$$ds^2 = a^2 \left((1 + 2\phi)d\eta^2 - (1 - 2\psi)d\mathbf{x}^2 \right). \quad (10.27)$$

Defining the so-called *Weyl* potential Ψ as

$$\Psi = \frac{\phi + \psi}{2} \quad (10.28)$$

this can be rewritten as

$$ds^2 \approx a^2 e^{\psi - \phi} \left((1 + 2\Psi)d\eta^2 - (1 - 2\Psi)d\mathbf{x}^2 \right), \quad (10.29)$$

where we have ignored terms of higher than linear order in the potentials (which is a very precise approximation on cosmological scales). The line element along light-like geodesics is $ds = 0$. Hence the trajectories of light rays are described by the equation

$$0 \stackrel{!}{=} (1 + \Psi)d\eta^2 - (1 - \Psi)d\mathbf{x}^2. \quad (10.30)$$

Formally, this is identical to line element of light rays in non-expanding space in the Newtonian approximation when Ψ is taken to be the Newtonian potential (cf. Padmanabhan, 2010a). However, in contrast to the non-expanding case, the source of the Weyl potential is not given by the density field $\rho(\eta, \mathbf{x})$. Instead, we have seen in section 5.2.2 that on sub-horizon scales

$$\Delta\Psi \approx \Delta\phi \approx \Delta\psi = 4\pi a^2 \bar{\rho}_m \delta_m, \quad (10.31)$$

where $\bar{\rho}_m$ is the average matter density in the universe and δ_m is the matter density contrast. Hence, the Poisson equation in the non-expanding Newtonian case has to be exchanged by

$$\Delta\phi_{\text{Newton}} = 4\pi\rho \leftrightarrow \Delta\Psi = 4\pi a^2 \bar{\rho}_m \delta_m, \quad (10.32)$$

i.e. $a^2 \bar{\rho}_m \delta_m$ acts as a source for the gravitational potential in an expanding universe filled with cold matter. Let us call $\delta_{\text{conf}} = a^2 \bar{\rho}_m \delta_m$ the *conformal matter density* (which can have

negative values). Using the fact that $\bar{\rho}_m = \bar{\rho}_{m,0}/a^3$ and the fact that the average matter density today is $\rho_{m,0} = 2H_0^2\Omega_{m,0}/8\pi$ this can be written as

$$\delta_{\text{conf}} = \frac{3H_0^2\Omega_{m,0}}{8\pi} \frac{\delta_m}{a} . \quad (10.33)$$

(See section 5.1 for the definitions of H_0 and $\Omega_{m,0}$.) Then, in an infinitesimal shell at co-moving distance w' around an observer we can define a conformal surface density as

$$\begin{aligned} d\Sigma_{\text{conf}}(\boldsymbol{\theta}, w') &= dw' \cdot \delta_{\text{conf}}[w'\boldsymbol{\theta}, w'] \\ &= dw' \cdot \frac{3H_0^2\Omega_0}{8\pi} \frac{\delta_m[f_k(w')\boldsymbol{\theta}, w']}{a(w')} . \end{aligned} \quad (10.34)$$

Note that for simplicity we have parametrized the time dependence of a by the co-moving distance w' along the light trajectory and not by the conformal time $\eta' = \eta_0 - w'$ as we did in chapter 9. The critical density at this shell compared to source galaxies at a co-moving distance w is

$$\Sigma_{\text{crit}}(w', w) = \frac{1}{4\pi} \frac{D_{so}(w)}{D_{lo}(w')D_{sl}(w, w')} = \frac{1}{4\pi} \frac{w}{w'(w - w')} \quad (10.35)$$

and the lensing convergence in that particular shell is then given by

$$d\kappa(\boldsymbol{\theta}, w', w) = \frac{d\Sigma(\boldsymbol{\theta}, w')}{\Sigma_{\text{crit}}(w', w)} . \quad (10.36)$$

If a light source is located at co-moving distance w then the lensing convergence it experiences on its way towards us is given by the sum of the contributions from each thin shell between us and the source. This gives the effective convergence felt by the light from the light source as

$$\begin{aligned} \kappa_{\text{eff}}(\boldsymbol{\theta}, w) &= \int_0^w dw' dk(\boldsymbol{\theta}, w', w) \\ &= \frac{3H_0^2\Omega_0}{8\pi} \int_0^w dw' \frac{1}{\Sigma_{\text{crit}}(w', w)} \frac{\delta_m[w'\boldsymbol{\theta}, w']}{a(w')} \\ &= \frac{3H_0^2\Omega_0}{2} \int_0^w dw' \frac{w'(w - w')}{w} \frac{\delta_m[w'\boldsymbol{\theta}, w']}{a(w')} . \end{aligned} \quad (10.37)$$

Note that we have assumed a flat universe in this derivation. As, e.g. , shown in Bartelmann & Schneider (2001); Schneider (2005) 10.37 can be generalized to a universe with non-zero spatial curvature by replacing

$$\begin{aligned} \frac{w'(w - w')}{w} &\leftrightarrow \frac{\Phi_k(w')\Phi_k(w - w')}{\Phi_k(w)} \\ \delta_m[w'\boldsymbol{\theta}, w'] &\leftrightarrow \delta_m[\Phi_k(w')\boldsymbol{\theta}, w'] , \end{aligned} \quad (10.38)$$

where Φ_k has been defined in equations 5.8 and depends on the spatial curvature of the universe (i.e. the curvature of $\eta = \text{const.}$ slices).

10.4.2 Sources with a redshift distribution and convergence power spectrum

If instead of a single light source at some redshift z we consider an ensemble of source with redshift distribution $p(z)$, then the average convergence field probed by these sources is

$$\kappa_{\text{eff}}(\boldsymbol{\theta}) = \int_0^{w_{\text{max}}} dw \ g(w) \ \kappa_{\text{eff}}(\boldsymbol{\theta}, w) . \quad (10.39)$$

Here $g(w)$ is the distribution in co-moving distance, i.e. $g(w)dw = p(z)dz$. Following Bartelmann & Schneider (2001) this can be reformulated into a projection integral of the matter density contrast δ_m as

$$\kappa_{\text{eff}}(\boldsymbol{\theta}) = \frac{3H_0^2\Omega_0}{2} \int_0^{w_{\text{max}}} dw \ W(w) \ \Phi_k(w) \ \frac{\delta_m[\Phi_k(w')\boldsymbol{\theta}, w]}{a(w)} , \quad (10.40)$$

where

$$W(w) = \int_w^{w_{\text{max}}} dw' \ g(w') \frac{\Phi_k(w' - w)}{\Phi_k(w')} . \quad (10.41)$$

We will from now on drop the subscript eff and denote the effective lensing convergence of a sample of source simply by κ . Equation 10.39 is in fact of the same form as equation 9.3, i.e. κ is a 2D projection of the 3D density contrast with the projection kernel

$$q(w) = \frac{3H_0^2\Omega_0}{2} \ \frac{W(w)\Phi_k(w)}{a(w)} . \quad (10.42)$$

Using our derivation of the Limber approximation in chapter 9 we can then express the convergence power spectrum in terms of the power spectrum of the 3D density contrast as

$$P_\kappa(\ell) = \frac{9H_0^4\Omega_{m,0}^2}{4} \ \int_0^{w_{\text{max}}} dw \ \frac{W(w)^2}{a(w)^2} \ P_\delta\left(w, \frac{\ell}{\Phi_k(w)}\right) . \quad (10.43)$$

(As for the scale factor a we have parametrized the time dependence of P_m here by the co-moving distance w along the light trajectory and not by the conformal time $\eta = \eta_0 - w$ as we did in chapter 9.)

10.4.3 Cosmic shear correlation functions

The lensing convergence is not a direct observable. But as explained in section 10.3, the ellipticities $\epsilon = \epsilon_1 + i\epsilon_2$ of galaxies provide a direct estimate of the reduced shear field $g = g_1 + ig_2$. In this section we will define suitable 2-point statistics of this field, that can be related to the 2-point correlation function and power spectrum of the convergence field κ and hence, via Limber's approximation, to the power spectrum of the 3D matter density contrast δ_m . This section follows the derivations of Schneider et al. (2002).

Consider two galaxies at angular positions $\boldsymbol{\theta}^a$ and $\boldsymbol{\theta}^b$ on the sky and with ellipticities ϵ^a and ϵ^b . Let us denote the separation vector of the two galaxies at $\boldsymbol{\theta}^{ab} = \boldsymbol{\theta}^b - \boldsymbol{\theta}^a$. It is

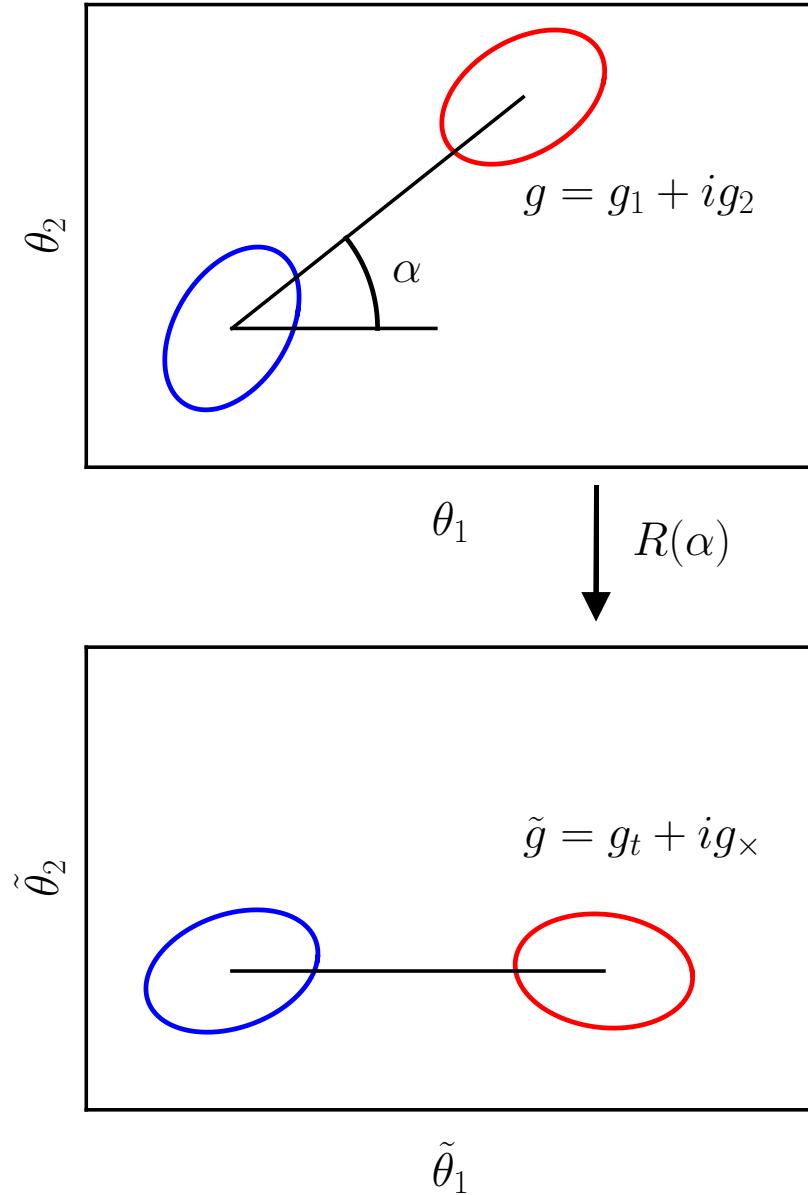


Figure 10.5: The shape g of a galaxy (e.g. the red ellipse in the upper panel) can be decomposed into a tangential and a cross component wrt. a reference point (e.g. wrt. the blue ellipse in the upper panel). To obtain the decomposition one can rotate that angular coordinate system θ_1, θ_2 to new coordinates $\tilde{\theta}_1, \tilde{\theta}_2$ in which the line connecting the galaxy and the reference point is parallel to the $\tilde{\theta}_1$ -axis (lower panel). The shape \tilde{g} in the rotated coordinates is given by $\tilde{g} = g_t + ig_\times$, where g_t is the tangential component and g_\times is the cross component of the galaxy's shape wrt. the reference point.

useful to define the so-called tangential- and cross-components ϵ_t and ϵ_\times of the ellipticities of these galaxies wrt. each other. This is done by rotating our coordinate system on the sky to new coordinates $\tilde{\boldsymbol{\theta}}$ in which the separation vector $\tilde{\boldsymbol{\theta}}^{ab}$ is parallel to the θ_1 -axis (cf. figure 10.5). This rotation will also change the complex ellipticities to new values $\tilde{\epsilon}^a$ and $\tilde{\epsilon}^b$. Using figure 10.2 we can now read off the tangential- and cross-components of the as the 1- and 2-components of the ellipticities in the new coordinate system (Schneider, 2005), i.e.

$$\begin{aligned}\tilde{\epsilon}^a &= -\epsilon_t^a - i\epsilon_\times^a \\ \tilde{\epsilon}^b &= -\epsilon_t^b - i\epsilon_\times^b .\end{aligned}\quad (10.44)$$

It should be stressed again, that these components have no absolute meaning and are only defined for a pair of galaxies. Also, we again employed here the flat sky approximation. It is however not difficult to generalize the above definition of ϵ_t and ϵ_\times to the curved sky (cf. Kilbinger et al., 2013).

Now assume that we have measured the ellipticity ϵ for a large set of galaxies on the sky. For each galaxy pair from that sample we can then define the tangential- and cross-components of their ellipticities as detailed above. Using those we can define the following 2-point statistics:

$$\begin{aligned}\hat{\xi}_{tt}(\theta_{\min}, \theta_{\max}) &= \left(\sum_{\substack{\text{pairs } i,j \\ \theta_{\min} < |\boldsymbol{\theta}^i - \boldsymbol{\theta}^j| < \theta_{\max}}} \epsilon_t^i \epsilon_t^j \right) / N_{\text{pair}}(\theta_{\min}, \theta_{\max}) \\ \hat{\xi}_{\times\times}(\theta_{\min}, \theta_{\max}) &= \left(\sum_{\substack{\text{pairs } i,j \\ \theta_{\min} < |\boldsymbol{\theta}^i - \boldsymbol{\theta}^j| < \theta_{\max}}} \epsilon_\times^i \epsilon_\times^j \right) / N_{\text{pair}}(\theta_{\min}, \theta_{\max}) .\end{aligned}\quad (10.45)$$

Here the sum is over all galaxy pairs whose separation $|\boldsymbol{\theta}^i - \boldsymbol{\theta}^j|$ lies within the interval $(\theta_{\min}, \theta_{\max})$ and $N_{\text{pair}}(\theta_{\min}, \theta_{\max})$ is the number of such galaxy pairs. In analyses of the cosmic shear field one typically considers the linear combinations

$$\begin{aligned}\hat{\xi}_+(\theta_{\min}, \theta_{\max}) &= \hat{\xi}_{tt}(\theta_{\min}, \theta_{\max}) + \hat{\xi}_{\times\times}(\theta_{\min}, \theta_{\max}) \\ \hat{\xi}_-(\theta_{\min}, \theta_{\max}) &= \hat{\xi}_{tt}(\theta_{\min}, \theta_{\max}) - \hat{\xi}_{\times\times}(\theta_{\min}, \theta_{\max}) .\end{aligned}\quad (10.46)$$

Schneider et al. (2002) showed that - as long as the intrinsic ellipticities of the galaxies are independent of each other - the statistics $\hat{\xi}_\pm(\theta_{\min}, \theta_{\max})$ are unbiased estimators of the so-called cosmic shear correlation functions $\xi_\pm(\theta)$ (when the argument θ is averaged over the bin $(\theta_{\min}, \theta_{\max})$). And the latter are related to the convergence power spectrum by

$$\begin{aligned}\xi_+(\theta) &= \frac{1}{2\pi} \int_0^\infty d\ell \ell J_0(\theta\ell) P_\kappa(\ell) \\ \xi_-(\theta) &= \frac{1}{2\pi} \int_0^\infty d\ell \ell J_4(\theta\ell) P_\kappa(\ell) ,\end{aligned}\quad (10.47)$$

where J_n are the Bessel functions of the first kind. Equations 10.47 again make use of the flat sky approximation. Stebbins (1996) has shown how to modify these expressions to an exact curved sky treatment. I summarize the curved sky expressions in appendix B .

At the same angular separation θ the functions $\xi_+(\theta)$ and $\xi_-(\theta)$ are sensitive to different physical scales. This is caused by the different filter functions $J_0(\theta\ell)$ and $J_4(\theta\ell)$ in the integrals on the right hand-side of 10.47. In figure 10.6 I explain which alignments of source galaxies typically give the largest contribution to each of the correlation functions and what configurations of lenses and sources cause such alignments. Also, I display the functions $J_0(x)$ and $J_4(x)$ in figure 10.7 to demonstrate how $\xi_+(\theta)$ and $\xi_-(\theta)$ for the same value of θ put weight on different modes ℓ in $x = \theta\ell$.

Measurements of the cosmic shear correlation function can be used to the theoretical predictions about the shape and amplitude of the matter power spectrum without the need to model the bias between galaxies and the total matter density field (cf. section 9.2). Also, measuring $\hat{\xi}_\pm$ for source galaxy samples at different ranges of redshift allows to study the time evolution of the power spectrum and hence to quantify the growth of structure in the universe. By now there have been a number of cosmological analyses based on cosmic shear (Kilbinger et al., 2013; Heymans et al., 2013; Hildebrandt et al., 2017; Troxel et al., 2017) or on combining cosmic shear correlation functions and other 2-point statistics of the large scale structure (van Uitert et al., 2018; DES Collaboration et al., 2017). In chapter 11 I will briefly review the results presented in Troxel et al. (2017) and DES Collaboration et al. (2017) which were obtained within the Dark Energy Survey.

10.4.4 Galaxy-galaxy lensing

Another way to probe the matter power spectrum with gravitational lensing is to measure the so called *galaxy-galaxy lensing* correlation function. To do so one considers 2 different galaxy samples: galaxies at a high redshift range acting as light sources (the *source galaxies*) and galaxies at a lower redshift range tracing the foreground matter that is producing the lensing effect (the *lens galaxies*). Given a lens galaxy at position $\boldsymbol{\theta}^l$ and a source galaxy at position $\boldsymbol{\theta}^s$ with ellipticity $\epsilon^s = \epsilon_1^s + i\epsilon_2^s$ we can define the component ϵ_t^s of the ellipticity that is tangential wrt. the position of the lens galaxy in the same way as described in the previous section (cf. figure 10.5).

Given a set of source and lens galaxies the galaxy-galaxy lensing correlation function γ_t in a finite angular bin $(\theta_{\min}, \theta_{\max})$ can then be estimated as

$$\begin{aligned} \hat{\gamma}_t(\theta_{\min}, \theta_{\max}) &= \left(\sum_{\substack{\text{pairs } l,s \\ \theta_{\min} < |\boldsymbol{\theta}^l - \boldsymbol{\theta}^s| < \theta_{\max}}} \epsilon_t^s \right) \Big/ N_{\text{lens-source pair}}(\theta_{\min}, \theta_{\max}) \\ &- \left(\sum_{\substack{\text{pairs } r,s \\ \theta_{\min} < |\boldsymbol{\theta}^r - \boldsymbol{\theta}^s| < \theta_{\max}}} \epsilon_t^s \right) \Big/ N_{\text{random-source pair}}(\theta_{\min}, \theta_{\max}) , \end{aligned} \quad (10.48)$$

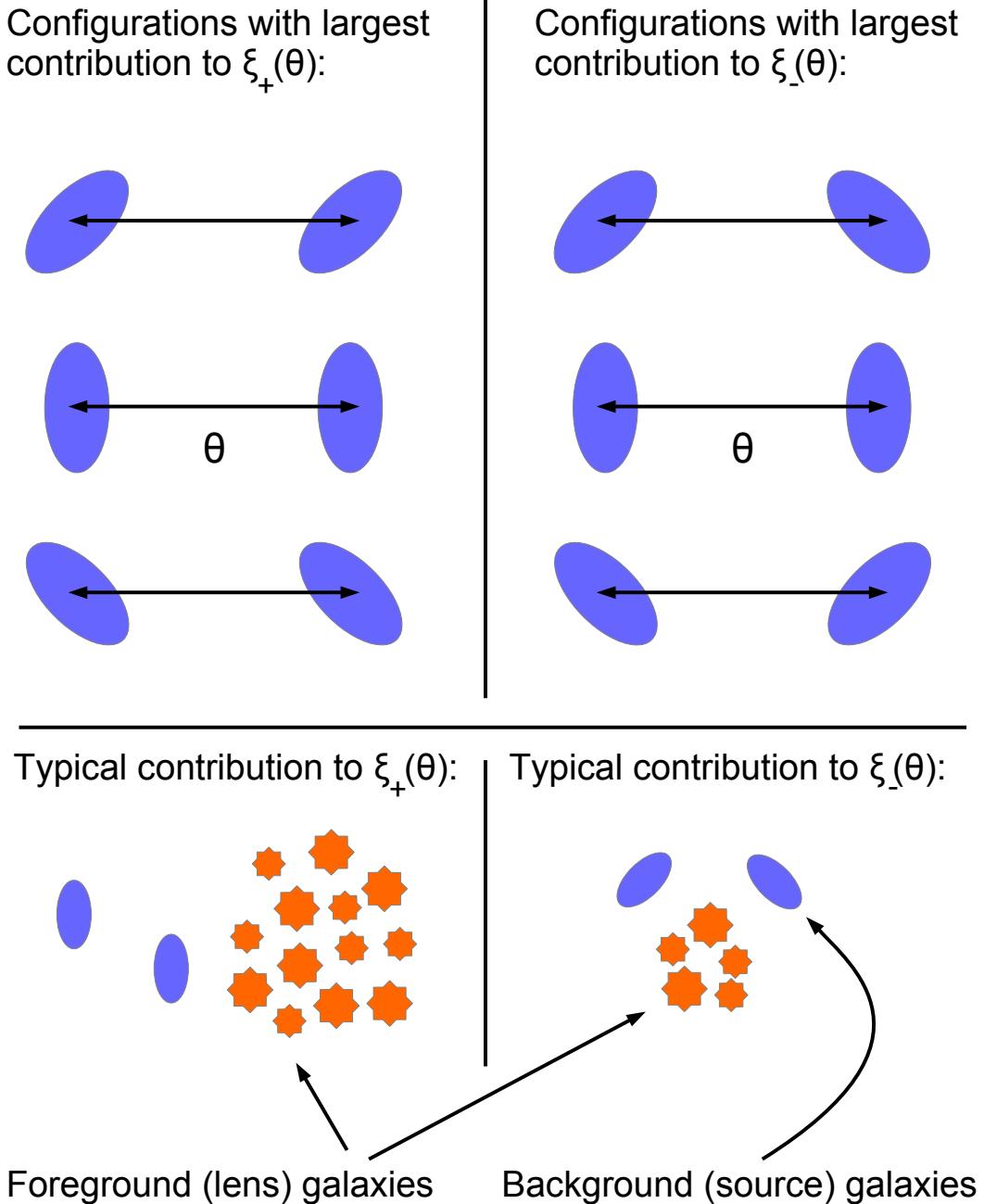


Figure 10.6: The largest contributions to the correlation function $\xi_+(\theta)$ come from source galaxy pairs that have a parallel alignment on the sky (upper left panel). The largest contributions to $\xi_-(\theta)$ come from source galaxy pairs whose alignments are mirrors of each other along the line connecting the two galaxies (upper right panel). At the same angular distance θ on the sky the function $\xi_+(\theta)$ typically probes larger physical structures than the function $\xi_-(\theta)$ (cf. lower left and right panels).

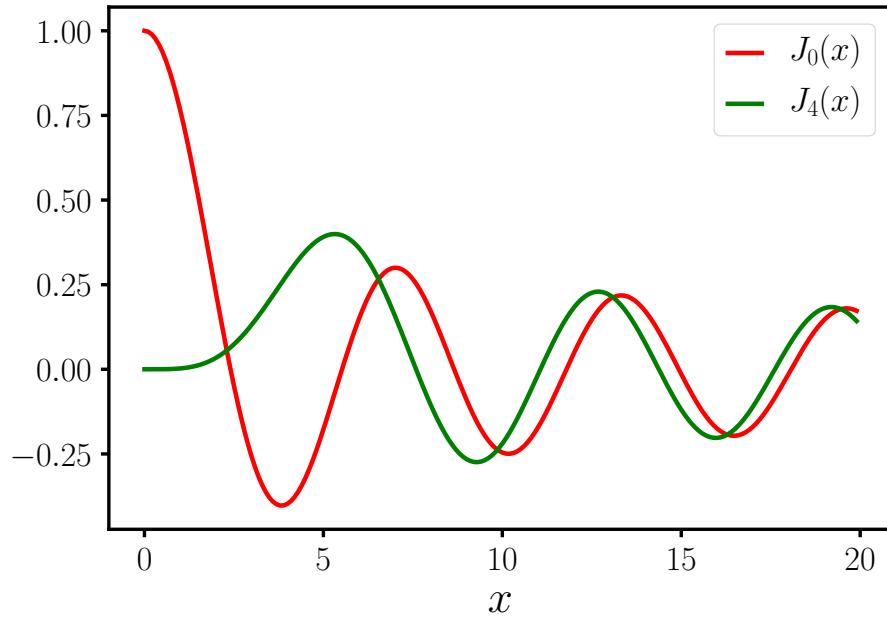


Figure 10.7: The filter functions J_0 and J_4 that appear in the relations between the cosmic shear correlation functions $\xi_{\pm}(\theta)$ and the convergence power spectrum $P_{\kappa}(\ell)$. At a fixed angular scale θ the filter $J_0(x) = J_0(\theta\ell)$ puts more weight on low frequencies ℓ than the filter $J_4(x) = J_4(\theta\ell)$. As a consequence, the function $\xi_+(\theta)$ is more sensitive to large scale density fluctuations than the function $\xi_-(\theta)$ at the same angular scale θ .

where the first sum is over all lens-source pairs whose angular distance lies within the angular bin $(\theta_{\min}, \theta_{\max})$ and in the second some the lens positions are replaced by a set of random points that are uniformly distributed across the sky (or rather: the part of the sky that is observed). When ensemble averaging over the intrinsic ellipticities of the source galaxies the second term actually averages to 0, but including it help to reduce the noise coming from field-to-field variations of the density field (the so-called *cosmic variance*, see, e.g. , Prat et al. 2017).

If the intrinsic ellipticities of source galaxies are independent of each other, then in a narrow angular bin around θ we have $\langle \hat{\gamma}_t(\theta) \rangle = \gamma_t(\theta)$ where $\gamma_t(\theta)$ is given by

$$\gamma_t(\theta) = \frac{1}{2\pi} \int_0^\infty d\ell \ell J_2(\theta\ell) P_{\kappa,g}(\ell) . \quad (10.49)$$

Here $P_{\kappa,g}(\ell)$ is the cross power spectrum between the lensing convergence field kappa and the 2D galaxy density contrast $\delta_{2D,g}$ (cf. section 9.2). Assuming a linear bias, $\delta_{2D,g} = b\delta_{2D,m}$, we can again use Limber's approximation to express $P_{\kappa,g}(\ell)$ in terms of the 3D matter power spectrum as

$$P_{\kappa,g}(\ell) = \frac{3H_0^2\Omega_{m,0}}{2} b \int_0^{w_{\max}} dw \frac{W(w)q_g(w)}{a(w)^2} P_\delta \left(w, \frac{\ell}{\Phi_k(w)} \right) . \quad (10.50)$$

Here $W(w)$ is the lensing kernel defined in equation 10.41 and $q_g(w)$ is the distribution of values of co-moving distance among the lens galaxy sample (cf. the notation in section 9.2).

As can be seen from equation 10.50 in the linear bias model $P_{\kappa,g}(\ell)$ and $\gamma_t(\theta)$ proportional to the linear galaxy bias b . This is in contrast to the 2D galaxy clustering correlation function $w(\theta)$ which was proportional to b^2 (cf. section 9.2). As a consequence of that, a combined measurement of w and γ_t breaks the degeneracy between galaxy bias and the amplitude of the matter power spectrum and can hence be used to test theoretical predictions for the growth of density fluctuations (see, e.g. , Mandelbaum et al., 2013; Kwan et al., 2017). The sketch in figure 10.8 provides a summary of the 2-point correlation functions of the cosmic large scale structure that were introduced in sections 9.2, 10.4.3 and 10.4.4: the galaxy clustering correlation function $w(\theta)$, the cosmic shear correlation functions $\xi_\pm(\theta)$ and the galaxy-galaxy-lensing correlation function $\gamma_t(\theta)$.

10.4.5 Intrinsic alignments

The estimators defined in 10.46 only give unbiased measurements of the cosmic shear correlation functions ξ_\pm if the intrinsic shapes of close by source galaxies are independent of each other. It has however been expected that galaxies forming in similar large scale tidal fields are correlated in their alignment. I am not discussing this topic further in this work and instead refer the reader to the papers Hirata et al. (2004); Troxel & Ishak (2015); Krause et al. (2016); Hilbert et al. (2017); Blazek et al. (2017) and the references therein.

In recent analyses of lensing by the large scale structure the correlation of intrinsic galaxy alignments has been incorporated into the models of the measured 2-point statistics

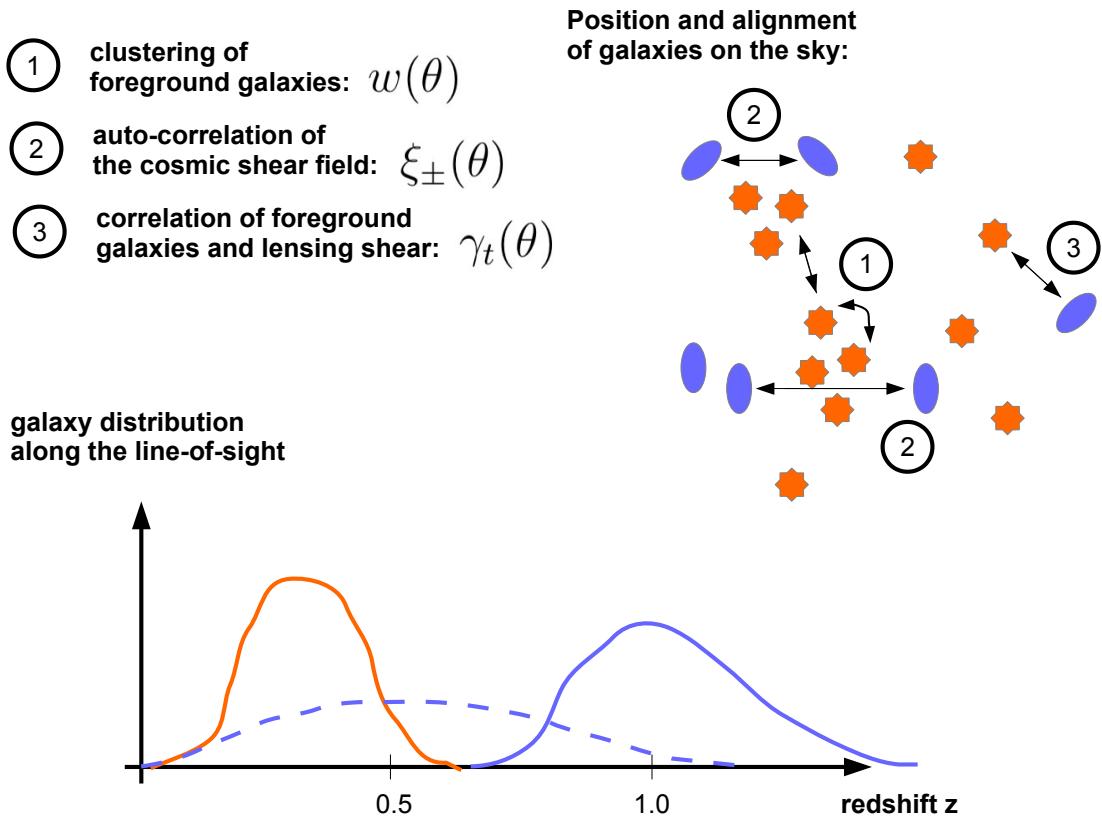


Figure 10.8: Summary of the 2-point statistics of the large scale structure relevant for this work. The shapes in the upper right part of the sketch indicate the position of lens galaxies (orange) and the position and alignment of source galaxies (blue). Note that I ignore here the fact that galaxies have an intrinsic ellipticity and assume the shape of the source galaxies is entirely dictated by the lensing shear field. The figure in the lower part of the sketch represents typical redshift distributions of lens galaxy samples (solid orange line) and source galaxy samples (solid blue line) in use in cosmological studies today. Note however, that modern surveys simultaneously analyze a number of lens and source redshift bins (Heymans et al., 2013; van Uitert et al., 2018; Hildebrandt et al., 2017; Prat et al., 2017; Elvin-Poole et al., 2017; Troxel et al., 2017; DES Collaboration et al., 2017). The dashed blue line indicated the typical shape of the lensing kernel W defined in equation 10.41 that would correspond to a source redshift distribution given by the solid blue line.

(see van Uitert et al., 2018; Köhlinger et al., 2017; Troxel et al., 2017, and combinations thereof with other 2-point statistics). In these works only mild detections of the amplitude of shape correlations have been obtained but notably, the studies of van Uitert et al. (2018) and Köhlinger et al. (2017) find an opposite sign of this correlation even though they are based on the same data set. In that context it has also been found that errors in photometric redshift estimates (cf. section 8.2) can to some degree mimic the signal expected from intrinsic alignments (Efstathiou & Lemos, 2018).

Chapter 11

The Dark Energy Survey

The *Dark Energy Survey* (DES) is an optical imaging survey that will eventually cover 5000deg^2 of the southern sky and will provide redshift and shape estimates for galaxies up to $z \sim 1.5$. It is run by the DES collaboration (see <https://www.darkenergysurvey.org/> and Flaugher 2005) and is carried out at the Victor M. Blanco 4-meter Telescope at the Cerro Tololo Inter-American Observatory (CTIO) in Chile. It images the sky with the *Dark Energy Camera* (DECam, see Flaugher et al. 2015) using five different optical bands (see the left panel of figure 11.1 taken from DES 2018b). The average seeing (i.e. the blurring of images due to atmospheric turbulences) differs from band to band but is overall located around 1 arcsecond (see the right panel of figure 11.1 which is also taken from DES 2018b).

The DES data enable a variety of scientific studies from solar system research to analyses of the large scale structure of the universe (see DES 2016 for an overview). A main objective of DES is to quantify the growth of structure in the low-redshift universe in order to test theoretical predictions for the evolution of cosmic density fluctuations. Especially, it is supposed to test whether the same parameters of the Λ CDM model that are needed to describe the primary anisotropies of the Cosmic Microwave Background (i.e. its temperature and polarization fluctuations which originate from the universe at $z \approx 1100$) can also describe the large scale structure of the low-redshift universe. In DES (2018a) (and a number of companion papers, see references in DES 2018a) the DES collaboration has published intermediate results from that effort based on its year-1 data (DES Y1, data taken between August 2013 and December 2014). These results were obtained by analyzing measurements of the galaxy clustering 2-point correlation function $w(\theta)$, the cosmic shear correlation functions $\xi_{\pm}(\theta)$ and the galaxy-galaxy lensing correlation function $\gamma_t(\theta)$ that have been introduced in sections 9.2, 10.4.3 and 10.4.4. These functions were measured using overall 5 different lens galaxy samples and 4 different source galaxy samples. In figure 11.2 I show a plot from DES (2018a) presenting the redshift distributions of these samples. The measurements of the different 2-point correlation functions in the different bins are shown in figures 11.3 and 11.4 (also taken from DES 2018a).

I was involved in this effort through the following contributions:

- designing and carrying out tests of the analytic covariance matrix of the 2-point

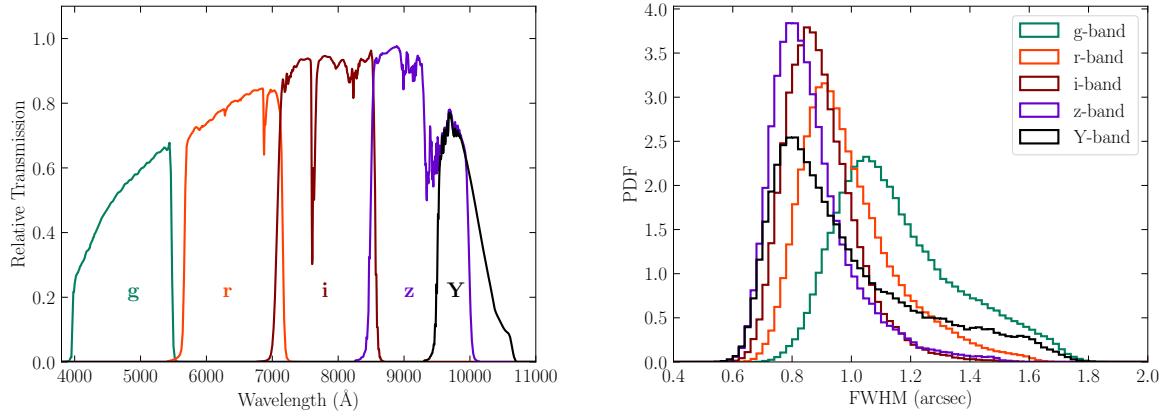


Figure 11.1: Taken from DES 2018b who present the public release of the DES year 1 data. The left panel shows the relative transmission of the 5 filters g, r, i, z, Y used for DES observations as a function of wavelength. The right panel displays the distribution of seeing conditions (i.e. widening of images due to atmospheric turbulence) in the different bands.

functions measurements that was used to draw cosmological conclusions from those measurements (see section III.B of Krause et al. 2017),

- supervising and supporting the creation of mock data based on the **FLASK** simulation tool (by Xavier et al., 2016) for tests of systematic uncertainties (see Krause et al., 2017; Prat et al., 2017; Troxel et al., 2017, for use of that data).

The analysis presented by DES (2018a) is the first one that was able to measure the parameters of the Λ CDM model to a precision comparable to that previously achieved through analyses of primary CMB anisotropies (Planck Collaboration et al., 2016). They find that the Λ CDM model describes the DES measurements of $w(\theta)$, $\xi_{\pm}(\theta)$ and $\gamma_t(\theta)$ well (cf. figure 11.3 and 11.4) and that there is no significant tension with the results obtain by Planck Collaboration et al. (2016). This especially means that current data of the large scale structure of the universe is well described by a model with a constant dark energy density resp. cosmological constant. Note that I don't show the results of that paper here because a proper discussion thereof is beyond the scope of this thesis.

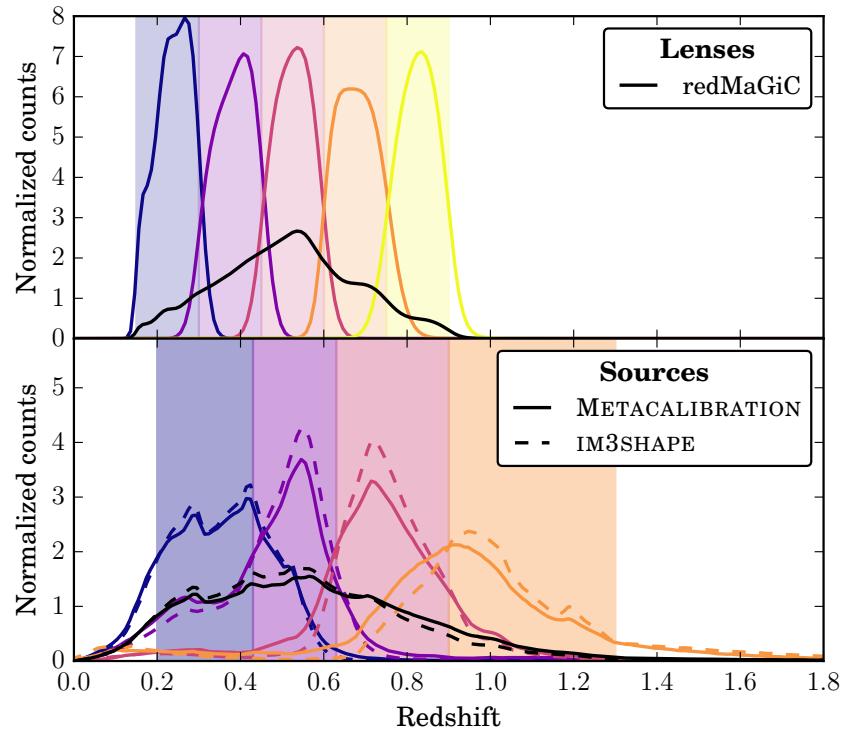


Figure 11.2: Taken from DES Collaboration et al. (2017) who present a cosmological analysis based on 2-point function measurements in DES year 1 data. The upper panel shows the redshift distributions of the 5 lens redshift bins used in their analysis. The lower panel shows the redshift distributions of their 4 source redshift bins (see Hoyle et al., 2018, for how the latter have been obtained).

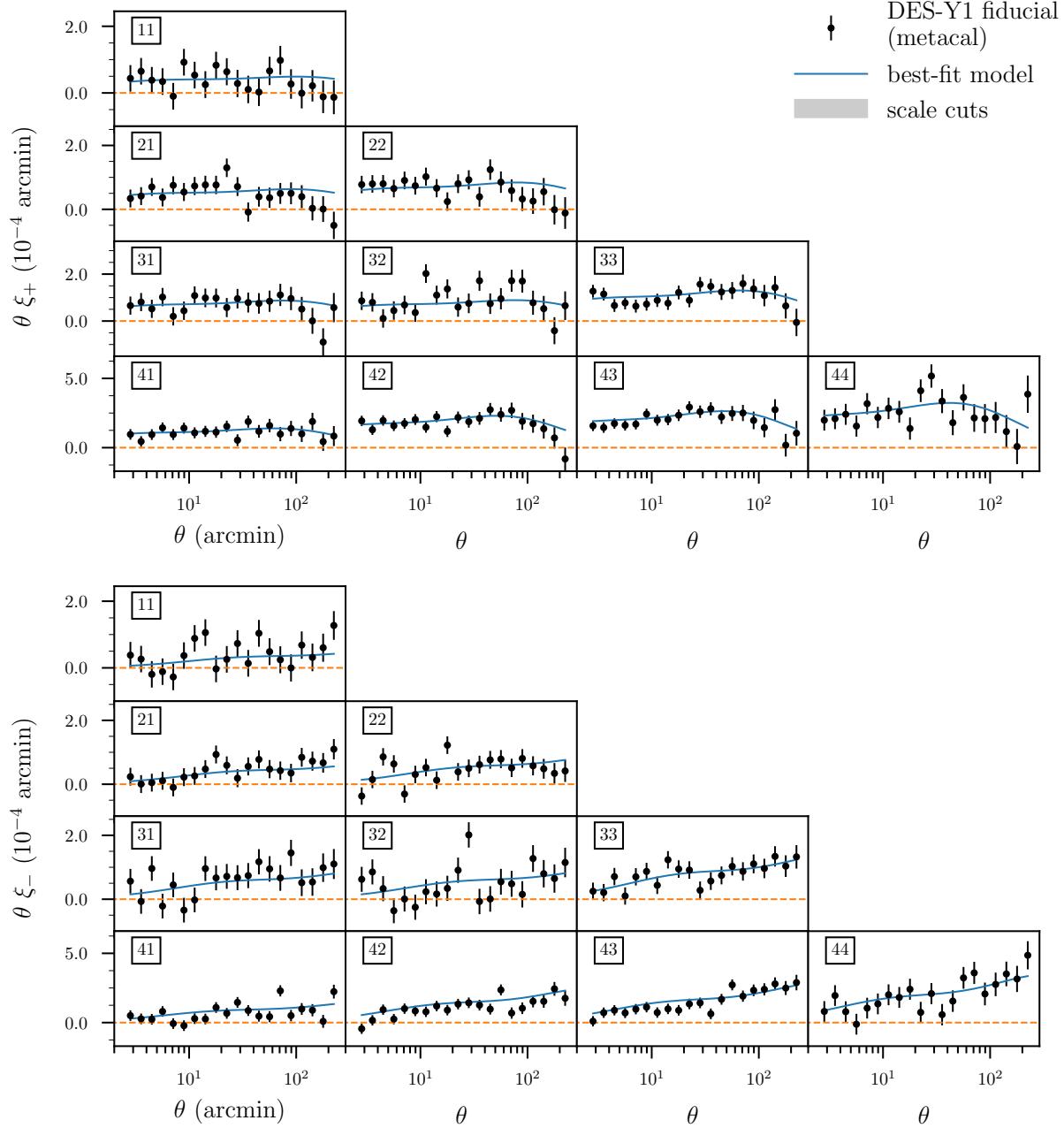


Figure 11.3: Taken from DES Collaboration et al. (2017) who present a cosmological analysis based on 2-point function measurements in DES year 1 data. The figure shows their measurements of the cosmic shear correlations functions $\xi_{\pm}(\theta)$ (cf. section 10.4). The solid blue line shows their best fit Λ CDM prediction. Gray areas have been cut from the analysis because of modeling uncertainties in the matter power spectrum and the galaxy bias model (see also Elvin-Poole et al., 2017).

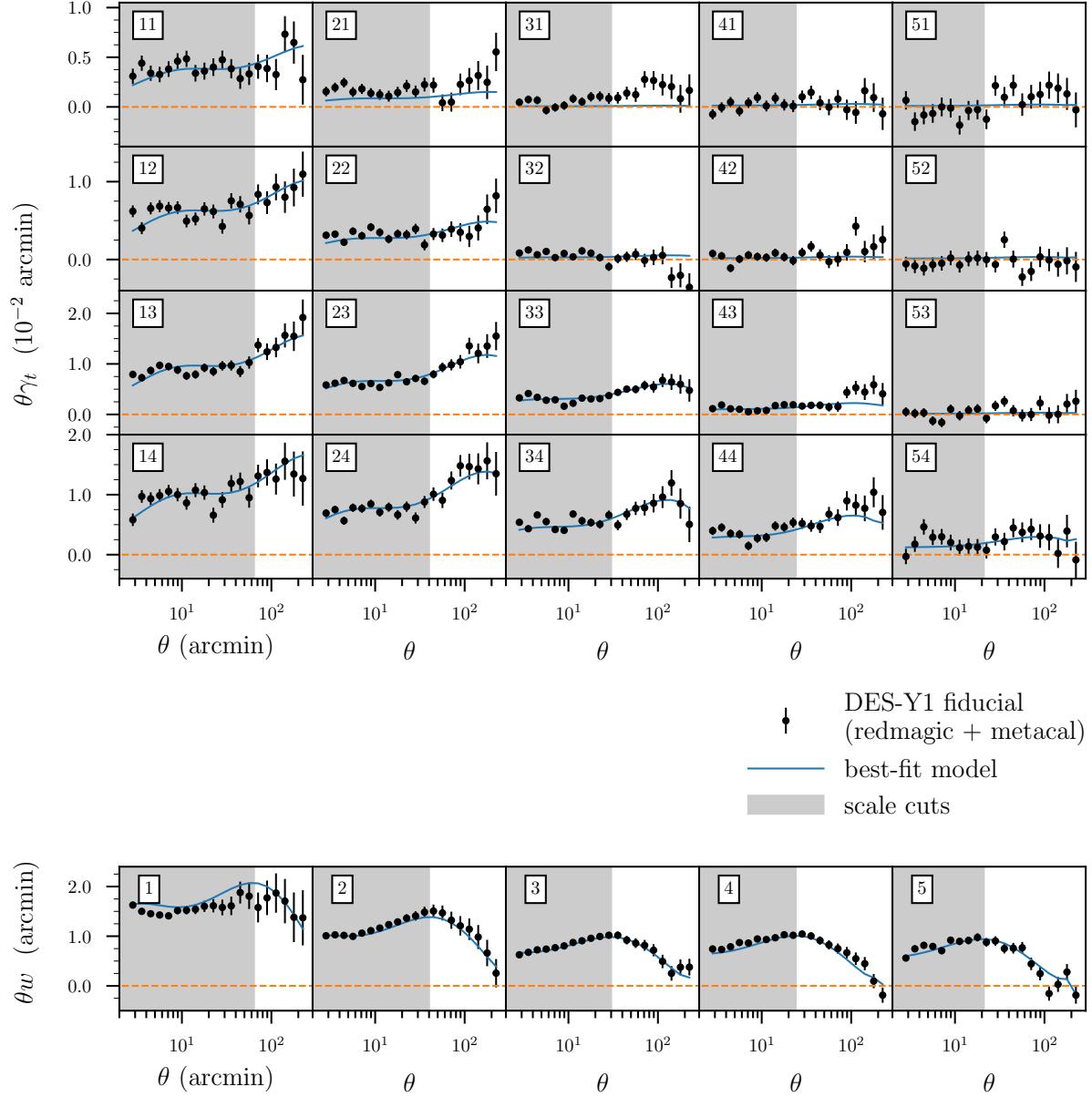


Figure 11.4: Taken from DES Collaboration et al. (2017) who present a cosmological analysis based on 2-point function measurements in DES year 1 data. The figure shows their measurements of the galaxy-galaxy lensing correlation function $\gamma_t(\theta)$ in different combinations of lens and source redshift bins as well as their measurements of the galaxy clustering correlation function $w(\theta)$ in different lens bins (cf. sections 9.2 and 10.4). The solid blue line shows their best fit Λ CDM prediction. Gray areas have been cut from the analysis because of modeling uncertainties in the matter power spectrum and the galaxy bias model (see also Elvin-Poole et al., 2017).

Part III

Estimation of Covariance Matrices for 2-point correlation functions

Chapter 12

Covariance matrices and the derivation of parameter constraints

This work is about higher order moments (moments of order > 2) of the cosmic density field. In this part of the thesis (chapters 12 to 14) we will encounter a 4th order statistic of the density field - the *covariance matrix* of 2-point correlation functions and power spectra.

A main objective of measuring the 2-point statistics $\xi_{\pm}(\theta)$, $\gamma_t(\theta)$ and $w(\theta)$ introduced in previous chapters is to test theoretical predictions for these data vectors. If a cosmological model is parametrized by a vector of parameters $\boldsymbol{\pi}$, then testing this model means to examine whether different measurements agree with predictions from that model for the same values of the parameters $\boldsymbol{\pi}$. In this chapter I briefly review different methods to decide whether a measured data vector agrees with a theoretical prediction and how this can be used to discriminate reasonable choices for the parameters $\boldsymbol{\pi}$ from unreasonable ones. To carry out such a discrimination one typically requires knowledge of the covariance matrix. Chapters 13 and 14 investigate different methods for estimating this matrix. Previous chapters that are required to understand this part of my thesis are

- Chapter 5 - Cosmology in General Relativity
(Especially section 5.1 for the definition of the Λ CMD model and its parameters.)
- Chapter 6 - Theory of Cosmic Structure Formation
(Especially sections 6.1 and 6.2 for the definition of the matter power spectrum and 2-point correlation function as well as the definition of the Λ CMD parameter σ_8 .)
- Chapter 9 - 2D Power Spectra and Correlation Functions
- Chapter 10 - Gravitational Lensing

12.1 Covariance matrices

Consider a random vector $\hat{\mathbf{x}}$ that consists of D data points. To say that $\hat{\mathbf{x}}$ that is drawn from a D -dimensional probability distribution function (PDF) $p(\mathbf{x})$ means to say that the

probability of finding $\hat{\mathbf{x}}$ within a volume $dV_{\mathbf{x}}$ around \mathbf{x} with size $|dV| = dx^D$ is given by

$$P(\hat{\mathbf{x}} \in dV) = dx^D p(\mathbf{x}) . \quad (12.1)$$

Please note that I have not specified here, what I mean by *probability*. This will be discussed further below.

Given the PDF $p(\mathbf{x})$ we can define the expectation value $\boldsymbol{\mu}$ of $\hat{\mathbf{x}}$ as

$$\begin{aligned} \boldsymbol{\mu} &= \langle \hat{\mathbf{x}} \rangle \\ &= \int dx^D \mathbf{x} p(\mathbf{x}) . \end{aligned} \quad (12.2)$$

This integral does not necessarily converge, but I will assume so in the following. Given the expectation value we can then define the covariance matrix \mathbf{C} of $\hat{\mathbf{x}}$ as

$$\begin{aligned} \mathbf{C} &= \langle [\hat{\mathbf{x}} - \boldsymbol{\mu}] \cdot [\hat{\mathbf{x}} - \boldsymbol{\mu}]^T \rangle \\ &= \int dx^D [\mathbf{x} - \boldsymbol{\mu}] \cdot [\mathbf{x} - \boldsymbol{\mu}]^T p(\mathbf{x}) . \end{aligned} \quad (12.3)$$

Here $[\cdot]^T$ means transposition and \cdot means matrix multiplication. For the individual components of the covariance matrix the definition of equation 12.3 reads

$$C_{ij} = \int dx^D [x_i - \mu_i][x_j - \mu_j] p(\mathbf{x}) . \quad (12.4)$$

From both definitions it can be seen that \mathbf{C} is a symmetric matrix, i.e. $\mathbf{C} = \mathbf{C}^T$.

Data vectors measured in cosmological analyses are often assumed to be *Gaussian* random vectors (Kilbinger et al., 2013; Heymans et al., 2013; Abbott et al., 2016; Hildebrandt et al., 2017; van Uitert et al., 2018; DES Collaboration et al., 2017). This means that one considers these measurements to be realizations of a random process with the PDF

$$p_{\text{Gauss.}}(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^D |\mathbf{C}|}} \exp \left\{ -\frac{1}{2} [\mathbf{x} - \boldsymbol{\mu}]^T \cdot \mathbf{C}^{-1} \cdot [\mathbf{x} - \boldsymbol{\mu}] \right\} . \quad (12.5)$$

Here \mathbf{C}^{-1} is the inverse of the covariance and $|\mathbf{C}|$ is its determinant. As for the cosmic shear correlation function one would, e.g., identify $\hat{\mathbf{x}}$ with $\hat{\xi}_{\pm}(\theta_{\min}, \theta_{\max})$ measured in a number of D angular bins $(\theta_{\min}, \theta_{\max})$ (or $D/2$ angular bins if both $\hat{\xi}_+$ and $\hat{\xi}_-$ are combined into one data vectors) and the expectation value $\boldsymbol{\mu}$ would correspond to the analytic functions ξ_{\pm} evaluated in these bins.

The assumption that measurements of cosmological 2-point functions have a Gaussian distribution can be justified with the central limit theorem (Anderson, 2003): 2-point statistics measured within a large area on the sky can be considered as averages of equivalent measurements carried out on sub-patches of that area. And independent of what the distribution of these sub-measurements is, their average should tend to a Gaussian distribution on the limit of a large total area. Note however, that is has recently been questioned whether the Gaussian assumption is accurate enough for current cosmic shear surveys (Sellentin & Heavens, 2018; Sellentin et al., 2018). This question is still undecided and is currently under investigation.

12.2 Frequentist approach on constraining model parameters

It is far beyond the scope of this work to discuss the differences (and dispute?) between the Bayesian and Frequentist approaches to statistics in sufficient detail. I will hence refer the interested reader to references such as Neyman (1937); Abroe et al. (2002); Yèche et al. (2006); Hobson et al. (2014); Trotta (2017) and will discuss here only my personal viewpoint on the subject. Especially, I will introduce a definition of frequentist parameter estimation that slightly differs from that adopted, e.g. , in Abroe et al. (2002); Yèche et al. (2006) and that in my opinion prevents Bayes-like features (cf. criticism in section 12.2.3).

12.2.1 Choice of the figure-of-merit as a contract between scientists

If the structure in the universe indeed originated from quantum fluctuations in the Inflaton field, then any data vector $\hat{\mathbf{x}}$ that was ever measured in any experiment can be legitimately considered as the results of a random process. And the ensemble of possible outcomes of any experiment is well defined through the ensemble of possible initial quantum fluctuations (+ the constraint that the experiment was indeed conducted with those initial conditions). I will hence in the following take any data vector $\hat{\mathbf{x}}$ to be a realization of a random process with PDF $p(\mathbf{x})$.

To define a program of frequentism we need the following assumptions:

- We are given a model $p(\mathbf{x}|\boldsymbol{\pi})$ for the PDF $p(\mathbf{x})$ which depends on a number of model parameters $\boldsymbol{\pi}$.
- For some values $\boldsymbol{\pi}_0$ of the parameters we have $p(\mathbf{x}|\boldsymbol{\pi}_0) \equiv p(\mathbf{x})$ (see section 12.2.3 for criticism of this assumption).

Frequentism can then be considered as a contract between scientists in which they

- agree on a scalar quantity

$$\hat{\mathcal{F}} [\hat{\mathbf{x}}, p(\cdot|\boldsymbol{\pi})] \quad (12.6)$$

that is to be formed out of the measurement $\hat{\mathbf{x}}$ and the model for the PDF of $\hat{\mathbf{x}}$ (the *figure-of-merit*).

- agree on a criterion to reject parameters values $\boldsymbol{\pi}$ based on the value of $\hat{\mathcal{F}}$. This criterion should be chosen such that if $\boldsymbol{\pi} = \boldsymbol{\pi}_0$ the probability of rejecting $\boldsymbol{\pi}$ has a fixed value α .

Note that our model $p(\mathbf{x}|\boldsymbol{\pi})$ for the PDF of $\hat{\mathbf{x}}$ will also lead to a model $p(\mathcal{F}|\boldsymbol{\pi})$ for the PDF of $\hat{\mathcal{F}}$. The criterion in the second bullet point above can, e.g. , be based on that PDF. Now, if all scientists conform to this contract, then the parameter values $\boldsymbol{\pi}_0$ will be rejected in a fraction α of all experiments ever conducted.

12.2.2 Prior knowledge and model parametrization

Assume that we have the prior information that certain parameters in the parameter vector $\boldsymbol{\pi}$ can only take values within specific ranges (e.g. we might know from theoretical considerations, that the mass of some particle cannot be negative, etc.). This prior knowledge can be incorporated in the above contract by simply rejecting all parameter values outside the ranges

$$\begin{aligned} \pi_{1,\min} < \pi_1 < \pi_{1,\max} \\ \pi_{2,\min} < \pi_2 < \pi_{2,\max} \\ \dots < \dots < \dots , \end{aligned} \tag{12.7}$$

where $(\pi_{i,\min}, \pi_{i,\max})$ is the allowed range for the parameter π_i .

Let N_p be the number of parameters. Then within the allowed ranges the parameters describe an N_p -dimensional manifold. A very convenient feature of the frequentist contract is the following: whether or not a point on the parameter manifold is rejected is independent of the particular set of coordinates chosen on the manifold (i.e. the particular parametrization we have chosen for our model of the PDF $p(\mathbf{x}|\boldsymbol{\pi})$). This is because our criterion for rejecting a point only depends on our data vector $\hat{\mathbf{x}}$ and our model PDF $p(\mathbf{x}|\boldsymbol{\pi})$ - i.e. it is a local criterion on this manifold. This feature is also not spoiled by the above way of incorporating prior knowledge.

12.2.3 Criticism

Even though individual experiments (such as observations of the large scale structure of the universe) cannot be repeated, the frequentist contract nevertheless allows a well defined interpretation of scientific results as a whole. This is through the statement that a fraction α of those experiments have rejected model parameters that correctly describe the PDF $p(\mathbf{x})$.

However, there are a number of practical reasons which (at least partially) exclude frequentism from an application in cosmology. I list them in the following.

- As a premise underlying the frequentist contract I assumed that there exists a point $\boldsymbol{\pi}_0$ in our model manifold that describes the distribution of $\hat{\mathbf{x}}$ perfectly. But no model is a perfect representation of reality. Even in principle, current cosmological models fail this assumption since model parameters such as the galaxy bias (cf. section 9.2) are from the beginning only nuisance parameters that are only meant to approximate more complex phenomena (such as the process of galaxy formation).
- It is questionable whether a universal scalar $\hat{\mathcal{F}}[\hat{\mathbf{x}}, p(\cdot|\boldsymbol{\pi})]$ and a universal rejection criterion can be defined that can be applied to all possible experiments. However, this may not be a severe problem since different criteria and different figures-of-merit could be defined for different types of data vectors.

- The frequentist program cannot be used to derive marginalized constraints on only a subset of the model parameters $\boldsymbol{\pi} = (\pi_1, \pi_2, \dots)$. It can only neglect or not neglect the parameter vector $\boldsymbol{\pi}$ as a whole and the rejection probability α only applies to that entire vector and not to subsets of it. There are cosmological studies that try to construct frequentist confidence intervals for individual parameters in a multi-parametric model (e.g. Abroe et al., 2002; Yèche et al., 2006) but they implicitly assume a (Bayesian) prior PDF for the remaining parameters. This is necessary because it is generally not possible to construct a model for the PDF of $\hat{\mathbf{x}}$ of the form $p(\mathbf{x}|\pi_1)$ from a model $p(\mathbf{x}|\boldsymbol{\pi})$ without assuming a specific distribution for the values of the parameters π_2, π_3, \dots . This has already been stated by Neyman (1937) who developed some of the foundations of frequentist parameter estimation.

The last point is in fact the most severe point of critique since in cosmology we are interested in constraining parameters such as Ω_m *after* marginalizing over a reasonable ensemble of values for nuisance parameters such as the galaxy bias b (and not in 2D constraints on the vector $[\Omega_m, b]$).

12.3 Bayesian approach on constraining model parameters

12.3.1 Probability as degree of personal belief

Given a data vector $\hat{\mathbf{x}}$ and a model $p(\mathbf{x}|\boldsymbol{\pi})$ for its PDF Bayesian statistics attempts to derive PDF $p(\boldsymbol{\pi}|\hat{\mathbf{x}})$ on the model manifold as parametrized by the parameters $\boldsymbol{\pi}$. For any volume V in the parameter space this PDF will determine the probability that $\boldsymbol{\pi}$ is indeed located in that volume.

Assigning probabilities to different ranges of the model parameters does not make sense from a frequentist perspective. This is because from the frequentist point of view there is only one true PDF of the data $p(\mathbf{x})$ and only a fixed set of parameter combinations $\boldsymbol{\pi}_0$ will provide exactly that true prediction $p(\mathbf{x}|\boldsymbol{\pi}_0) \equiv p(\mathbf{x})$. However, the definition of probability in Bayesian statistics is different from that in frequentist statistics. For Bayesianism, probability assignments over a set of statements \mathcal{D} are just assignments of real numbers $P(E)$ to certain subsets $E \subset \mathcal{D}$ such that $P(\cdot)$ fulfills the so-called Kolmogorov axioms (Anderson, 2003; Hobson et al., 2014; Trotta, 2017). The Bayesian interpretation of \mathcal{D} and $P(\cdot)$ is that \mathcal{D} represents a set of elementary and mutually exclusive statements and that $P(E)$ represents the degree of belief (of someone) that the true statement s is included in the subset $E \subset \mathcal{D}$ (Hobson et al., 2014; Trotta, 2017).

An important feature of Bayesian statistics is that any experiment can only update an already existing belief. Assume that a measurement $\hat{\mathbf{x}}$ was conducted to estimate the parameters $\boldsymbol{\pi}$. Then the Bayesian program assigns the so-called *a posteriori* PDF to the parameters as

$$p(\boldsymbol{\pi}|\hat{\mathbf{x}}) = \frac{p(\mathbf{x}|\boldsymbol{\pi}) \text{ pr}(\boldsymbol{\pi})}{\mathcal{N}} . \quad (12.8)$$

Here the PDF $p(\mathbf{x}|\boldsymbol{\pi})$ of the data vector is also called the *likelihood* of the parameters and $pr(\boldsymbol{\pi})$ is an a priori PDF that represent the belief *before* the experiment was carried out (the so-called *prior*). The factor \mathcal{N} ensures that $p(\boldsymbol{\pi}|\hat{\mathbf{x}})$ is normalized upon integration over $\boldsymbol{\pi}$ and is called the *evidence*. The assignment of the a posteriori PDF by equation 12.8 can be derived from *Bayes theorem* and from the fact that the Bayesian degree of belief is supposed to follow the Kolmogorov axioms (Hobson et al., 2014; Trotta, 2017).

12.3.2 Criticism

The main criticism of Bayesian statistics come from the fact that there is no obvious way to translate someones prior belief into a prior PDF $pr(\boldsymbol{\pi})$ for the parameter values and from the fact the a posteriori PDF $p(\boldsymbol{\pi}|\hat{\mathbf{x}})$ depends on how exactly that translation has been done (Abroe et al., 2002; Yèche et al., 2006). It is customary to simply use so-called flat priors (Kilbinger et al., 2013; Hildebrandt et al., 2017; Troxel et al., 2017; Gruen et al., 2018, and many other studies). This means that for a given parameter π_i one determines reasonable lower and upper boundaries $\pi_{i,\min}$ and $\pi_{i,\max}$ such that when fixing all other parameters we get

$$\begin{aligned} pr(\boldsymbol{\pi}) &= \text{const.} & \text{for } \pi_i \in (\pi_{i,\min}, \pi_{i,\max}) \\ &= 0 & \text{else.} \end{aligned} \tag{12.9}$$

This procedure is however highly ambiguous: if we choose to parametrize the model manifold with different parameters $\tilde{\boldsymbol{\pi}}$ that are non-linear functions of the old parameters, then a flat prior in the old parameters $\boldsymbol{\pi}$ will be a non flat one on the new parameters $\tilde{\boldsymbol{\pi}}$. Hobson et al. (2014); Trotta (2017) provide references to several different recipes for choosing a sensible prior. In this work we employ two different kind of priors: for some parameters we assume a flat prior distribution between certain fixed minimum and maximum values and for some we assume Gaussian prior distributions around a value of highest prior confidence (see chapters 16, 17 and 14 for details).

Chapter 13

Performance of internal covariance estimators for cosmic shear correlation functions

As explained in the previous chapter, if the noise of a measured data vector can be considered to have a Gaussian distribution, then the covariance matrix of that noise is everything we need to know in order to test the agreement of that data vector with theoretical predictions. In the following chapter I will investigate the easiest, but also the most messy way to get the covariance matrix: estimating it from the data itself.

The present chapter has been published as Friedrich, Seitz, Eifler & Gruen (2016) in MNRAS. I exclusively performed the analysis in this article and also developed the numerical tools to generate the mock data sets used in this study. Tim Eifler provided the cosmology tool package `CosmoLike` that was used to perform the simulated likelihood analyses. All authors contributed through discussions and proofreading.

Permission for non-commercial re-use of the material included in this thesis has been granted by the MNRAS editorial office. Oxford University Press holds the copyright on the paper.

ABSTRACT

Data re-sampling methods such as delete-one jackknife, bootstrap or the sub-sample covariance are a common tool for estimating the covariance of large scale structure probes. We investigate different implementations of these methods in the context of cosmic shear two-point statistics. Using log-normal simulations of the convergence field and the corresponding shear field we carry out realistic tests of internal covariance estimators. For a survey of $\sim 5000 \text{ deg}^2$ we find that jackknife, if implemented in what we call the *galaxy-scheme*, provides the most reliable covariance estimates. Bootstrap, in the common implementation of duplicating sub-regions of galaxies, strongly overestimates the statistical uncertainties.

In a forecast for the complete 5-year DES survey we show that internally estimated covariance matrices can provide a large fraction of the true uncertainties on cosmological parameters in a 2D cosmic shear analysis. The volume inside contours of constant likelihood in the Ω_m - σ_8 plane as measured with internally estimated covariance matrices is on average $\gtrsim 85\%$ of the volume derived from the true covariance matrix. The uncertainty on the parameter combination $\Sigma_8 \sim \sigma_8 \Omega_m^{0.5}$ derived from internally estimated covariances is $\sim 90\%$ of the true uncertainty.

13.1 Introduction

Two-point statistics of cosmological random fields such as the cosmic shear correlation functions or the galaxy clustering angular correlation function are common probes of the large scale structure of the universe. Recent measurements of these correlation functions are e.g. reported in Thomas et al. (2011); Kilbinger et al. (2013); de Simoni et al. (2013); Becker et al. (2016). In order to use these statistics for constraining cosmological models one needs a quantitative description of the joint distribution of the correlation function estimators. When assuming multivariate Gaussian errors, this is given by the covariance matrix. On large angular scales this covariance matrix can - both for cosmic shear and galaxy clustering - be well described by a Gaussian approximation for the involved fields (Schneider et al., 2002; Crocce et al., 2011). It has, however, been shown, that the Gaussian approximation fails to describe the true PDF of the weak lensing convergence field (Taruya et al., 2002; Vale & White, 2003) and that it underestimates the true covariance of the cosmic shear correlation functions on small scales, which can be alleviated by an empirical re-scaling (Semboloni et al., 2007; Sato et al., 2011), a log-normal approximation (Hilbert et al., 2011), or halo model approaches (e.g. Cooray & Hu, 2001; Takada & Jain, 2009; Eifler et al., 2014).

Alternatives to modeling the covariance matrix are to estimate it from many independent realizations of cosmological N-body simulations or to estimate it internally, i.e. from the data itself. The latter method is independent of assuming a particular cosmological model and is hence often used to complement the other methods (Kilbinger et al. 2013; Wang et al. 2013; Becker et al. 2016.).

So far the performance of internal covariance estimators has only been systematically studied for the galaxy clustering 2-pt function (in most detail by Norberg et al. 2009) or for cross-correlations of the Cosmic Microwave Background (CMB) and the galaxy field (Cabré et al., 2007). In our paper, we will concentrate on cosmic shear correlation functions. We will show that the shape noise part of the covariance can be very accurately estimated internally while the cosmic variance part is generally underestimated. Gaussian simulations of the convergence field hence yield an overly optimistic test of internal covariance estimators, since the Gaussian model underpredicts the cosmic variance contribution to the covariance. We overcome this problem by employing log-normal simulations of the convergence field.

In our paper we want to study the performance of internal covariance estimators such as bootstrap, jackknife or the sub-sample covariance. There is no complete agreement in the literature yet on whether internal covariance estimates can be used to constrain cosmological parameters from measured 2pt-correlations or whether they are a mere tool to generate reasonable errorbars in plots of correlation functions (see e.g. Norberg et al. 2009; Wang et al. 2013; de Simoni et al. 2013; Taylor et al. 2013). We want to address the questions of how many internal re-samplings are required in order to get a stable covariance matrix, whether internal estimators over- or underestimate the covariance matrix and whether/how internal covariance estimates can yield unbiased estimates of the *inverse* covariance matrix.

Our paper is organized as follows: In section 13.2 we introduce the cosmic shear correlation functions and explain the Gaussian and the log-normal model for the covariance of 2-pt. function estimators. In section 13.3 we describe the simulations we use to generate mock shape catalogs that follow any given input power spectrum and whose underlying convergence field has a log-normal PDF. These are the simulations with which we will test the performance of internal covariance estimators.

In section 13.4 we introduce two distinct ways of performing jackknife estimation of the covariance of two-point measures - the pair-jackknife and the galaxy-jackknife. Furthermore, we are explaining why jackknife, bootstrap and subsample covariance are almost equivalent.

In section 13.5 we apply internal covariance estimators to simulated cosmic shear surveys. We show that in the pair-scheme all estimators are almost identical and we demonstrate the systematic effects of the different estimation schemes when varying the number of re-samplings. Our method to find optimal estimation schemes has to be re-run for any specific survey, because the performance of internal estimators depends crucially on the depth and area of a survey. In the end of section 13.5 we configure our simulations to match the complete, 5-year Dark Energy Survey (DES, The Dark Energy Survey Collaboration 2005; Flaugher 2005) and test the accuracy of jackknife covariance matrices for this particular setting. The code used for our simulations is made publicly available¹.

In section 13.6 we discuss the results of our work.

¹www.usm.uni-muenchen.de/people/oliverf/, the code also contains many other useful features, that e.g. enable the user to create mock data suitable for galaxy-galaxy lensing or galaxy clustering measurements.

13.2 Cosmic Shear Basics

13.2.1 Cosmic Shear Correlation Functions

Cosmic shear measures the correlated distortion of galaxy shapes due to gravitational lensing by the large scale structure of the universe as a function of the angular distance of galaxy pairs on the sky. We follow here the notation of Schneider et al. (2002) and employ the flat-sky-approximation, i.e. we assume a tangential Cartesian coordinate system $\boldsymbol{\vartheta} = (\vartheta_1, \vartheta_2)$ on the sky.

In this coordinate system the cosmic shear field is at each point characterized by a complex number $\boldsymbol{\gamma}(\boldsymbol{\vartheta}) = \gamma_1 + i\gamma_2$. If the separation vector $\Delta\boldsymbol{\vartheta} = \boldsymbol{\vartheta}_2 - \boldsymbol{\vartheta}_1$ of two points on the sky has the polar angle ϕ then the *tangential* and *cross* components of $\boldsymbol{\gamma}$ at $\boldsymbol{\vartheta}_2$ and $\boldsymbol{\vartheta}_1$ (with respect to each other) are defined as

$$\gamma_t = -\mathcal{R}\text{e}(\boldsymbol{\gamma}e^{-2i\phi}) \quad ; \quad \gamma_{\times} = -\mathcal{I}\text{m}(\boldsymbol{\gamma}e^{-2i\phi}) . \quad (13.1)$$

The *cosmic shear correlation functions* $\xi_{\pm}(\theta)$ are defined as the expectation values

$$\xi_{\pm}(\theta) = \langle \gamma_{t,1}\gamma_{t,2} \rangle \pm \langle \gamma_{\times,1}\gamma_{\times,2} \rangle , \quad (13.2)$$

where θ is the absolute value of $\Delta\boldsymbol{\vartheta}$. It can be computed in terms of the power spectrum $P_{\kappa}(\ell)$ of the scalar *convergence field* $\kappa(\boldsymbol{\vartheta})$ as

$$\xi_{\pm}(\theta) = \frac{1}{2\pi} \int d\ell \ell P_{\kappa}(\ell) J_{0/4}(\ell\theta) , \quad (13.3)$$

where $J_0(x)$ ($J_4(x)$) is the 0-th order (4-th order) Bessel function.

The shape of a galaxy can be characterized by a complex number $\boldsymbol{\epsilon}$ which is to first order the sum of the intrinsic shape $\boldsymbol{\epsilon}^{\text{in}}$ of the galaxy and the distortion caused by gravitational lensing, i.e. the value $\boldsymbol{\gamma}(\boldsymbol{\vartheta})$ at the location $\boldsymbol{\vartheta}$ of the galaxy,

$$\boldsymbol{\epsilon} = \boldsymbol{\epsilon}^{\text{in}} + \boldsymbol{\gamma} . \quad (13.4)$$

In a cosmic shear survey the shapes $\boldsymbol{\epsilon}_i$ of many galaxies are measured and (cf. Schneider et al. 2002) an estimator for the correlation function can be constructed as

$$\hat{\xi}_{\pm}(\theta) = \frac{\sum_{ij} w_i w_j (\epsilon_{t,i} \epsilon_{t,j} \pm \epsilon_{\times,i} \epsilon_{\times,j}) \Delta_{\theta}(ij)}{\sum_{ij} w_i w_j \Delta_{\theta}(ij)} , \quad (13.5)$$

where we have allowed for some weighting scheme w_i for the shape measurements and where the filter $\Delta_{\theta}(ij)$ selects all galaxy pairs (i, j) in the survey whose angular separation lies in some finite bin around θ . The normalization in equation 13.5 is the effective number of galaxy pairs in a bin around θ , which we will abbreviate as

$$N_p(\theta) = \sum_{ij} w_i w_j \Delta_{\theta}(ij) . \quad (13.6)$$

13.2.2 Covariance of the Correlation Functions

The covariance matrix of the estimator in equation 13.5 is defined as

$$\begin{aligned} C_{\pm, \pm}(\theta_1, \theta_2) &= \langle (\hat{\xi}_{\pm}(\theta_1) - \xi_{\pm}(\theta_1))(\hat{\xi}_{\pm}(\theta_2) - \xi_{\pm}(\theta_2)) \rangle \\ &= \langle \hat{\xi}_{\pm}(\theta_1) \hat{\xi}_{\pm}(\theta_2) \rangle - \xi_{\pm}(\theta_1) \xi_{\pm}(\theta_2) . \end{aligned} \quad (13.7)$$

In order to compute this covariance matrix it is convenient to split $\xi_{\pm}(\theta)$ into the three different contribution

$$\begin{aligned} \hat{\xi}_{\pm}^{nn}(\theta) &= \frac{\sum_{ij} w_i w_j (\epsilon_{t,i}^{\text{in}} \epsilon_{t,j}^{\text{in}} \pm \epsilon_{\times,i}^{\text{in}} \epsilon_{\times,j}^{\text{in}}) \Delta_{\theta}(ij)}{N_p(\theta)} , \\ \hat{\xi}_{\pm}^{ss}(\theta) &= \frac{\sum_{ij} w_i w_j (\gamma_{t,i} \gamma_{t,j} \pm \gamma_{\times,i} \gamma_{\times,j}) \Delta_{\theta}(ij)}{N_p(\theta)} , \\ \hat{\xi}_{\pm}^{sn}(\theta) &= \frac{\sum_{ij} w_i w_j (\epsilon_{t,i}^{\text{in}} \gamma_{t,j} \pm \epsilon_{\times,i}^{\text{in}} \gamma_{\times,j}) \Delta_{\theta}(ij)}{N_p(\theta)} \end{aligned} \quad (13.8)$$

which are the autocorrelation of the intrinsic shape noise, the autocorrelation of the shear signal and their cross correlation. The whole estimator 13.5 is given in terms of these as

$$\hat{\xi}_{\pm}(\theta) = \hat{\xi}_{\pm}^{nn}(\theta) + \hat{\xi}_{\pm}^{ss}(\theta) + 2 \cdot \hat{\xi}_{\pm}^{sn}(\theta) .$$

Under the assumption that the shear signal and the shape noise are independent of each other it is obvious that

$$\langle \hat{\xi}_{\pm}^{nn}(\theta_1) \hat{\xi}_{\pm}^{sn}(\theta_2) \rangle = 0 = \langle \hat{\xi}_{\pm}^{ss}(\theta_1) \hat{\xi}_{\pm}^{sn}(\theta_2) \rangle .$$

If the intrinsic shape of any two galaxies is assumed to be uncorrelated, we can also conclude that

$$\langle \hat{\xi}_{\pm}^{nn} \rangle = 0 \text{ for } \theta > 0 \quad (13.9)$$

and hence

$$\langle \hat{\xi}_{\pm}^{nn}(\theta_1) \hat{\xi}_{\pm}^{ss}(\theta_2) \rangle = \langle \hat{\xi}_{\pm}^{nn}(\theta_1) \rangle \cdot \langle \hat{\xi}_{\pm}^{ss}(\theta_2) \rangle = 0 \text{ for } \theta_1, \theta_2 > 0 .$$

The covariance matrix can thus be split into three different contributions,

$$C_{\pm, \pm} = C_{\pm, \pm}^{nn} + C_{\pm, \pm}^{ss} + C_{\pm, \pm}^{sn} , \quad (13.10)$$

namely

$$\begin{aligned} C_{\pm, \pm}^{nn}(\theta_1, \theta_2) &= \langle \hat{\xi}_{\pm}^{nn}(\theta_1) \hat{\xi}_{\pm}^{nn}(\theta_2) \rangle , \\ C_{\pm, \pm}^{ss}(\theta_1, \theta_2) &= \langle \hat{\xi}_{\pm}^{ss}(\theta_1) \hat{\xi}_{\pm}^{ss}(\theta_2) \rangle - \xi_{\pm}(\theta_1) \xi_{\pm}(\theta_2) , \\ C_{\pm, \pm}^{sn}(\theta_1, \theta_2) &= 4 \cdot \langle \hat{\xi}_{\pm}^{sn}(\theta_1) \hat{\xi}_{\pm}^{sn}(\theta_2) \rangle . \end{aligned} \quad (13.11)$$

The $C_{\pm, \pm}^{ss}$ term depends on 4-point functions of the shear field and is called the *cosmic variance* term. In order to evaluate it, further assumptions on the probability distribution

function (PDF) of the shear or the convergence field are needed and we will discuss two possible models for the convergence PDF in sections 13.2.2 and 13.2.2 - the Gaussian and the log-normal model.

The contributions $C_{\pm\pm}^{nn}$ and $C_{\pm\pm}^{sn}$ can be computed without additional assumptions. In Joachimi et al. (2008) it is derived that they are given by²

$$\begin{aligned} C_{\pm\pm}^{sn}(\theta_1, \theta_2) &= \frac{\sigma_\epsilon^2}{\pi A \bar{n}} \int d\ell \ell J_{0/4}(\ell\theta_1) J_{0/4}(\ell\theta_2) P_\kappa(\ell) , \\ C_{++}^{nn}(\theta_1, \theta_2) &= C_{--}^{nn}(\theta_1, \theta_2) \\ &= \frac{\sigma_\epsilon^4}{N_p(\theta_1)} \cdot \delta_{\theta_1, \theta_2} , \\ C_{+-}^{nn}(\theta_1, \theta_2) &= 0 , \end{aligned} \quad (13.12)$$

where A is the survey area, \bar{n} is the number density of galaxies, σ_ϵ is the dispersion of the intrinsic ellipticity which is defined as

$$\sigma_\epsilon^2 := \langle \boldsymbol{\epsilon}^{\text{in}} \boldsymbol{\epsilon}^{\text{in}*} \rangle , \quad (13.13)$$

and P_κ is again the convergence power spectrum.

Gaussian Approximation

In the paper series by Schneider et al. (2002), Kilbinger & Schneider (2004) and Joachimi et al. (2008) the covariance matrix is studied in the Gaussian approximation, i.e. assuming that the convergence field has a Gaussian PDF such that its 4-point correlation functions can be expressed in terms of its 2-point correlation functions.

For the case where the survey geometry is much larger than the angular scales considered in the correlation functions, Joachimi et al. (2008) derive the following expressions for the cosmic variance term:

$$C_{\pm\pm}^{ss}(\theta_1, \theta_2) = \frac{1}{\pi A} \int d\ell \ell J_{0/4}(\ell\theta_1) J_{0/4}(\ell\theta_2) P_\kappa^2(\ell) . \quad (13.14)$$

However, due to the finite geometry of any given survey equation 13.14 generally overestimates the covariance of Gaussian field as was demonstrated in Sato et al. (2011). This *finite area effect* according to Sato et al. is not important for surveys larger than 1000 deg^2 . For smaller surveys a method developed in Kilbinger & Schneider (2004) which doesn't employ an ensemble average over galaxy positions should be used to evaluate the Gaussian covariance. This method was for example used in the analysis of CHFTLenS data in Kilbinger et al. (2013). The finite area effect is also important for internal covariance estimation and will be further discussed in section 13.4.2.

²as in Schneider et al. (2002) they employ an ensemble average over the galaxy positions to derive their expressions.

Shifted Log-Normal Approximation

As e.g. reported by Taruya et al. (2002), Vale & White (2003) or by Hilbert et al. (2011) the Gaussian model fails to describe the true PDF of the convergence and especially on small separations poorly represents the true covariance of the cosmic shear 2-point functions.

Hilbert et al. (2011) propose a different model for the convergence PDF, namely that of a *zero-mean shifted log-normal distribution*. In this approach the convergence at a given point on the sky is assumed to be of the form

$$\kappa(\boldsymbol{\theta}) = \exp[n(\boldsymbol{\theta})] - \kappa_0 \quad (13.15)$$

where $n(\boldsymbol{\theta})$ is a Gaussian random field (not necessarily with a vanishing mean) and the *minimal convergence parameter* κ_0 is chosen such that $\langle \kappa \rangle = 0$. Hilbert et al. (2011) show that from the corresponding PDF a model for the shear-shear contribution to the covariance matrix can be derived. Considering only the most important terms they also provide a simplified log-normal covariance, which reads

$$\begin{aligned} C_{\pm\pm}^{ss}(\theta_1, \theta_2) &= \frac{1}{\pi A} \int d\ell \ell J_{0/4}(\ell\theta_1) J_{0/4}(\ell\theta_2) P_\kappa^2(\ell) \\ &+ \frac{8\pi}{\kappa_0^2 A} \xi_\pm(\theta_1) \xi_\pm(\theta_2) \int_0^{\theta_A} d\theta \theta \xi_+(\theta) , \end{aligned} \quad (13.16)$$

where θ_A represents the 'radius' of the survey, given by

$$\theta_A = \sqrt{\frac{A}{\pi}} . \quad (13.17)$$

Comparing equation 13.16 to equation 13.14 one can see that the simplified log-normal approximation to $C_{\pm\pm}^{ss}$ consists of only one correction term to the Gaussian model. In our paper, we will simulate log-normally distributed convergence fields and use equation 13.16 to compute the cosmic variance part of our model covariance.

Finite bin width

The expressions presented above for the covariance of $\hat{\xi}_\pm$ are derived under the assumption of small angular bins (Schneider et al., 2002). However, in section 13.5.2 we need correct covariance expressions also for data vectors where the relative bin width is ~ 0.3 , i.e. where the assumption of small bins does not hold. This is in fact the more realistic case, since broad bins are commonly used to reduce the number of data points (see e.g. Kilbinger et al. (2013), Becker et al. (2016)).

Hence, in section 13.5.2 we proceed as follows: We first compute the log-normal model for the covariance, eqn. 13.16, for a set of very small angular bins $\tilde{\theta}_i, i = 1, \dots, \tilde{N}$. Then we apply a linear transformation that takes the large data vector of the small angular bins

to a smaller data vector by putting together p neighboring bins of the old data vector,

$$\begin{aligned}\theta_j &= \sum_{i=p \cdot (j-1)+1}^{p \cdot j} \tilde{\theta}_i N_p(\tilde{\theta}_i) / \sum_{i=p \cdot (j-1)+1}^{p \cdot j} N_p(\tilde{\theta}_i) \\ \hat{\xi}(\theta_j) &= \sum_{i=p \cdot (j-1)+1}^{p \cdot j} \hat{\xi}(\tilde{\theta}_i) N_p(\tilde{\theta}_i) / \sum_{i=p \cdot (j-1)+1}^{p \cdot j} N_p(\tilde{\theta}_i),\end{aligned}\tag{13.18}$$

where $N_p(\tilde{\theta}_i)$ is the number of pairs in the i th bin of the finer data vector.

The same linear transformation is then applied to the covariance matrix of the large data vector to get the covariance matrix of the compressed data vector. We find that for $\hat{\xi}_-$ this decreases the mixed- and cosmic variance part of the covariance by $\gtrsim 30\%$, while for $\hat{\xi}_+$ it makes almost no difference. The reason is that adjacent bins in ξ_+ are much more correlated than adjacent bins in ξ_- . Hence, if two bins in ξ_+ are joined, the variance of the joined bin is almost identical to that of the individual bins and eqn. 13.16 can still be applied³.

13.3 Log-normal Simulations

Simon et al. (2004) describe a quick method to simulate cosmic shear surveys based on a Gaussian convergence field for any given convergence-power-spectrum. On a quadratic grid in 2D-Fourier space they generate at each point ℓ of the grid a value of the convergence

$$\hat{\kappa}(\ell) = \kappa_1(\ell) + i\kappa_2(\ell)$$

where the components $\kappa_i(\ell)$ are drawn from a Gaussian distribution with zero mean and variance

$$\sigma_\ell^2 = \frac{1}{2V} P_\kappa(\ell).$$

Here P_κ is the desired convergence power-spectrum and V is the volume of the grid in angular space which is given in terms of the grid spacing $\Delta\ell$ as

$$V = \left(\frac{2\pi}{\Delta\ell} \right)^2.\tag{13.19}$$

In order to achieve a convergence field that is real valued in angular space one has to impose the condition

$$\hat{\kappa}(\ell) = \hat{\kappa}^*(-\ell)$$

and in Fourier space the shear field is related to the convergence field by the equation⁴

$$\hat{\gamma}(\ell) = \frac{\ell_1^2 - \ell_2^2 + 2i\ell_1\ell_2}{\ell^2} \hat{\kappa}(\ell).\tag{13.20}$$

³a similar reasoning can be applied for the off-diagonal terms of the covariance

⁴see eqn. 2.1.11 of Kaiser & Squires (1993) or eqn. 25 of Simon et al. (2004)

A Fourier transform then gives the shear field in angular space.

The main idea in generating a log-normal random field is to generate a Gaussian field $n(\boldsymbol{\theta})$ with the method of Simon et al. (2004) and transform it into $\kappa(\boldsymbol{\theta})$ via equation 13.15. According to Martin et al. (2012); Takahashi et al. (2014) the power spectrum of $n(\boldsymbol{\theta})$, P_n , can be computed from P_κ as follows:

First, the 2-pt. function of $\kappa(\boldsymbol{\theta})$ is given in terms of the power spectrum P_κ by

$$\xi_\kappa(\theta) = \frac{1}{2\pi} \int_0^\infty d\ell \ell P_\kappa(\ell) J_0(\ell\theta) .$$

Next, the 2-pt. function ξ_n is related to the 2-pt. function of $n(\boldsymbol{\theta})$ via (see e.g. equation B.8 of Hilbert et al. 2011)

$$\xi_n(\theta) = \ln \left(\xi_\kappa(\theta) / \kappa_0^2 + 1 \right) ,$$

where κ_0 is the minimal convergence parameter from eqn. 13.15. Finally, the power spectrum of the Gaussian field $n(\boldsymbol{\theta})$ by

$$P_n(\ell) = 2\pi \int_0^\infty d\theta \theta \xi_n(\theta) J_0(\ell\theta) . \quad (13.21)$$

The field $n(\boldsymbol{\theta})$ can now be generated as described by Simon et al. (2004). However, this way $n(\boldsymbol{\theta})$ will have a mean value of zero. In order to ensure that $\langle \kappa \rangle = 0$ the mean value

$$\mu = \kappa_0 - \frac{\sigma^2}{2} \quad (13.22)$$

has to be added, where σ^2 is the variance of the Gaussian field. The convergence field $\kappa(\boldsymbol{\theta})$ now has to be transformed into Fourier space. Using equation 13.20 one can then compute the Fourier modes of the shear field and another Fourier transform gives the desired shear field in angular space.

13.3.1 Setup and Validation of the Simulations

The harmonic space grid we are using has a total number of $(2^{16})^2$ grid points and a grid spacing of $\Delta\ell = 2$. Hence, in each axis it ranges from $-\ell_{\max} = -2^{16}$ to $\ell_{\max} = +2^{16}$. All modes $\gamma(\ell)$ with $|\ell| > \ell_{\max}$ (i.e. the corners of the grid) are set to zero. The mode γ_0 is also set to 0 and all other modes are generated as explained above. Especially, we have to fix a cosmology and assume a certain redshift distribution of sources, $p(z)$, to compute the convergence power spectrum P_κ .

Following eqn. 13.19 the grid in angular space has a volume of $V = 2\pi/2 \approx 10^4 \text{ deg}^2$. Out of the center of that volume we will cut out a sub-grid of size A . Onto that sub-grid we are uniformly placing galaxies with a certain number density n_{gal} . The shear of each individual galaxy is then determined by quadratic interpolation of the grid onto the galaxy position. Finally, a Gaussian intrinsic shape noise with an ellipticity dispersion σ_ϵ is added to get the total shape of the galaxy. Note that we simply added the shear signal and intrinsic ellipticity, hereby ignoring the effects of *reduced shear*.

| setup | A [deg] | n_{gal} | z_{median} | κ_0 |
|-------|------------|------------------|---------------------|------------|
| I | 4900 | 20 | 1.0 | 0.032 |
| IIa | ~ 150 | 6 | 0.7 | 0.019 |
| IIb | 5000 | 6 | 0.7 | 0.019 |
| IIc | 5000 | 10 | 0.7 | 0.019 |

Table 13.1: The different configurations of mock catalogs used in this paper.

In this work we always keep the cosmology fixed to that of Hilbert et al. (2011), i.e. a flat Λ CDM universe with $(\Omega_m, \Omega_b, \sigma_8, h_{100}, n_s) = (0.25, 0.045, 0.9, 0.73, 1.0)$. To compute the convergence power spectrum we employ *halofit* (Smith et al., 2003) using the open source code NICAEA⁵. The source distribution is taken to have the form

$$p(z) = \frac{3z^2}{2z_0^3} e^{-\left(\frac{z}{z_0}\right)^{3/2}}, \text{ where } z_0 = \frac{z_{\text{median}}}{1.412}. \quad (13.23)$$

This is the same form that was also assumed by Hilbert et al. (2011). The ellipticity dispersion is always set to 0.3 per component, i.e. $\sigma_\epsilon = \sqrt{2} \cdot 0.3$. All other quantities, i.e. area A , source density n_{gal} and median redshift z_{median} , will be varied throughout section 13.5. The different setups are summarized in table 13.1.

The redshift distribution of setup I is exactly that of Hilbert et al. (2011) and imitates a rather deep survey comparable, e.g. , to *euclid*. In this setup, we measure the 2-pt. correlation functions in 35 logarithmic bins from $\theta_{\text{min}} = 1'$ to $\theta_{\text{max}} = 150'$. The area A was taken to be a square of $70 \text{ deg} \times 70 \text{ deg}$. The minimal convergence parameter κ_0 was chosen to be 0.032 as suggested by Hilbert et al. for this redshift distribution.

The area, galaxy density and redshift distribution of setup IIa are chosen to be similar to that of DES science verification data (DES-SV) which was used in Becker et al. (2016). In this setup, we measure the 2-pt. correlation functions in 15 logarithmic bins from $\theta_{\text{min}} = 2'$ to $\theta_{\text{max}} = 300'$, which is also exactly the data vector used by Becker et al. (2016). In this setup we also reproduce the irregular shape of DES-SV, i.e. we use an SV-shaped healpix mask to cut out the sub-volume A .

The setups IIb and IIc are aimed at a forecast for the final 5-year DES data. In IIb we are assuming the same source density as in DES-SV and in IIc a slightly higher one. Note that in principle, when adjusting the source density, we should also adjust the source median redshift of the sources. But we will ignore this point, since our redshift distribution is anyway only a rough match to that of DES. Thus, for all setups IIa, IIb and IIc we take a median redshift of 0.7. Furthermore, for all these setups we use an empirical formula $\kappa_0(z)$ found by Hilbert et al. (2011) to fix the minimal convergence parameter. Inserting the mean redshift of $z_{\text{mean}} \approx 0.745$ gives a value of $\kappa_0 = 0.019$. The area in setups IIb and IIc are simply taken to be square shaped.

To validate our simulations, we generate 1000 independent realizations of setup I. In order to speed up the computations we decrease the number of galaxies with respect to

⁵by Kilbinger et al., www.cosmostat.org/software/nicaea/

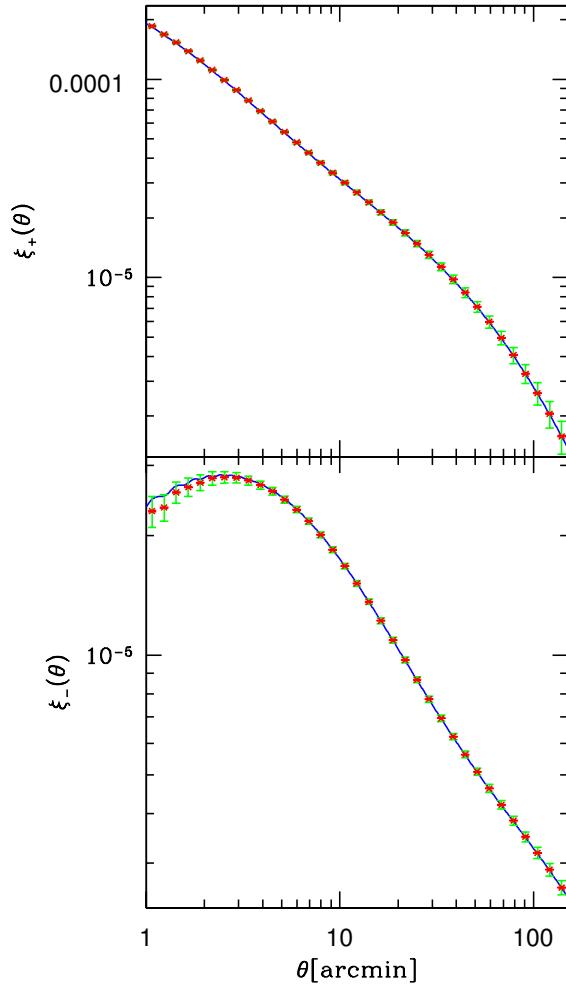


Figure 13.1: Comparison of the mean correlation functions from 1000 simulations (red dots) and the input model (blue line). The the red error bars show the standard deviation of the mean and the green errorbars show the standard deviation of the single correlation function measurements. We used the redshift distribution of Hilbert et al. (2011) to compute the input power spectrum and we also used their value of κ_0 to generate the log-normal convergence. Note that in section 13.5.2 we will use a different configuration.

our jackknife analysis by a factor of 5, i.e. to $n_{\text{gal}} = 4/\text{arcmin}^2$, while at the same time decreasing the ellipticity dispersion by a factor of $\sqrt{5}$. This way the covariance expressions in equation 13.16 stay unaffected.

In figure 13.1 we show the mean measured correlation functions in the mock surveys. The correlation function measurement was carried out using the `TREECORR` tree code⁶. The measured correlation functions and those derived from the input model agree well on most scales. Only at small angular scales the measured value of ξ_- differs significantly from the input model. The reason is the artificial cut-off at high ℓ -values in our Fourier grid which both in the model and the simulation introduces artifacts - as can be seen from the oscillatory behavior of ξ_- . To keep our analyses in section 13.5 free from these artifacts we will only consider those bins in ξ_- that have $\theta \gtrsim 4.5'$. For ξ_+ we continue to use a range of $1' < \theta < 150'$. Also, for the setups *IIa* to *IIc* (not shown here) the discrepancy in ξ_- turns out to be less significant. Hence for these setups we stay with $\theta_{\min} = 2'$.

Figure 13.2 compares the sample covariance of the 1000 simulations to the predictions from equation 13.16. The relative deviation between measured variance and the log-normal model is $\leq 20\%$ for ξ_+ and $\leq 15\%$ for ξ_- . For both correlation functions these deviations seem to be significant given the uncertainties of the sample covariance estimate. However, the sample variance values at different angular scales are highly correlated, which makes a 'χ-by-eye' judgment of the fit impossible. When transforming the covariance matrices into the eigenbasis of the model covariance (right-hand panel of figure 13.2), the variance values become uncorrelated and the agreement of the covariance matrices becomes more evident. The eigenvalues at which the log-normal covariance significantly differs from the sample covariance of our simulations are 3 orders of magnitude smaller than the biggest eigenvalues for ξ_+ and more than 2 orders of magnitude smaller than the biggest eigenvalues for ξ_- (c.f. right-hand panel of figure 13.2). Finally, our analyses in section 13.5 remain unchanged when the log-normal covariance is exchanged by the sample covariance of the 1000 independent realizations, which validates the simulations for our purposes (cf. appendix 13.A, figure 13.14).

13.4 Internal Covariance Estimation for two-point correlation functions

Suppose the correlation functions ξ_{\pm} have been measured in finite bins around a set of angular distances θ_i , $i = 1, \dots, d$. Let $\hat{\xi}$ be either one of the data vectors $[\xi_{\pm}(\theta_1), \dots, \xi_{\pm}(\theta_d)]$ or the joint data vector of both correlation functions.

If $\xi[\pi]$ is a model for the measurement $\hat{\xi}$ which depends on a set of parameters π , then a common statistic for constraining the possible values of π is the χ^2 statistic (Kilbinger & Schneider, 2004), i.e.

$$\chi^2[\pi] = (\hat{\xi} - \xi[\pi])^T C^{-1} (\hat{\xi} - \xi[\pi]) , \quad (13.24)$$

⁶by Jarvis et al., github.com/rmjarvis/TreeCorr

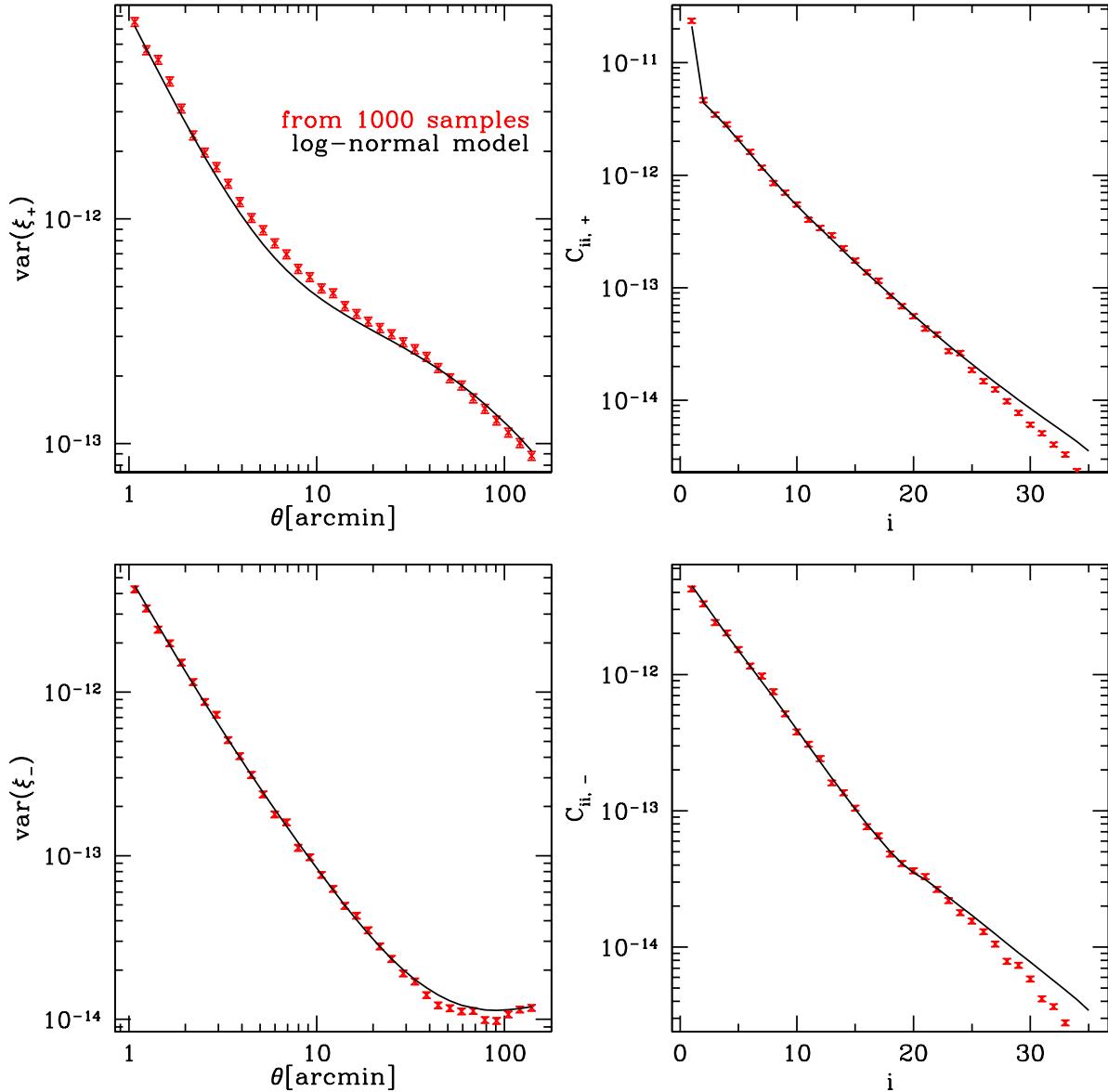


Figure 13.2: Left: sample variance from 1000 independent simulations compared to the log-normal input model. The errorbars are assuming a Wishart distribution, note however that the different sample variance values are correlated. Right: in the diagonal basis of the model covariance matrix the sample variance values should independently follow a χ^2 -distribution. The model and the simulations are consistent for the ≈ 20 largest eigenvalues of the model covariance matrix.

where C is the covariance matrix of $\hat{\xi}$. One way to get the covariance matrix is to model it theoretically. As we have seen in section 13.2.2 the modeling of the covariance depends crucially on the PDF of the convergence field (Schneider et al., 2002; Hilbert et al., 2011; Sato et al., 2011) and neither the Gaussian nor the log-normal approximation match a realistic convergence PDF. Also, the model covariance matrix will depend on cosmological parameters itself which, at least for small surveys, has to be taken into account when deriving parameter constraints (Eifler et al., 2009).

A way to get around modeling the covariance matrix directly is to use the sample covariance of measurements of the correlation functions in a set of independent N-body simulations (cf. Takahashi et al. 2009; Sato et al. 2009; Harnois-Déraps & van Waerbeke 2015 or for an application to data Kilbinger et al. 2013) which however still depends on the model parameters, i.e. on the assumption of a particular cosmological model. Another alternative to modeling the covariance matrix is to estimate it from the data itself. In the following we will introduce three different internal covariance estimation methods - the *sub-sample covariance*, the *delete-one-jackknife* and the *bootstrap* (cf. Norberg et al. 2009; Loh 2008).

13.4.1 Subsample Covariance

Let us split the area A of our cosmic shear survey into N equally shaped and sized subregions of the area $A_S = A/N$. In each subregion $\alpha = 1, \dots, N$, a measurement of the data vector $\hat{\xi}^\alpha$ can be carried out. Assuming that each sub-region has approximately the same number of galaxies and that the correlation functions are measured on scales much smaller than $\sqrt{A_S}$ the measurement of $\hat{\xi}$ in the whole survey is given by

$$\hat{\xi} \approx \bar{\xi} := \frac{1}{N} \sum_{\alpha=1}^N \hat{\xi}^\alpha , \quad (13.25)$$

i.e. it is the mean values of the measurements in the sub-regions. If the measurements $\hat{\xi}^\alpha$ are independent, then the ij -th element of their covariance matrix can be estimated by

$$\langle \Delta \hat{\xi}_i^\alpha \Delta \hat{\xi}_j^\alpha \rangle \approx \frac{1}{N-1} \sum_{\beta=1}^N (\hat{\xi}^\beta - \bar{\xi})_i (\hat{\xi}^\beta - \bar{\xi})_j , \quad (13.26)$$

where $\Delta \hat{\xi}^\alpha$ is the difference between $\hat{\xi}^\alpha$ and its expectation value

$$\xi = \langle \hat{\xi}^\alpha \rangle = \langle \hat{\xi} \rangle . \quad (13.27)$$

Accordingly, if the assumption of independent sub-regions were true, the covariance of the total measurement $\hat{\xi}$ could be estimated by

$$\hat{C}_{\text{SC}} = \frac{1}{N(N-1)} \sum_{\alpha=1}^N (\xi^\alpha - \bar{\xi})^T (\xi^\alpha - \bar{\xi}) . \quad (13.28)$$

We will call the estimator in equation 13.28 the *sub-sample covariance* (Norberg et al., 2009). The main systematic effects of internal covariance estimation can be most easily understood in terms of this estimator. Hence, before introducing the jackknife and bootstrap estimator, we will explain these systematics in the following two sections.

13.4.2 Correlation of sub-samples

The sub-sample covariance estimator relies on the assumption that the data is split into independent sub-samples, i.e. that there is no correlation of the measurements of the correlation functions in the different sub-regions,

$$\langle \Delta \hat{\xi}_i^\alpha \Delta \hat{\xi}_j^\beta \rangle \stackrel{!}{=} 0, \text{ for } \alpha \neq \beta. \quad (13.29)$$

This can be seen from the fact that eqn. 13.28 simply rescales the sub-field-to-sub-field covariance by a factor of $1/N$ to get the covariance of the whole survey. If the sub-samples are correlated, this will underestimate the true covariance matrix (Nordman & Lahiri, 2007).

Another way to think about this is as follows: the sub-sample covariance estimator assumes that the covariance matrix of $\hat{\xi}$ is inversely proportional to the survey area A . Hence it estimates the covariance of sub-regions of the size A_S within the data and then rescales it to the total area,

$$C = \frac{A_S}{A} \cdot C_S = \frac{1}{N} \cdot C_S, \quad (13.30)$$

where N is again the number of sub-regions. But already from the log-normal model for the covariance it can be seen, that this rescaling is not valid. The log-normal correction term to the Gaussian covariance matrix is given by

$$C_{\pm\pm}^{ss,\log}(\theta_1, \theta_2) = \frac{8\pi}{\kappa_0^2 A} \xi_\pm(\theta_1) \xi_\pm(\theta_2) \int_0^{\theta_A} d\theta \theta \xi_+(\theta).$$

This term may be proportional to $1/A$, but the upper integral boundary also depends on the survey diameter θ_A . As A increases, the covariance therefore decreases slower than $1/A$. Hence, assuming $1/A$ scaling when extrapolating from the covariance of the smaller sub-fields to the covariance of the full area underestimates the full covariance. Also, note that even the Gaussian covariance term in eqn. 13.16 is only an approximation for large survey sizes A . It also suffers from a finite area effect as can be seen from its derivation in Schneider et al. (2002) or its form given in Hilbert et al. (2011).

The fact that sub-samples should be as uncorrelated as possible is also the reason why the re-sampling of the data should be done into spatially connected patches. If instead the data would be randomly divided into sub-samples then the shear correlations in the sub-samples would be almost identical. Hence, only the shape-noise contributions to the covariance would be measured by such an estimator.

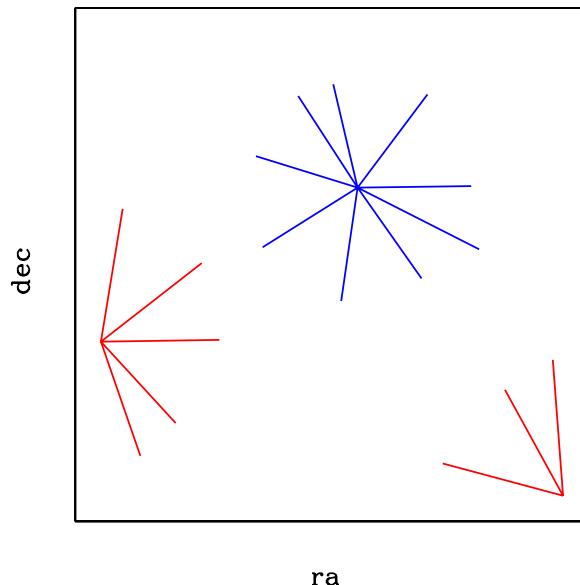


Figure 13.3: Galaxies at the edge of a sub-region (in red) contribute less pairs to the measurement of the correlation functions (i.e. to equation 13.5 applied to the sub-sample) than galaxies in the center of the sub-region (in blue). Consequently, the area of the sub-patch is not uniformly probed by the galaxy pairs. This increases the cosmic variance between sub-regions and biases the covariance estimates high. Hence, it has an opposite effect to the correlation of sub-samples, which biases the covariance estimates low. As seen from the left-hand panel of figure 13.5, at large angular scales this can even lead to an overestimation of the cosmic variance of $\hat{\xi}_-$ (in the galaxy-scheme, c.f. also figure 13.4).

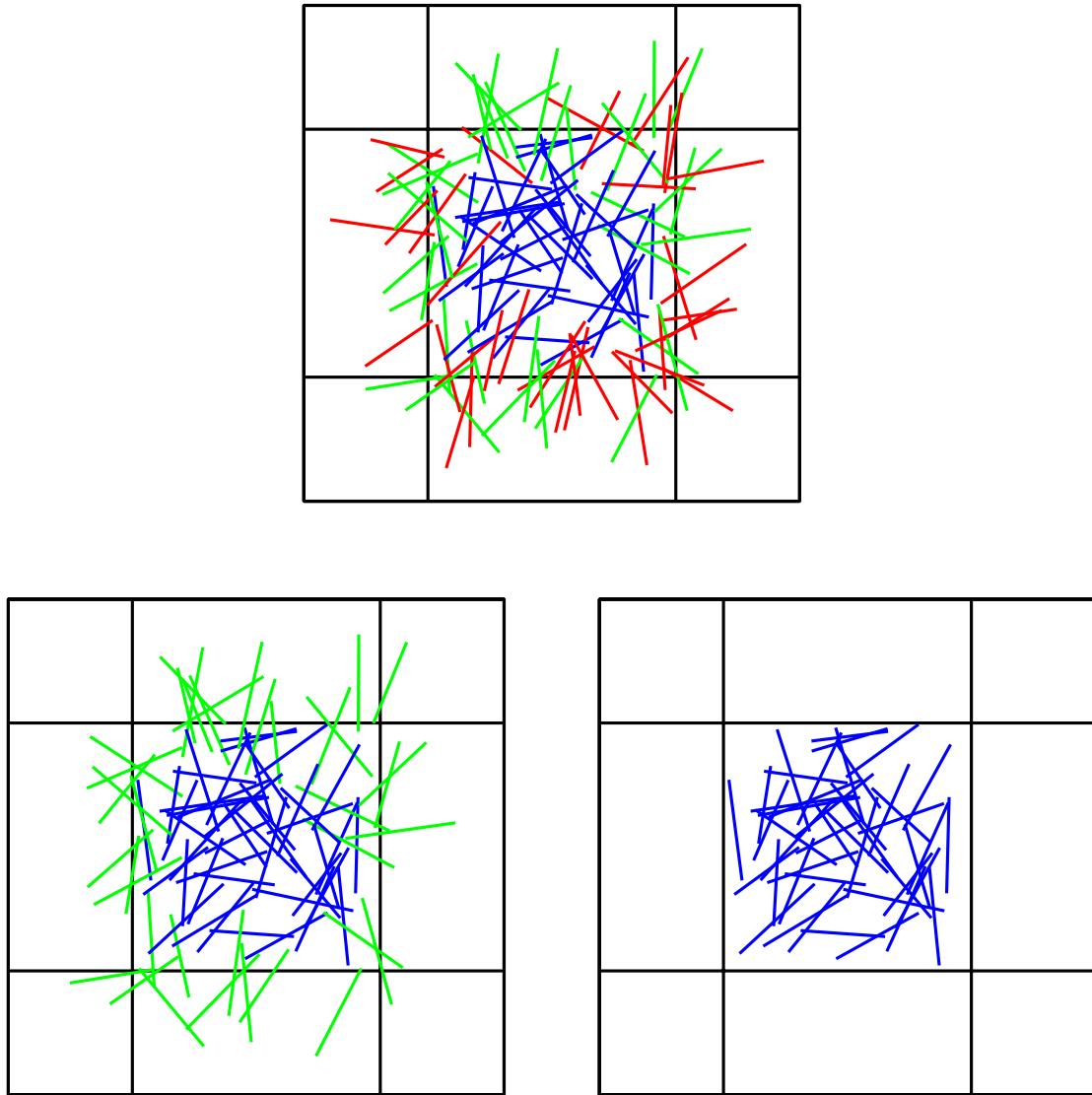


Figure 13.4: Two basic schemes for dividing a set of galaxy pairs into sub-samples: For each sub-region of the survey, there will be galaxy pairs crossing from that region into another (upper panel, green and red). In the galaxy-scheme ξ^α is computed by considering only pairs that completely lay within the sub-region α (lower right panel). In the pair-jackknife scheme (lower left panel) half of the pairs that cross from α to another region (drawn in green) are taken into account for computing ξ^α while only the other half (red) is discarded.

13.4.3 Galaxy pairs crossing between sub-samples

A problem specific to the internal covariance estimation for two-point correlation functions is the question of what to do with pairs of galaxies where each galaxy lies in a different sub-region of the survey.

In fact, the pieces of information in a cosmic shear survey are not the individual galaxy shapes but the pairs of galaxy shapes. If the pairs crossing between sub-regions are completely ignored when computing the sub-measurements $\hat{\xi}^\alpha$, then one is re-sampling a data set that has *less* information than the total measurement of ξ_\pm and hence a larger variance. Note, that this does not only influence the shape-noise part of the covariance but also the cosmic variance part. The reason is that galaxies at the edge of a sub-region contribute less terms to the correlation function measurement than galaxies in the center of the sub-region (c.f. figure 13.3), i.e. the area of the sub-patch is not uniformly probed by the galaxy pairs and the measured shear correlations are dominated by the inner part of the patch. In contrast to the correlation of sub-samples discussed before, this increases the cosmic variance between the sub-samples and can bias the covariance estimate high - especially on large angular scales.

This effect can in principle be resolved by re-sampling the set of pairs (instead of the set of galaxy shapes), i.e. by defining the sub-measurement $\hat{\xi}^\alpha$ as

$$\hat{\xi}_\pm^\alpha(\theta) = \frac{\sum_{\text{pairs in } \alpha} (\epsilon_t^i \epsilon_t^j \pm \epsilon_x^i \epsilon_x^j) + \sum_{\text{half of cross pairs}} (\epsilon_t^i \epsilon_t^j \pm \epsilon_x^i \epsilon_x^j)}{N_{\text{pairs}}}.$$
(13.31)

How this re-sampling of galaxy pairs can be done is illustrated in figure 13.4. Especially one has to make sure that each galaxy pair enters exactly one of the $\hat{\xi}^\alpha$. We call this procedure the *pair-scheme* while we will call the standard procedure of considering only the individual galaxies in sub-region α when computing $\hat{\xi}^\alpha$ as *galaxy-scheme*.

In figure 13.5 we demonstrate this effect along with the effect of correlated sub-samples that was discussed before. The left-hand panel shows sub-sample estimates of the variance of $\hat{\xi}_\pm$ in a simulated survey (corresponding to setup I in table 13.1) where the shape noise was put to zero and where 400 sub-samples were used. Both the variance of $\hat{\xi}_+$ and $\hat{\xi}_-$ are severely underestimated on small scales, which is due to the correlation of sub-samples. At large angular scales, the galaxy-scheme yields systematically higher value for the variance than the pair-scheme and at least for $\hat{\xi}_-$ it can even overestimate the variance. This is due to the missing cross-pairs in the re-sampling.

For the right-hand panel of figure 13.5 we have generated a catalog of pure shape noise ($\sigma_\epsilon, A, n_{\text{gal}}$ as in setup I). This is the only situation where the assumption of uncorrelated sub-samples is valid. You can see that in this case the pair-scheme is able to estimate the variance without bias. The galaxy-scheme overestimates the variance for the reasons

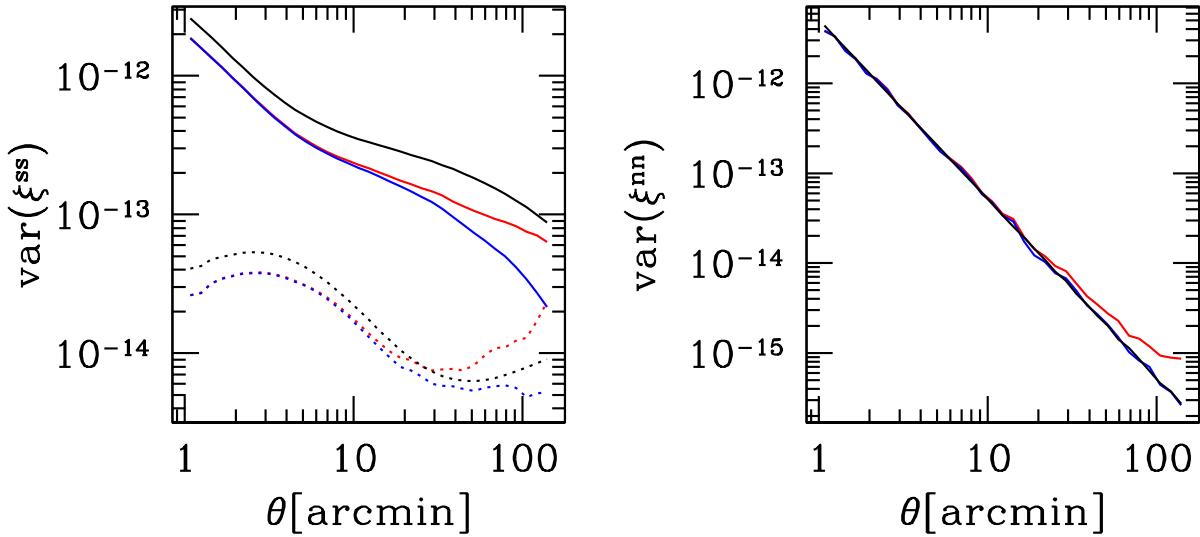


Figure 13.5: Different variance estimates using the sub-sample covariance estimator and 400 sub-samples. Left: Variance estimates for $\hat{\xi}_+$ (solid lines) and $\hat{\xi}_-$ (dotted lines) in a mock catalog without shape noise that is otherwise following setup I. The red lines show the galaxy-scheme estimate (c.f. section 13.4.3), the blue lines show the pair-scheme estimate and the black lines show the log-normal input model. Right: Sub-sample estimates of the variance of $\hat{\xi}_+$ in a mock catalog that only consists of shape noise and has the same area and density as in setup I. It is only in this situation (and in the pair-scheme) that internal estimation of the covariance yields unbiased results.

explained before. A downside of the pair-scheme is that the shear signals in the sub-measurements $\hat{\xi}^\alpha$ become even more correlated, as can also be seen from the left-hand panel of figure 13.5.

13.4.4 Jackknife

Another method of covariance estimation that Norberg et al. (2009) investigate is the *delete-one-jackknife*. Instead of estimating the covariance of the measurements ξ^α and rescaling it to the size of the whole survey the jackknife is considering the measurements

$$\hat{\xi}_\pm^{*\alpha}(\theta) = \frac{\sum_{\{i,j \text{ not in } \alpha\}} (\epsilon_i^i \epsilon_t^j \pm \epsilon_x^i \epsilon_x^j) \cdot \Delta_\theta(|\theta_i - \theta_j|)}{\sum_{\{i,j \text{ not in } \alpha\}} \Delta_\theta(|\theta_i - \theta_j|)} , \quad (13.32)$$

i.e. the *jackknife-sample* α is generated by cutting out the subregion α and measuring the correlation functions in the rest of the survey. The jackknife estimate for the covariance

matrix is then given by (Efron, 1982; Norberg et al., 2009)

$$\hat{C}_{\text{jack}} = \frac{N-1}{N} \sum_{\alpha=1}^N (\boldsymbol{\xi}^{*\alpha} - \bar{\boldsymbol{\xi}}^*)^T (\boldsymbol{\xi}^{*\alpha} - \bar{\boldsymbol{\xi}}^*) , \quad (13.33)$$

where $\bar{\boldsymbol{\xi}}^*$ is the mean of all jackknife measurements.

If we again assume that all subregions have the same galaxy density and that the correlation functions are measured on scales much smaller than the sub-region size then $\boldsymbol{\xi}^{*\alpha}$ is approximately given by

$$\boldsymbol{\xi}^{*\alpha} \approx \frac{1}{N-1} \sum_{\beta \neq \alpha} \hat{\boldsymbol{\xi}}^\beta . \quad (13.34)$$

From this it also follows that

$$\begin{aligned} \boldsymbol{\xi}^{*\alpha} - \bar{\boldsymbol{\xi}}^* &\approx \frac{1}{N-1} \sum_{\beta \neq \alpha} \hat{\boldsymbol{\xi}}^\beta - \frac{1}{N} \sum_{\beta} \boldsymbol{\xi}^{*\beta} \\ &= \frac{N \cdot \bar{\boldsymbol{\xi}} - \hat{\boldsymbol{\xi}}^\alpha}{N-1} - \frac{1}{(N-1) \cdot N} \sum_{\beta} \sum_{\gamma \neq \beta} \hat{\boldsymbol{\xi}}^\gamma \\ &= \frac{N \cdot \bar{\boldsymbol{\xi}} - \hat{\boldsymbol{\xi}}^\alpha}{N-1} - \frac{N-1}{(N-1) \cdot N} \sum_{\gamma} \hat{\boldsymbol{\xi}}^\gamma \\ &= \frac{N \cdot \bar{\boldsymbol{\xi}} - \hat{\boldsymbol{\xi}}^\alpha}{N-1} - \bar{\boldsymbol{\xi}} \\ &= \frac{\bar{\boldsymbol{\xi}} - \hat{\boldsymbol{\xi}}^\alpha}{N-1} . \end{aligned} \quad (13.35)$$

Inserting this into the definition of \hat{C}_{jack} gives exactly the subsample covariance \hat{C}_{SC} , i.e. on small angular scales the two methods are approximately equivalent⁷.

In jackknife estimation one can in principle also differentiate between a pair scheme and a galaxy scheme. Using eq. 13.32 for $\boldsymbol{\xi}^{*\alpha}$ corresponds to the galaxy scheme. This is equivalent to disregarding all pairs in the top panel of figure 13.4 when computing $\boldsymbol{\xi}^{*\alpha}$. The pair-scheme is given by disregarding all pairs in the lower left panel of figure 13.4 when computing $\boldsymbol{\xi}^{*\alpha}$. In the pair scheme jackknife and sub-sample covariance become exactly equivalent when (assuming that each sub-patch has the same number of galaxies).

13.4.5 Bootstrap Covariance

The so called *block bootstrap* estimator of the covariance also divides the data into sub-samples. If the data is split into N sub-regions, then a number of N_{boot} bootstrap resamplings of the data are generated by randomly drawing with replacement N of the sub-samples and combining then into one re-sampled data set (Norberg et al., 2009; Nordman

⁷This is no general statement on the jackknife method. It holds only in our particular situation.

& Lahiri, 2007; Loh, 2008; Efron, 1982). If the correlation function measured in the re-sampled data i ($i = 1, \dots, N_{\text{boot}}$) is called $\xi^{\text{boot},i}$, then the bootstrap estimate of the covariance is given by

$$\hat{C}_{\text{boot}} = \frac{1}{N_{\text{boot}} - 1} \sum_{i=1}^{N_{\text{boot}}} (\xi^{\text{boot},i} - \bar{\xi}^{\text{boot}})^T (\xi^{\text{boot},i} - \bar{\xi}^{\text{boot}}), \quad (13.36)$$

where $\bar{\xi}^{\text{boot}}$ is now the mean of all $\xi^{\text{boot},i}$.

Again, the question arises of whether one should consider the single galaxies or the galaxy pairs as the actual data (cf. section 13.4.3). In what we will call *galaxy-bootstrap* one simply adds a copy of all galaxies in a sub-region α to the re-sampled data set i each time the sub-region α gets drawn.

In the *pair-bootstrap* one adds all pairs associated to sub-region α to the list of pairs that is used to compute $\xi^{\text{boot},i}$. The difference between the two bootstrap schemes is mainly the following: if the sub-region α gets drawn n times, then each pair in α gets a weight of n in the pair-scheme and a weight of n^2 in the galaxy-scheme.

Note that the pair-bootstrap is very similar to what Loh (2008) describes as *marked point bootstrap*, the only difference being, that we chose to split pairs between sub-regions evenly among these regions.

We will see in section 13.5 that the galaxy-bootstrap severely overestimates the covariance. The other covariance estimators perform very similar to each other and suffer in similar ways from the systematics explained in subsection 13.4.2 and 13.4.3.

13.4.6 Stability and Inversion of the Covariance Estimate

All effects that bias the internal covariance estimate can in principle be minimized by dividing the data into very large sub-regions. This decreases both the correlation of the different sub-regions and the influence of pairs crossing between sub-regions. However, this also decreases the possible number of re-samplings and hence increases the variance of the covariance estimator itself.

In order to derive constraints on the number of re-samplings let us assume that we are in the limit where the correlations between sub-regions are small. Small here means that

$$\langle \Delta \hat{\xi}_i^\alpha \Delta \hat{\xi}_j^\beta \rangle \ll \langle \Delta \hat{\xi}_i^\alpha \Delta \hat{\xi}_i^\alpha \rangle, \text{ for } \alpha \neq \beta. \quad (13.37)$$

As explained before, this is the only limit in which internal covariance estimation is valid. In this limit the sub-sample covariance is just a rescaling of the sample covariance of independent realizations of the sub-regions. Hence - in the limit considered here and under the assumption that the data vector behaves Gaussian - the sub-sample covariance estimates are distributed according to a Wishart distribution (cf. Taylor et al. 2013). Also, the pair-jackknife is almost equivalent to the pair-version of the sub-sample covariance, i.e. to equation 13.28 when $\hat{\xi}^\alpha$ is computed with equation 13.31. Hence, also the pair-jackknife estimates should approximately follow a Wishart distribution.

The most important consequence of this is that the inverse of the covariance matrix estimate will be a biased estimate of the true inverse covariance matrix, and the bias is approximately given by (Hartlap et al., 2007; Taylor et al., 2013):

$$\langle \hat{C}_{\text{SC}}^{-1} \rangle \approx \frac{N-1}{N-d-2} C_{\text{true}}^{-1}, \quad (13.38)$$

where N is the number of sub-regions and d is the number of data points in $\hat{\xi}$. Especially, this factor has to be accounted for when computing the χ^2 statistic, eq. 13.24, i.e. it has an influence on the constraints derived on cosmological parameters when using internal covariance estimation.

Taylor et al. (2013) also give constraints on N with respect to d when a certain accuracy in the final parameter constraints is demanded.⁸ We take their criterion,

$$N > \frac{2}{\epsilon^2} + (d + 4), \quad (13.39)$$

where ϵ is the required fractional accuracy on parameter constraints, as a guideline also for internal covariance estimation. This is however under the assumption of an exact Wishart distribution, i.e. that the data vector is Gaussian and that the sub-regions are large enough to not cause systematic biases in the covariance estimate. Demanding a fractional accuracy of $\epsilon = 0.2$ for the parameter constraints, this yields a necessary number of $N > 54 + d$ re-sampling. Below this number there is no chance for internal covariance estimation to yield parameter constraints that are accurate to more than 20%.

13.5 Testing internal Covariance Estimators on simulated cosmic Shear Surveys

We will now use the simulations described in section 13.3 to test the performance of internal covariance estimators. First, we will use setup I (cf. 13.1) corresponding to a rather deep survey. We carry out 50 independent realizations of this survey. In each survey we measure the correlation functions in the range and binning that was explained in section 13.3. We then estimate the covariance of the measured correlation functions using the different internal estimation schemes that were introduced in section 13.4. Throughout this section - except for subsection 13.5.2 - we consider the log-normal model that was explained in section 13.2.2 as the *true* covariance of the simulated surveys. This is justified by the fact that our results don't change if we instead use the sample covariance of 1000 independent realizations that were presented in section 13.3 (cf. appendix 13.A, figure 13.14).

In figure 13.6 we compare the sub-sample, jackknife and bootstrap estimates of the diagonal elements of the covariance matrix (both in the galaxy- and pair-scheme) when splitting the survey into $N = 225$ sub-regions. The most impressive finding is, that in

⁸However, they are ignoring the impact that the variance in the inverted covariance estimate has on parameter constraints, which is investigated by Taylor & Joachimi (2014).

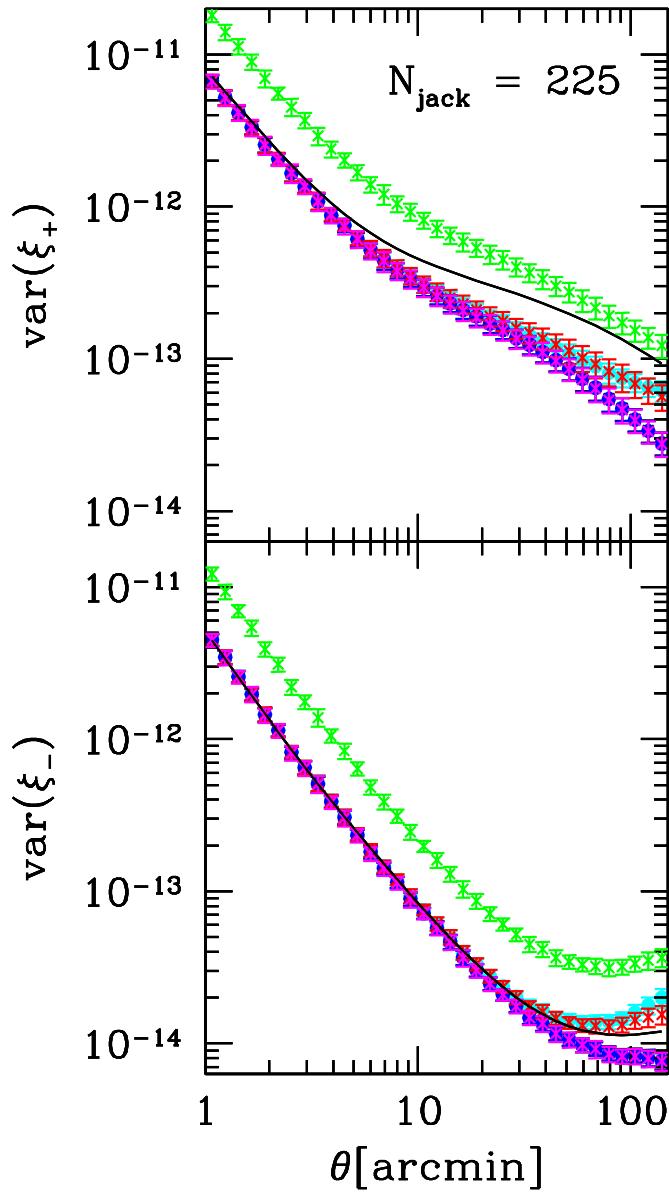


Figure 13.6: A comparison of the different internal estimation schemes when splitting the survey into $N = 225$ sub-regions. Green: galaxy-bootstrap, purple: pair-bootstrap, red: galaxy-jackknife, blue: pair-jackknife and cyan: sub-sample covariance compared to the analytical covariance (black line). We show the sub-sample covariance only in the galaxy-scheme because in the pair-scheme it is almost identical to jackknife and bootstrap. As explained in section 13.4, at large angular scales the different treatment of galaxy pairs crossing between sub-region leads to an overestimation of the variance by the galaxy-scheme and an underestimation of the variance by the pair-scheme.

the galaxy-scheme the bootstrap severely overestimates the variance. This is in agreement with the findings of Norberg et al. (2009) for galaxy clustering correlation functions. The duplication of whole sub-volumes of galaxies creates bootstrap samples that are in fact unrealistic, i.e. these bootstrap samples contain regions with no sources at all and on the other hand regions with a very high source density. Each original galaxy pair gets weighted by a factor of n^2 when the corresponding region is drawn n times. This puts a very high weight on small sub-areas of the bootstrap sample and creates an unphysically high variance between the bootstrap samples.

In the pair-scheme however, all three internal estimators perform almost identical. This is not surprising, because in that scheme the bootstrap is just an approximation to the sub-sample covariance and sub-sample and jackknife covariance are almost identical in the pair-scheme. As explained in section 13.4.3, in the galaxy-jackknife scheme the two effects of correlated sub-regions and false re-sampling of pairs partly cancel each other. Hence the galaxy-jackknife comes closest to the true variance at large scales. The performance of the sub-sample covariance (in the galaxy-scheme) only slightly differs from that.

Because of the strong similarity between the different estimator we will restrict the following analyses to the pair-jackknife and the galaxy-jackknife. We now investigate the influence of sub-region size on internal covariance estimation. Hence we split the surveys into 3 different numbers of sub-regions: 10^2 , 15^2 and 20^2 corresponding to sub-region areas of approximately $7.0 \times 7.0 \text{ deg}^2$, $4.67 \times 4.67 \text{ deg}^2$ and $3.5 \times 3.5 \text{ deg}^2$. In figure 13.7 we compare the mean value of the 50 jackknife estimates of the variance of $\hat{\xi}_\pm$ (the diagonal elements of the covariance matrix) to the true underlying log-normal model. A comparison of the off-diagonal behavior of the jackknife estimates to that of the input-covariance can be found in appendix 13.A. The errorbars in figure 13.7 represent the standard deviation of the 50 jackknife estimates, i.e. they illustrate the noise of the internal estimators. You can see in this figure the biases in the jackknife estimates that we explained in the previous section. For ξ_+ , both jackknife schemes underestimate the variance. At large scales, this is in the galaxy-jackknife scheme partly compensated by the false re-sampling of galaxy pairs. For ξ_- , the pair-jackknife underestimates the variance while the galaxy-jackknife overestimates it. ξ_- is a much more local measure in the sense that the different sub regions are less correlated in $\hat{\xi}_-$ and that the covariance matrix is much more dominated by the shape noise contributions. Hence, the severe systematic underestimation of the variance that can be seen for ξ_+ does not appear as strongly for ξ_- .

When increasing the number of sub-regions for the jackknife estimators, the noise in the variance estimates becomes smaller but the deviations from the true variance also become stronger. This is because for smaller sub-regions the estimated $\hat{\xi}^\alpha$ become more correlated and because there will be more galaxy pairs crossing from one sub-region to another.

13.5.1 Constraints on cosmological Parameters

We will now take the 50 simulations as mock observations and try to constrain the dark matter density parameter Ω_m and the power spectrum normalization σ_8 . To do so we sample the Ω_m - σ_8 plane on a fine grid while keeping the other cosmological parameters

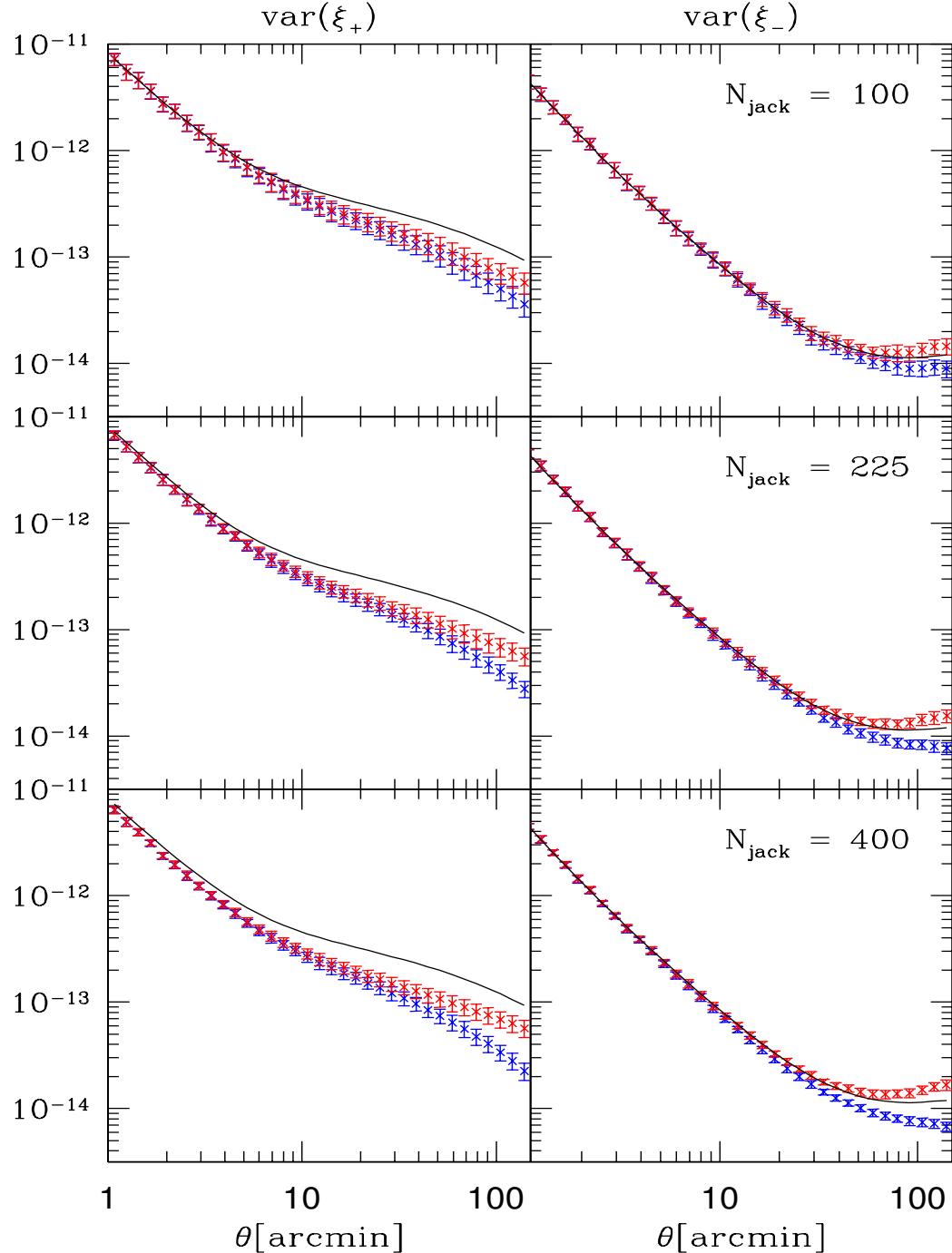


Figure 13.7: Mean values of 50 jackknife estimates of the variance of ξ_+ (left) and ξ_- (right). Galaxy-jackknife was used for the red points while pair-jackknife was used for the blue points and the errorbars show the sample standard deviation of the single estimate (as estimated from the 50 jackknife matrices). The black line corresponds to the log-normal input model of the simulations.

fixed. Following a Bayesian approach we take the probability density in the parameter space to be proportional to the likelihood,

$$p(\boldsymbol{\pi}) \sim \mathcal{L}(\boldsymbol{\pi}) \sim \exp\left(-\frac{1}{2}\chi^2[\boldsymbol{\pi}]\right), \quad (13.40)$$

where we assume our data vector $\hat{\boldsymbol{\xi}}$ to be Gaussian such that

$$\chi^2[\boldsymbol{\pi}] = (\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}[\boldsymbol{\pi}])^T C^{-1} (\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}[\boldsymbol{\pi}]). \quad (13.41)$$

Here $\boldsymbol{\xi}[\boldsymbol{\pi}]$ are our model predictions for $\langle \hat{\boldsymbol{\xi}} \rangle$ which we again compute with the NICAEA package. We are assuming a prior of $\Omega_m \in [0.1, 0.4]$ and $\sigma_8 \in [0.8, 1.1]$, which is well centered around our input cosmology. For C we will either insert the log-normal model covariance or the jackknife estimates of the covariance. We will de-bias the inverse of the latter in the way explained in section 13.4.6. Note that the reasoning in section 13.4.6 is in principle only valid for the pair-jackknife. And also for the pair-jackknife it is only valid in the case of almost uncorrelated sub-regions. We will nevertheless carry out the de-biasing in the same way for both jackknife schemes. Furthermore, we will also ignore the variance of the inverted covariance estimate (Taylor & Joachimi, 2014), as explained in the end of section 13.4. Our data vector $\hat{\boldsymbol{\xi}}$ will be either $\hat{\boldsymbol{\xi}}_+$ or $\hat{\boldsymbol{\xi}}_-$ or the joint data vector of both correlation functions, in which case we will also take into account the cross covariance between the two.

For each mock observation $\hat{\boldsymbol{\xi}}$ and for each available covariance matrix we use equation 13.40 to compute marginalized 1σ constraints on Ω_m and σ_8 , i.e. we consider the marginalized probability densities

$$\begin{aligned} p_\Omega(\Omega_m) &= \int d\sigma_8 p(\Omega_m, \sigma_8) \\ p_\sigma(\sigma_8) &= \int d\Omega_m p(\Omega_m, \sigma_8) \end{aligned} \quad (13.42)$$

and we define 1σ confidence interval to be that interval around the best fit parameter value which encloses $\sim 68\%$ of the probability and which has equal values of the probability density at each interval boundary⁹.

Because of the strong degeneracy between Ω_m and σ_8 (Kilbinger et al., 2013; Kilbinger & Schneider, 2004), even little uncertainties in the modeling of $\boldsymbol{\xi}[\boldsymbol{\pi}]$ ¹⁰ or in our simulations could shift the best-fit values of the parameters along the degeneracy. Fortunately, this does not affect our analysis because we only have to compare the constraints derived from the jackknife covariance estimates to the constraints obtained from the true (log-normal) covariance matrix. Furthermore, our results don't change noticeably, if instead of the log-normal covariance matrix we use the sample covariance estimated from 1000 simulations

⁹Without the last statement the definition of the 1σ confidence interval would be ambiguous.

¹⁰In our modeling we are for example not considering the finite bin width in our measurement of $\hat{\boldsymbol{\xi}}$.

(c.f. section 13.3). Hence in any case, our analysis provides a fair test of internal covariance estimators.

In figure 13.8 we show the mean values of the upper and lower boundaries on Ω_m and σ_8 as well as their mean best fit value for different numbers of jackknife re-samplings (red points and errorbars). The mean is taken with respect to all 50 confidence intervals we computed from the 50 mock observations. We also compare the jackknife constraints to those we get when using the true covariance matrix (blue lines). These figures only show the results for the galaxy-jackknife, which in the situation considered here yields the best agreement with the true covariance.

We compare galaxy-jackknife and pair-jackknife in figure 13.9. Here we show the mean width of the confidence intervals obtained with galaxy-jackknife, pair-jackknife and the true covariance matrix. For ξ_- , the width of the confidence intervals agrees well with the confidence intervals obtained from the true covariance matrix. This is because the covariance matrix of ξ_- is dominated by its shape noise component, which is very accurately captured by jackknife. In fact, even for the pair-scheme and even for 400 jackknife re-samplings the width of the confidence intervals from ξ_- alone is not underestimated. This seems to contradict figure 13.7, where the pair-scheme systematically underestimates the covariance. One reason for this is probably, that the variance in the inverted covariance estimate increases parameter uncertainties (Taylor & Joachimi, 2014). Note especially, that this is not the same effect as the de-biasing in eqn. 13.4.6. For ξ_+ , the strong underestimation of the covariance matrix by jackknife also leads to an underestimation of the uncertainties on Ω_m and σ_8 . Again one can see that the variance in the width of the confidence intervals (the errorbars in figure 13.9) becomes smaller, when more jackknife re-samplings are used. In turn, this increases the overall underestimation of the uncertainties. If both correlation functions are combined and 225 re-samplings are used, the parameter uncertainties are underestimated by $\sim 10\%$.

We have not shown results from the pair-jackknife estimates in figure 13.8, but the best fit values of Ω_m and σ_8 agree very well between the two jackknife schemes (i.e. within the green errorbars in figure 13.8), if only $\hat{\xi}_+$ or $\hat{\xi}_-$ are used to constrain the parameters. In figure 13.10 we compare the pair-jackknife and galaxy-jackknife best fit values when using the *full* data vector. Here the pair-jackknife seems to yield a stronger bias of the best fit values with respect to the true covariance.

The above results indicate that internal covariance estimation can reproduce the constraints on parameters from the true covariance quite well, especially when the galaxy-jackknife scheme is used. However, these results are not generalizable. In general, internal estimation of the covariance works best if the covariance matrix is shape noise dominated. Hence, the answer to what is the best estimation scheme and how well it can reproduce the true errorbars on cosmological parameters depends on the depth of the considered survey. A shallower survey not only has a smaller source density and hence a bigger shape noise. It also has a smaller convergence power spectrum which in turn reduces the cosmic variance part of the covariance.

The procedure we presented above to investigate the performance of internal covariance estimators thus has to be re-run for each survey under consideration. One can consider

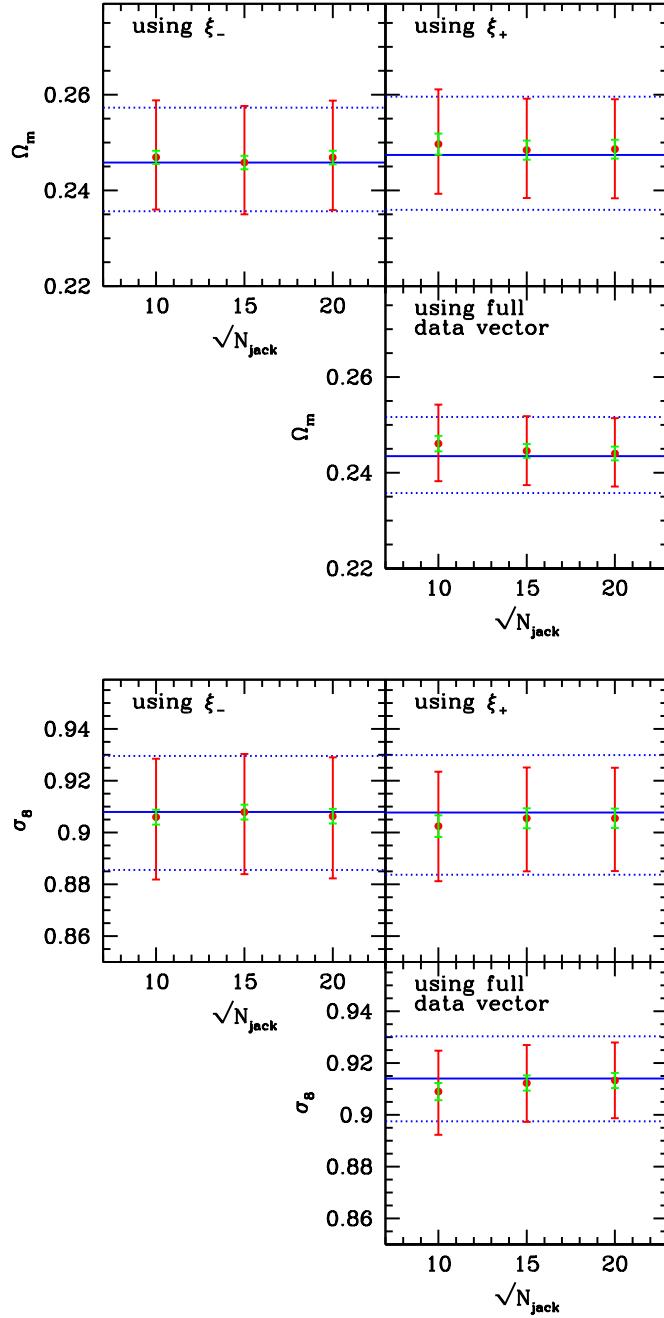


Figure 13.8: Mean 1σ constraints on Ω_m and σ_8 using galaxy-jackknife (red errorbars). The green errorbars show the standard deviation of the mean best-fit values (i.e. the standard deviation of the best fit values divided by $\sqrt{50}$). The blue lines indicate the constraints that are obtained when the true covariance is used in each mock catalog. Note that the error bars are very symmetric. For surveys as big as our simulations the constraining power becomes large enough to turn the - usually *banana* shaped - degeneracy between Ω_m and σ_8 into almost elliptical contours in the parameter plane (cf. appendix 13.B).

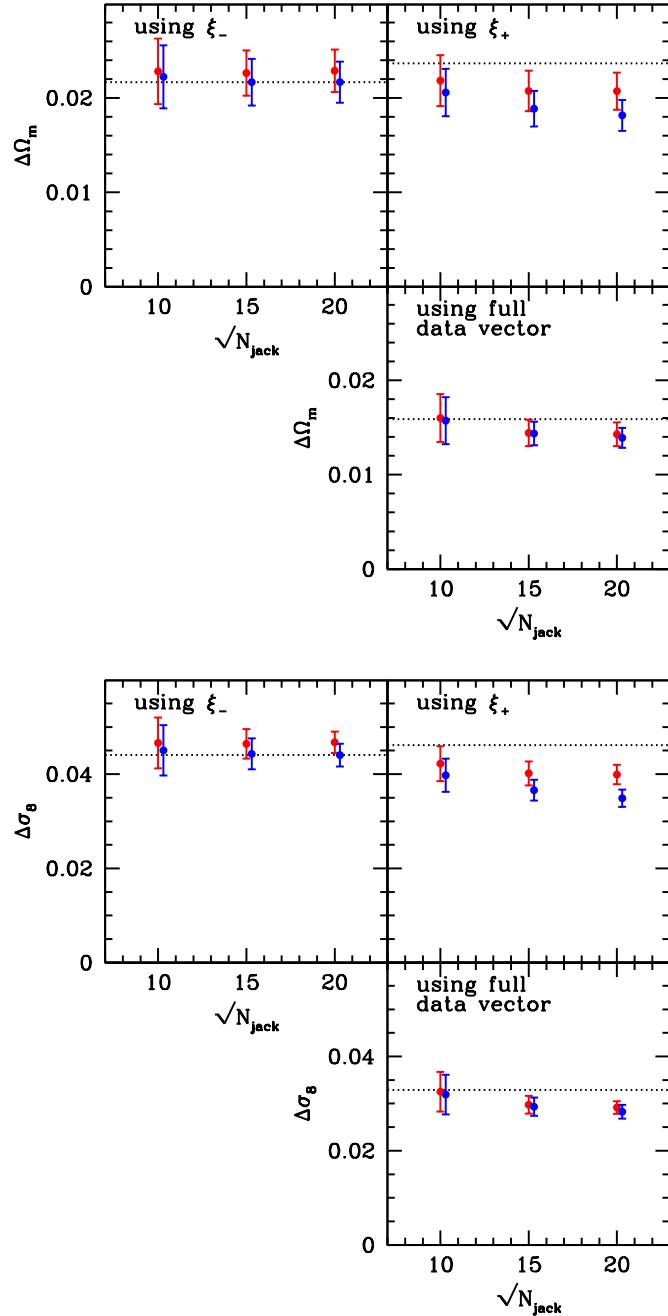


Figure 13.9: Mean width of the 1σ uncertainty on Ω_m and σ_8 using pair-jackknife (blue) and galaxy-jackknife (red). The errorbars show the standard deviation the 50 estimated confidence intervals. The black dotted line indicates the mean width of the confidence intervals when the true covariance is used in each mock catalog.

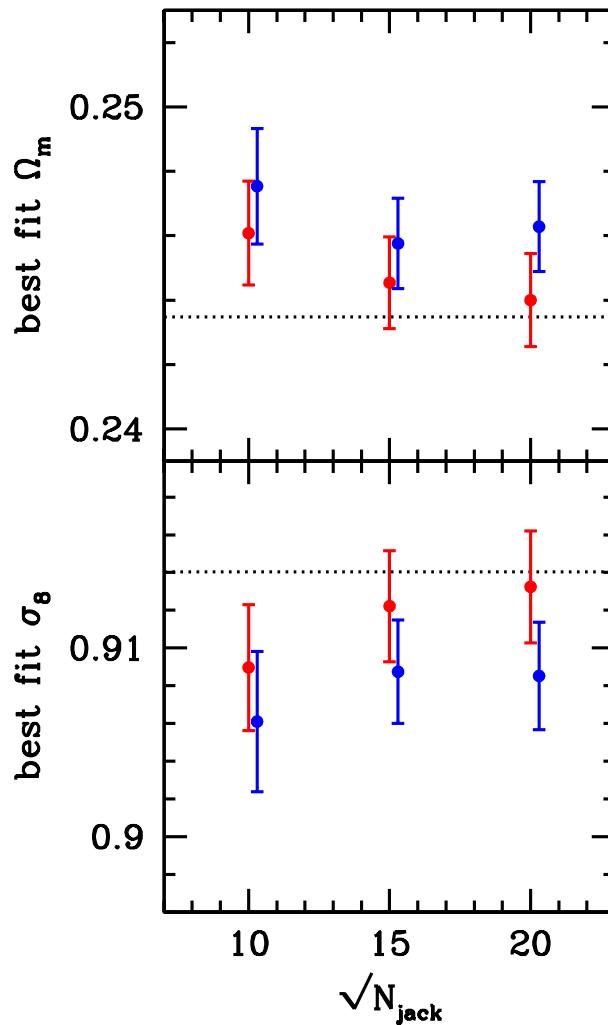


Figure 13.10: Mean best fit values of Ω_m and σ_8 using pair-jackknife (blue) and galaxy-jackknife (red). The errorbars show the standard deviation of the mean, as estimated from the 50 best fit values. The black dotted line indicates the mean best fit value when the true covariance is used in each mock catalog.

the log-normal model as a good model for the true covariance of our simulations for mock catalogs with an area of $\gtrsim 1000 \text{ deg}^2$ and a simple, connected geometry. For smaller surveys the finite-area-effect should not be ignored (Sato et al., 2011; Kilbinger et al., 2013). However, these surveys can be simulated fast enough with our public code to generate a large sample of independent realisations of the mock data which provides a good sample covariance estimate of the true covariance matrix. This estimate can then be compared to an ensemble of internal covariance estimates as we have done it above.

13.5.2 Matching the procedure to DES science verification and year 5 Data

We will now present an application of our method. Our attempt is to determine the performance of internal covariance estimation for

- setup IIa: Dark Energy Survey science verification data (DES-SV)
- setup IIb: DES year five data (DES-Y5) assuming a low source density
- setup IIc: DES year five data assuming a high source density.

For the area, shape noise, source density and source redshift distribution cf. table 13.1 and section 13.3.1. A source density of 10 arcmin^{-2} is forecasted for the final DES data while a density of $\sim 6 \text{ arcmin}^{-2}$ roughly corresponds to the current status of DES science verification data. Note also, that we are using a mask similar to the footprint of DES-SV to simulate mock shape catalogs for setup IIa. Setups IIb and IIc are simply simulated to be square shaped.

We adjust our data vector to that used by Becker et al. (2016), i.e. for both ξ_+ and ξ_- we now use 15 logarithmic bins ranging from $\theta = 2 \text{ arcmin}$ to $\theta = 300 \text{ arcmin}$. We will cut the survey into 100 sub-regions for setup IIa. Note that this way our biggest angular scales by far exceed the diameter of our subregions which is $\sim 45 \text{ arcmin}$. Hence, this can be considered an on-the-edge test of internal covariance estimators. A good tool to define sub-regions in an irregular survey geometry is the *kmeans* algorithm¹¹. For setups IIb and IIc we decide to split the survey into 225 sub-regions which corresponds to a diameter of $\sim 4.7 \text{ arcmin}$. This should give a more stable estimate of the covariance while still yielding much larger sub-regions than in setup IIa.

In figure 13.11 we compare the internal variance estimates to the true covariance. The latter is taken to be the log-normal model for the Y5 simulations and a sample variance computed from 1000 independent realizations for the SV simulations. Because of the fewer number of bins we are now using the procedure described in section 13.2.2 to compute the log-normal covariance matrix. As you can see, for $\hat{\xi}_-$ the pair-jackknife now becomes the best estimator of the variance. For $\hat{\xi}_+$ the situation is similar to what we have seen before, i.e. both schemes mostly underestimate the variance and the galaxy-jackknife is

¹¹implemented by Erin Sheldon for python, www.github.com/esheldon/kmeans_radec

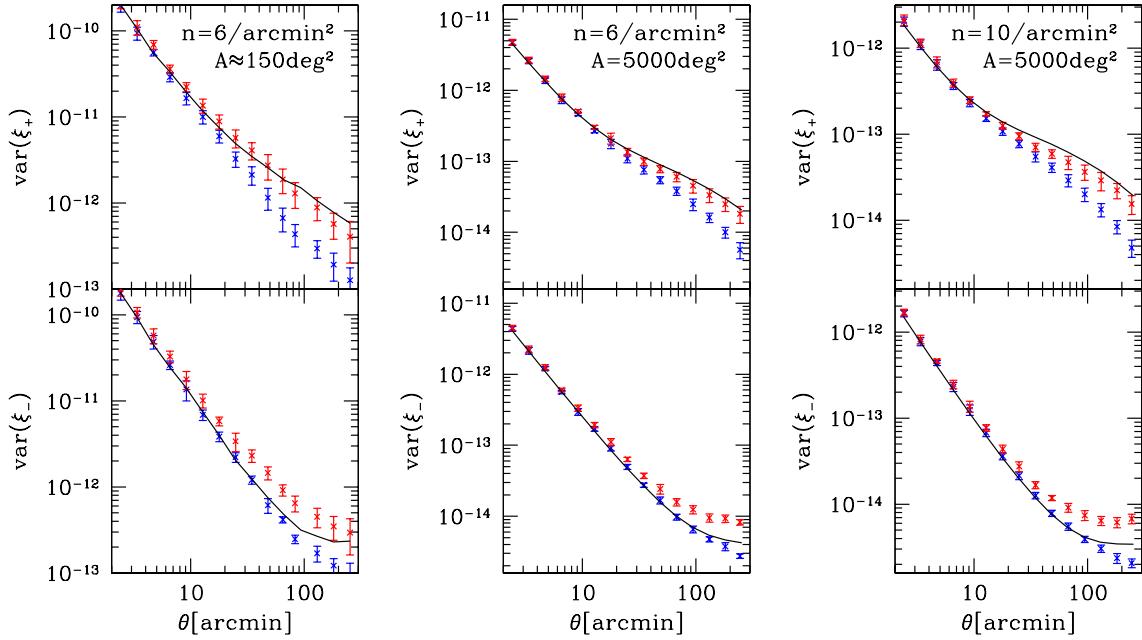


Figure 13.11: Variance estimates for DES-SV like data (left), DES-Y5 like data with a low density (middle) and with a high density (right). Red dots show the galaxy-jackknife estimates and blue dots the pair-jackknife estimates. For the Y5 case the log-normal model together with eqn. 13.18 was taken as a reference covariance (black lines) while for the SV case we estimated the true covariance from 1000 independent realizations of the mock data in order to account for the finite-area-effect. The errorbars indicate the standard deviation of the single estimates as obtained from 10 independent measurements.

overall closer to the true variance. Hence, judging from figure 13.11 we conclude that galaxy-jackknife should be used in order to not underestimate the true uncertainties in the data vector. However, these statements only hold for the diagonal elements of the covariance matrix. A convenient way to compare the complete covariance estimates is to derive likelihood contours from them in the desired parameter space.

We carry out a likelihood analysis in the Ω_m - σ_8 plane for the 10 simulations that have a Y5-like area and a source density of 6 arcmin⁻² which is the highest density currently achieved in DES science verification data (Becker et al., 2016). In figure 13.12 we show the likelihood contours obtained from one of the simulations when using galaxy-jackknife, pair-jackknife and the log-normal model for the covariance matrix. The contours were obtained from Monte-Carlo-Markov-Chains (150.000 steps) using the COSMOLIKE package by Eifler et al. (2014). We present the likelihood contours from the other 9 independent simulations in appendix 13.B. As expected, jackknife estimation underestimates the uncertainties. The input cosmology lies within the 1- σ contour in 6 of 10 simulation, when the log-normal covariance is used. It lies within the 1- σ contour in 5 of 10 simulation, when the covariance is estimated with jackknife (either scheme).

In table 13.2 we show the average ratio of the volume in the Ω_m - σ_8 plane enclosed by the 1- σ - and 2- σ -contours when using jackknife to that when using the true covariance matrix. Since the 1- σ - and 2- σ -ellipses obtained from jackknife and from the true covariance lie well on top of each other, this ratio can be considered as the fraction of the true uncertainties that is recovered by the jackknife covariance matrices. You can see from table 13.2 that the volume inside contours of constant likelihood in the Ω_m - σ_8 plane estimated with galaxy-jackknife is on average $\gtrsim 85\%$ of the true volume while the volume estimated with pair-jackknife recovers only $\gtrsim 70\%$ of the true volume. This agrees with the impression (from figures 13.15 and 13.16) that the contours obtained with galaxy-jackknife match better to the contours obtained from the true covariance. Note also, that the ellipses obtained from pair-jackknife have in some cases a strong off-set along the degeneracy between Ω_m and σ_8 compared to the true covariance and the galaxy-jackknife estimates. This is probably because pair-jackknife strongly underestimates the variance of $\hat{\xi}_\pm$ at large angular scales, which causes even small fluctuations at these scales to shift the contours considerably.

Finally, we want to see how well jackknife matrices recover the uncertainties perpendicular to the degeneracy between Ω_m and σ_8 . To do so, we consider the parameter combination

$$\Sigma_8 := \frac{\sigma_8}{0.9} \left(\frac{\Omega_m}{0.25} \right)^{0.5}. \quad (13.43)$$

Contours of constant Σ_8 are roughly parallel to the degeneracy that can be seen in figures 13.12, 13.15 and 13.16. For each of our 10 realizations we bin our MCMC's in Σ_8 to estimate its probability density. Table 13.3 displays the average ratio of the 1- σ and 2- σ uncertainties obtained from jackknife to the uncertainties obtained from the true covariance. This time, we find that galaxy-jackknife on average yields $\sim 90\%$ of the true uncertainties while pair-jackknife yields $\sim 85\%$. Hence, when the degeneracy between Ω_m and σ_8 is broken by other probes (such as the power spectrum of temperature fluctuation in the cosmic microwave

background) the performance of jackknife covariance matrices slightly improves.

Judging from the above numbers and from the contours in appendix 13.B we deem that $\gtrsim 85\%$ of the true uncertainties on Ω_m and σ_8 in a 2D cosmic shear analysis can be recovered without the use of large suits of N-body simulations or covariance models. When other probes like the CMB are used to break the degeneracy between the two parameters, the performance of jackknife even increases, because the deviations from the true covariance mostly take place along the direction of degeneracy between Ω_m and σ_8 .

13.6 Conclusions

We have explored the performance of internal covariance estimation for cosmic shear 2-pt. correlation functions. We devised two different jackknife schemes and explained in detail when these schemes underestimate the true covariance and when overestimation takes place. Furthermore, we explained why the sub-sample covariance and the pair-bootstrap covariance yield results that are very similar to jackknife estimation of the covariance matrix. Based on the pair-jackknife scheme we have argued that the Anderson-Hartlap-Kaufman (Kaufman, 1967; Hartlap et al., 2007) de-biasing factor should also be applied when inverting jackknife covariance matrices. Based on empirical findings we also recommend the use of this factor for the galaxy-jackknife scheme.

We have demonstrated our findings in an exemplary study using log-normal simulations of the convergence field and the corresponding shear field. We found the performance of all internal covariance estimators - except for the bootstrapping of galaxies - to be very similar. For the investigated cases, jackknife covariance matrices could provide accurate uncertainties on cosmological parameters as compared to the true covariance matrix of our simulations. Our conclusions regarding the two possible re-sampling schemes are the following:

- galaxy-bootstrap severely overestimates the covariance, which is in agreement with the finding of Norberg et al. (2009) for galaxy clustering correlation functions.
- from ξ_- alone, the pair-jackknife scheme reconstructs the parameter constraints most faithfully (cf. figure 13.8).
- from ξ_+ alone and when combining the two correlation functions, we find that the parameter constraints are best reconstructed by the galaxy-jackknife.

The performance of the galaxy-scheme turns out to be better in most situations, because two systematic errors (cf. sections 13.4.2 and 13.4.3) cancel each other partly in the that scheme. The pair-jackknife suffers from only one of these systematics and hence always yields lower (absolute) values for the covariance than the galaxy-jackknife and always underestimates the (absolute) values of the true covariance matrix.

Our results can not be generalized to arbitrary surveys, i.e. our paper rather demonstrates a general method to find a good covariance estimation scheme for any particular

| | galaxy-jackknife | pair-jackknife |
|---|------------------|-----------------|
| $V_{1\sigma,\text{jack}}/V_{1\sigma,\text{true}}$ | 0.86 ± 0.08 | 0.72 ± 0.09 |
| $V_{2\sigma,\text{jack}}/V_{2\sigma,\text{true}}$ | 0.87 ± 0.08 | 0.74 ± 0.09 |

Table 13.2: Ratio of the volume within the 1σ and 2σ contours in the $\Omega_m - \sigma_8$ plane obtained from jackknife and true covariance (setup IIb). The errors are given by the standard deviation of a sample of 10 independent simulations. The combined data vector of ξ_+ and ξ_- was used.

| | galaxy-jackknife | pair-jackknife |
|---|------------------|-----------------|
| $\Delta\Sigma_8 \text{ } 1\sigma,\text{jack} / \Delta\Sigma_8 \text{ } 1\sigma,\text{true}$ | 0.91 ± 0.08 | 0.86 ± 0.10 |
| $\Delta\Sigma_8 \text{ } 2\sigma,\text{jack} / \Delta\Sigma_8 \text{ } 2\sigma,\text{true}$ | 0.90 ± 0.08 | 0.85 ± 0.09 |

Table 13.3: Ratio of the 1σ and 2σ uncertainties on $\Sigma_8 \sim \sigma_8 \Omega_m^{0.5}$ obtained from jackknife and true covariance (setup IIb). The errors are given by the standard deviation of a sample of 10 independent simulations. Again, the combined data vector of ξ_+ and ξ_- was used.

survey. In making our simulation code public we provide our readers with a tool to redo the presented analyses for their desired set-up. As an application example we tested jackknife estimation of the covariance for a 2D cosmic shear analysis of the Dark Energy Survey. We found that for the complete, 5-year DES survey internal covariance estimators can provide reliable parameter constraints in a 2D cosmic shear analysis. We recommend a scheme of $\sim 15 \times 15$ jackknife re-samplings to yield a stable covariance matrix. Judging from figures 13.12, 13.15 and 13.16, we find as before that the likelihood contours in the $\Omega_m - \sigma_8$ plane are best reconstructed by the galaxy-jackknife scheme, if both correlation functions ξ_+ and ξ_- are combined. This way, on average $\gtrsim 85\%$ of the true uncertainties are captured by the internally estimated covariance matrix. If the degeneracy between Ω_m and σ_8 is broken, this value increases to $\sim 90\%$. Hence, up to $\sim 90\%$ of the true uncertainties in a 2D cosmic shear analysis can be provided from internally estimated covariance matrices.

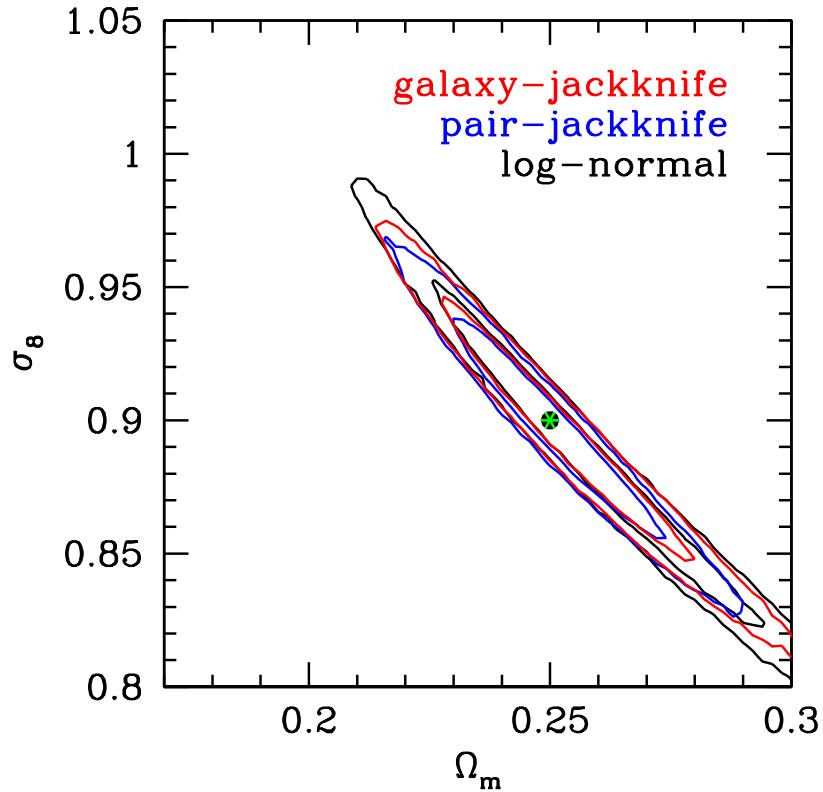


Figure 13.12: $1-\sigma$ and $2-\sigma$ contours in the Ω_m - σ_8 plane obtained from the two jackknife schemes (red and blue) and the true covariance (log-normal covariance, black) and using the combined data vector $(\hat{\xi}_+, \hat{\xi}_-)$. The input cosmology lies within the $1-\sigma$ contour in 6 of 10 simulation, when the log-normal covariance is used. It lies within the $1-\sigma$ contour in 5 of 10 simulation, when the covariance is estimated with jackknife (either scheme). In appendix 13.B we show the contours obtained from the other simulations. The underestimation of the uncertainties by jackknife mainly takes place along the direction of the degeneracy between Ω_m and σ_8 .

Acknowledgments

This work was supported by SFB-Transregio 33 ‘The Dark Universe’ by the Deutsche Forschungsgemeinschaft (DFG). We also acknowledge the support by the DFG Cluster of Excellence ”Origin and Structure of the Universe”. The simulations have been carried out on the computing facilities of the Computational Center for Particle and Astrophysics (C2PAP). Part of the research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

This paper has gone through internal review by the DES collaboration. We thank David Bacon, Gary Bernstein, Stefan Hilbert, Klaus Honscheid, Benjamin Joachimi and Bhuvnesh Jain for very helpful comments and discussions during the review process.

Funding for the DES Projects has been provided by the U.S. Department of Energy, the U.S. National Science Foundation, the Ministry of Science and Education of Spain, the Science and Technology Facilities Council of the United Kingdom, the Higher Education Funding Council for England, the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign, the Kavli Institute of Cosmological Physics at the University of Chicago, the Center for Cosmology and Astro-Particle Physics at the Ohio State University, the Mitchell Institute for Fundamental Physics and Astronomy at Texas A&M University, Financiadora de Estudos e Projetos, Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro, Conselho Nacional de Desenvolvimento Científico e Tecnológico and the Ministério da Ciência, Tecnologia e Inovação, the Deutsche Forschungsgemeinschaft and the Collaborating Institutions in the Dark Energy Survey. The DES data management system is supported by the National Science Foundation under Grant Number AST-1138766.

The Collaborating Institutions are Argonne National Laboratory, the University of California at Santa Cruz, the University of Cambridge, Centro de Investigaciones Enérgéticas, Medioambientales y Tecnológicas-Madrid, the University of Chicago, University College London, the DES-Brazil Consortium, the University of Edinburgh, the Eidgenössische Technische Hochschule (ETH) Zürich, Fermi National Accelerator Laboratory, the University of Illinois at Urbana-Champaign, the Institut de Ciències de l’Espai (IEEC/CSIC), the Institut de Física d’Altes Energies, Lawrence Berkeley National Laboratory, the Ludwig-Maximilians Universität München and the associated Excellence Cluster Universe, the University of Michigan, the National Optical Astronomy Observatory, the University of Nottingham, The Ohio State University, the University of Pennsylvania, the University of Portsmouth, SLAC National Accelerator Laboratory, Stanford University, the University of Sussex, and Texas A&M University.

The DES participants from Spanish institutions are partially supported by MINECO under grants AYA2012-39559, ESP2013-48274, FPA2013-47986, and Centro de Excelencia Severo Ochoa SEV-2012-0234. Research leading to these results has received funding from the European Research Council under the European Unions Seventh Framework Programme (FP7/2007-2013) including ERC grant agreements 240672, 291329, and 306478.

13.A Correlation Matrices and Constraints from empirical Covariance

To see how jackknife estimates of the covariance matrix capture the cross-correlations between different bins of $\hat{\xi}_{i\pm}$ we are looking at the *correlation matrix*. This matrix is given in terms of the covariance matrix elements Cov_{ij} as

$$\text{Corr}_{ij} = \frac{\text{Cov}_{ij}}{\sqrt{\text{Cov}_{ii}\text{Cov}_{jj}}} . \quad (13.44)$$

In the top panel of figure 13.13 we compare the correlation matrix obtained from the log-normal model covariance to the correlation matrix obtained from averaging all 50 jackknife estimates of the covariance matrix that were presented in section 13.5, i.e. using setup I from table 13.1 (we show here the jackknife with 400 re-samplings which divided the survey into the smallest sub-regions). The lower left corner shows the auto correlations of $\hat{\xi}_+$ and the upper right corner shows the auto correlations of $\hat{\xi}_-$. The upper left and lower right corners show the cross-correlations between the two correlation functions. Furthermore, the lower right half of the plot shows the correlations obtained from the log-normal model and the upper left half shows the correlations obtained from the average jackknife covariance estimate. Each column and row of pixels represents one angular bin and the bins range from $1'$ to $150'$, starting on the lower left corner.

The top panel of figure 13.13 indicates, that jackknife is able to capture the general structure of the correlation matrix of the 2-pt correlation functions. Given that internal covariance estimators mostly underestimate the variance of $\hat{\xi}_{\pm}$ one can hence conclude that the covariance elements Cov_{ij} are approximately underestimated by the same amount as the square root of $\text{Cov}_{ii}\text{Cov}_{jj}$ (cf. eqn. 13.44).

In the bottom panel of figure 13.13 we show the same plot but using the empirical covariance matrix obtained from 1000 independent realizations of setup I in the lower right half of the plot. The empirical covariance matrix is obviously noisier than the log-normal model covariance matrix. In order to confirm, that there are nevertheless no significant deviations of our simulations from the log-normal input model, we show in figure 13.14 again the constraints on Ω_m and σ_8 that were presented in figure 13.8, but this time using the empirical covariance to compute the reference constraints.

13.B Likelihood Contours

Figures 13.15 and 13.16 show the 1- and 2- σ contours in the Ω_m - σ_8 plane computed with COSMOLIKE when using galaxy-jackknife and pair-jackknife to estimate the covariance matrix (red and blue lines) and compare them to the same contours obtained from the true covariance matrix (black lines). The simulations are configured to mimic the complete, 5 year Dark Energy Survey (cf. section 13.5.2 or table 13.1, setup IIb). The only thing that differs from simulation to simulation is the random seed that was used to generate the

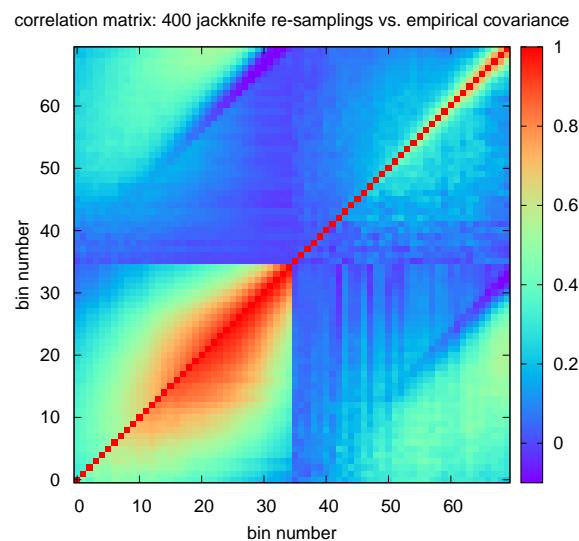
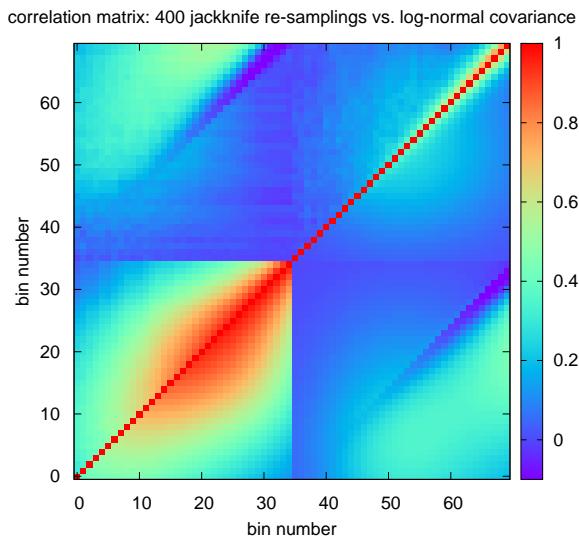


Figure 13.13: Correlation matrix of $\hat{\xi}_\pm$. Bins 0 to 34 belong to $\hat{\xi}_+$ and bins 35 to 69 belong to $\hat{\xi}_-$. For each 2-pt function the bins range from $1'$ to $150'$, starting on the lower left corner. Top panel: the lower right half of the plot displays the correlation coefficients of $\hat{\xi}_\pm$ obtained from the log-normal model and the upper left half displays the correlation coefficients obtained from the average jackknife covariance estimate (using 400 re-samplings). Bottom panel: same plot, but this time the empirical covariance from 1000 realizations of setup I. Note that we are averaging over 50 different jackknife estimates of the covariance matrix here, which makes the empirical covariance matrix seem noisier than the jackknife covariance.

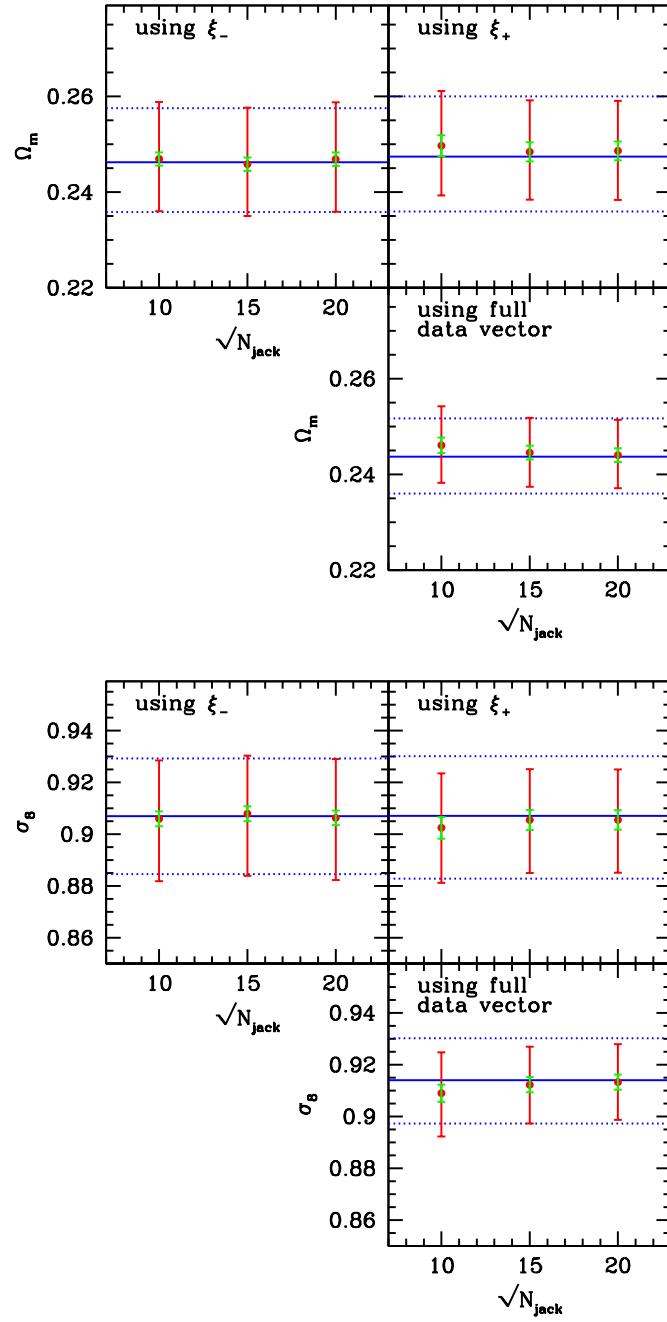


Figure 13.14: These plots are identical to the ones shown in figure 13.8, except that this time the blue lines indicate the mean best fit values and $1 - \sigma$ constraints obtained from the empirical covariance matrix, i.e. from the sample covariance of 1000 independent realizations of setup I.

log-normal fields and the shape noise. The green dots represent the input cosmology of the simulations.

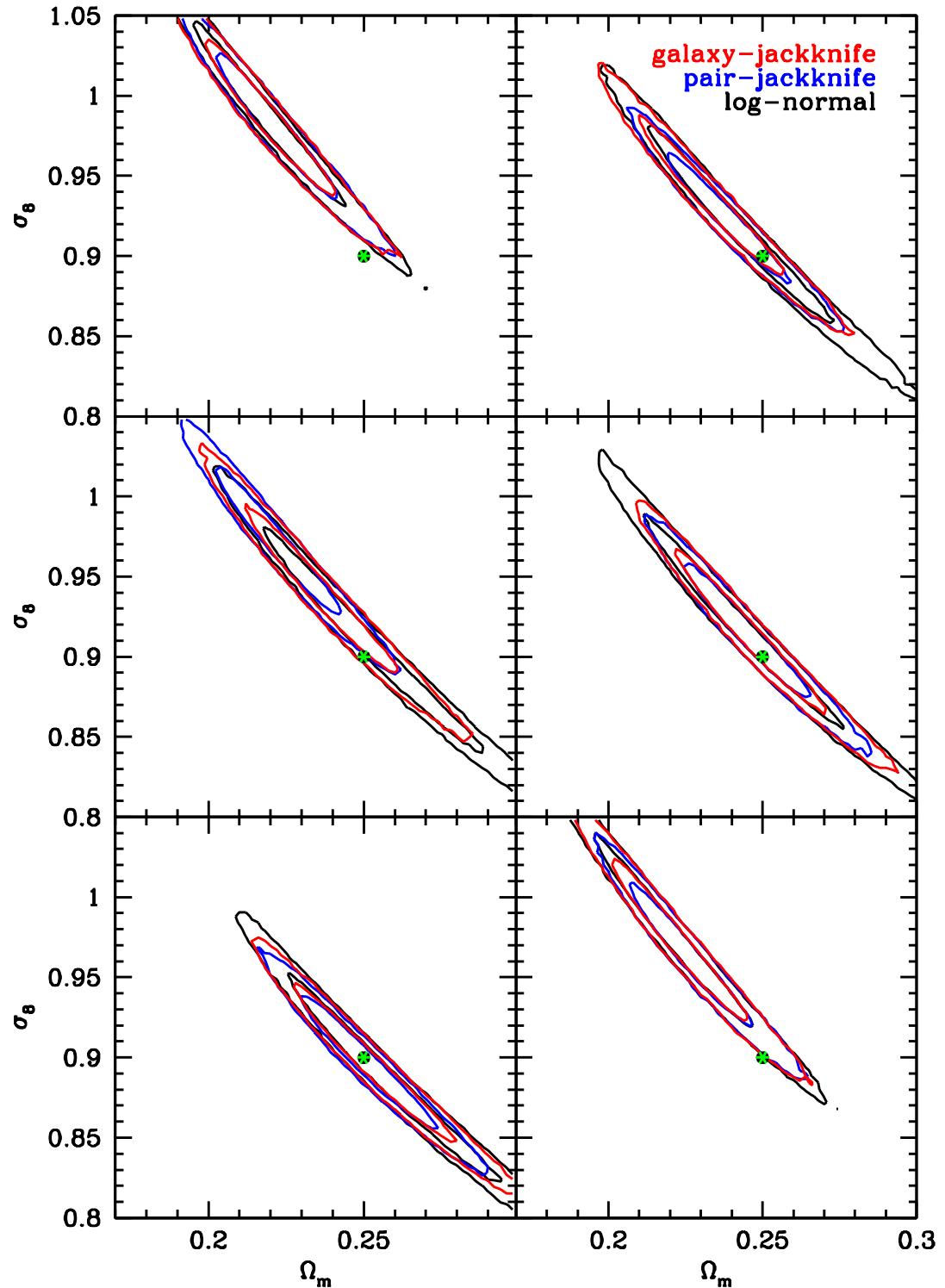


Figure 13.15: 1- and 2- σ contours in the Ω_m - σ_8 plane obtained from the first 6 simulations.

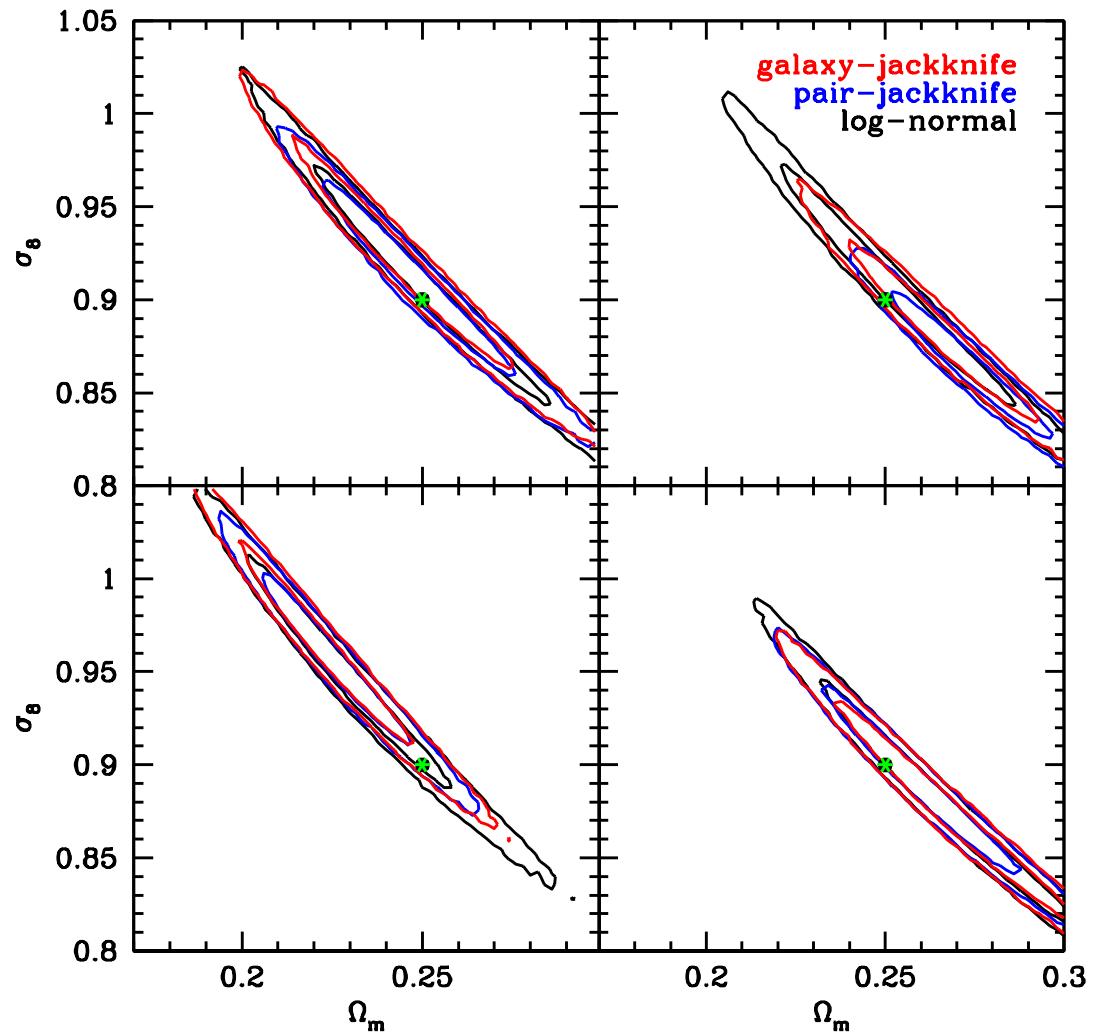


Figure 13.16: 1- and 2- σ contours in the Ω_m - σ_8 plane obtained from the remaining 4 simulations.

Chapter 14

Precision matrix expansion – efficient use of numerical simulations in estimating errors on cosmological parameters

The data we are gathering about the large scale structure of the universe is increasing rapidly, and so is the size of data vectors used in cosmological analyses (~ 400 data points were used in the analysis of DES Collaboration et al. 2017). To estimate the inverse covariance matrix (the *precision matrix*) of such huge data vectors from simulated data, one would require thousands of high resolution N-body simulations. In the following chapter I show, how this requirement can be reduced drastically when using information from a theoretical first guess of the covariance matrix.

This chapter has been published as Friedrich & Eifler (2018) in MNRAS. I have exclusively developed the method presented in this article and carried out the analysis of its performance. Tim Eifler provided the cosmology tool package `CosmoLike` that was used to perform the simulated likelihood analyses and he computed the model covariance matrices used in this work. His comments also helped to significantly improve an early draft of the article.

Permission for non-commercial re-use of the material included in this thesis has been granted by the MNRAS editorial office. Oxford University Press holds the copyright on the paper.

ABSTRACT

Computing the inverse covariance matrix (or precision matrix) of large data vectors is crucial in weak lensing (and multi-probe) analyses of the large scale structure of the universe. Analytically computed covariances are noise-free and hence straightforward to invert, however the model approximations might be insufficient for the statistical precision of future cosmological data. Estimating covariances from numerical simulations improves on these approximations, but the sample covariance estimator is inherently noisy, which introduces uncertainties in the error bars on cosmological parameters and also additional scatter in their best fit values. For future surveys, reducing both effects to an acceptable level requires an unfeasibly large number of simulations.

In this paper we describe a way to expand the true precision matrix around a covariance model and show how to estimate the leading order terms of this expansion from simulations. This is especially powerful if the covariance matrix is the sum of two contributions, $\mathbf{C} = \mathbf{A} + \mathbf{B}$, where \mathbf{A} is well understood analytically and can be turned off in simulations (e.g. shape-noise for cosmic shear) to yield a direct estimate of \mathbf{B} . We test our method in mock experiments resembling tomographic weak lensing data vectors from the Dark Energy Survey (DES) and the Large Synoptic Survey Telescope (LSST). For DES we find that 400 N-body simulations are sufficient to achieve negligible statistical uncertainties on parameter constraints. For LSST this is achieved with 2400 simulations. The standard covariance estimator would require $> 10^5$ simulations to reach a similar precision. We extend our analysis to a DES multi-probe case finding a similar performance.

14.1 Introduction

Wide area surveys such as the currently running Dark Energy Survey (DES, Flaugher, 2005) or the upcoming Large Synoptic Survey Telescope (LSST, Ivezić et al., 2008) will collect vast amounts of data about the large scale structure of the universe. In cosmological analyses this data can e.g. be compressed into measurements of 2-point correlation functions of galaxy clustering or cosmic shear. In a redshift-tomographic analysis this will easily accumulate to data vectors with several hundreds of data points. Testing cosmological models from a measurement of such a large data vector requires precise knowledge of the inverse covariance matrix of the noise in this data vector. There has been extensive research on the impact of errors associated with covariance estimation on the constraints derived on cosmological parameters. Hartlap et al. (2007) discussed the fact that the inverse of an unbiased covariance estimator is not an unbiased estimator for the inverse covariance matrix (*the precision matrix*). They also described a way to correct for this when assuming that the covariance estimate follows a Wishart distribution (see also Kaufman 1967 and Anderson 2003). The noise properties of this corrected precision matrix estimator and its impact on the constraints derived on cosmological parameters was e.g. investigated by Taylor et al. (2013); Dodelson & Schneider (2013); Taylor & Joachimi (2014).

Sellentin & Heavens (2016, hereafter SH16a) have presented a different approach: given a covariance estimate they marginalize over the posterior distribution of the true precision matrix to compute the likelihood in parameter space. Assuming that the covariance estimate follows a Wishart distribution they have derived a simple, closed-form expression for the resulting likelihood function. In Sellentin & Heavens (2017) they have extended these results to derive the information loss in parameter space due to noisy covariance estimates. A fully non-Gaussian treatment of the effects discussed in Dodelson & Schneider (2013, hereafter DS13) is however still missing.

Prior knowledge on the sparsity of the covariance matrix and the precision matrix was used by Paz & Sánchez (2015) and Padmanabhan et al. (2016) to improve estimates of the precision matrix from few simulations. Pope & Szapudi (2008) investigated shrinkage estimators of the covariance, i.e. a mixing of estimated and modeled covariance matrices. This however raises the task of finding an equivalent to the Kaufman-Hartlap correction for such a mixture of estimated and analytic matrices. More recently, Joachimi (2017) describes a non-linear extension of that estimator which combines covariance estimates from two sets of independent data vector realizations and hence does not require a covariance model.

In this paper we describe a way to expand the true precision matrix around a covariance model as a power series in the deviation between model and true covariance. Assuming a Wishart realization for the true covariance and using the results on invariant moments of the Wishart distribution by Letac & Massam (2004) we derive an unbiased estimator for the up to second order expansion of the true precision matrix. This becomes especially powerful if parts of the covariance matrix that are well understood analytically can be turned off in simulations in order to yield a direct estimate of the remaining covariance parts. In Sect. 14.3 we recap the main problems of estimating parameter constraints from noisy covariance estimates and present our method of "Precision Matrix Expansion" (PME). In Sect. 14.4 we perform numerical experiments that mimic data from the Dark Energy Survey (DES) and the Large Synoptic Survey Telescope (LSST) likelihood analyses to test the performance of our idea. Sect. 14.5 concludes with a discussion of our results.

14.2 Parameter constraints from noisy covariance estimates

We begin by outlining the main task of this paper. Let $\hat{\boldsymbol{\xi}}$ be a vector of N_d data points measured from observational data and let $\boldsymbol{\xi}[\boldsymbol{\pi}]$ be a model for this data vector that depends on a vector of N_p parameters $\boldsymbol{\pi}$. If \mathbf{C} is the covariance matrix of $\hat{\boldsymbol{\xi}}$ then a standard way to constrain the parameters $\boldsymbol{\pi}$ is to assign a posterior distribution $p(\boldsymbol{\pi}|\hat{\boldsymbol{\xi}})$ to them as

$$p(\boldsymbol{\pi}|\hat{\boldsymbol{\xi}}) \sim \exp\left(-\frac{1}{2}\chi^2[\boldsymbol{\pi} | \hat{\boldsymbol{\xi}}, \mathbf{C}]\right) p(\boldsymbol{\pi}) \quad (14.1)$$

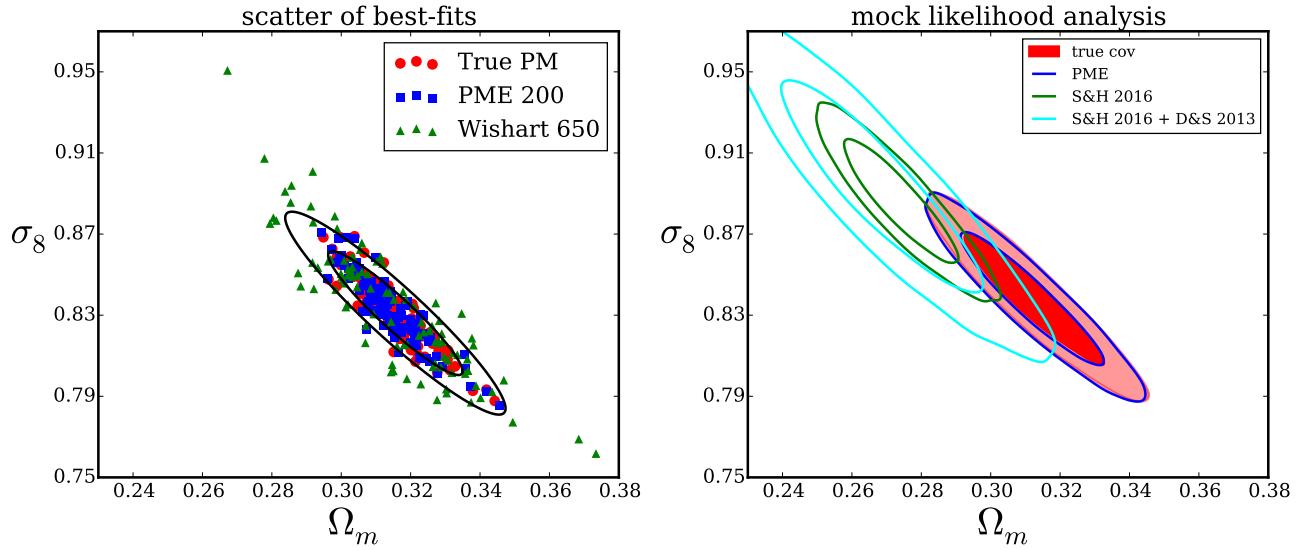


Figure 14.1: Left: Best fit parameter pairs (Ω_m, σ_8) obtained from random realizations of a DES-like weak lensing data vector with 450 data points when using different approaches to compute the precision matrix. The red points assume that the true covariance matrix is known while for the green points we draw a Wishart realization of the covariance ($N_s = 450 + 200 = 650$ simulations) for each data vector. The blue points are obtained with the method of precision matrix expansion (and allowing only 200 simulations to estimate the expansion). The black contours display the 1σ and 2σ Fisher contours derived from our fiducial covariance. Right: For one of the random realizations we perform a complete likelihood analysis and show the 1σ and 2σ contours in the $\Omega_m - \sigma_8$ plane after marginalizing over w_0 and w_a (see Sec. 14.4 for details). The contours obtained from the Wishart realization of the covariance are clearly offset from those obtained from the true covariance matrix. We recommend to account for this by expanding the likelihood around its maximum (of the full parameter space, which in this figure is 4-dimensional) with the factor derived by DS13. This leads to a decreased constraining power of our mock survey. The use of PME manages to significantly decrease this contour offset.

with

$$\chi^2 [\boldsymbol{\pi} \mid \hat{\boldsymbol{\xi}}, \mathbf{C}] = (\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}[\boldsymbol{\pi}])^T \mathbf{C}^{-1} (\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}[\boldsymbol{\pi}]) \quad (14.2)$$

and $p(\boldsymbol{\pi})$ being a prior density incorporating a priori knowledge or assumptions on $\boldsymbol{\pi}$. These expressions in fact ignore that \mathbf{C} also can be dependent on $\boldsymbol{\pi}$. We will do this throughout this paper and refer the reader to Eifler et al. (2009) who investigated the impact of cosmology dependent covariance matrices on cosmic shear likelihood analyses. Another assumption that goes into Eq. 14.1 is that the measured data vector $\hat{\boldsymbol{\xi}}$ is drawn from a multi-variate Gaussian distribution. In wide area surveys this is justified in the limit where one can consider the survey to consist of many independent sub-regions, such that the measurements in those regions add up to a Gaussian data vector by means of the central limit theorem.

If the covariance matrix \mathbf{C} is not exactly known, it can e.g. be estimated from N-body simulations. If $\hat{\boldsymbol{\xi}}_i$, $i = 1 \dots N_s$, are a number of independent measurements of $\boldsymbol{\xi}$ in simulations then an unbiased estimate of \mathbf{C} is given by

$$\hat{\mathbf{C}} := \frac{1}{\nu} \sum_{i=1}^{N_s} (\hat{\boldsymbol{\xi}}_i - \bar{\boldsymbol{\xi}}) (\hat{\boldsymbol{\xi}}_i - \bar{\boldsymbol{\xi}})^T, \quad (14.3)$$

where $\nu = N_s - 1$ and $\bar{\boldsymbol{\xi}}$ is the sample mean of the $\hat{\boldsymbol{\xi}}_i$. We will assume $\hat{\mathbf{C}}$ to have a Wishart distribution with ν degrees of freedom which follows from our assumption that $\hat{\boldsymbol{\xi}}$ and the $\hat{\boldsymbol{\xi}}_i$ are Gaussian distributed (cf. Taylor et al. 2013).

To compute the likelihood in Eq. 14.1 we need to know the precision matrix, i.e. is the inverse covariance matrix $\boldsymbol{\Psi} = \mathbf{C}^{-1}$. According to Kaufman (1967, see also Hartlap et al. 2007; Taylor et al. 2013) an unbiased estimator for $\boldsymbol{\Psi}$ can be constructed from $\hat{\mathbf{C}}$ as

$$\hat{\boldsymbol{\Psi}} = \frac{\nu - N_d - 1}{\nu} \hat{\mathbf{C}}^{-1} \quad (14.4)$$

and we will call the factor of $(\nu - N_d - 1)/\nu$ the Kaufman-Hartlap-correction (N_d being again the dimension of the data vector).

Given a measurement $\hat{\boldsymbol{\xi}}$ of the data vector one can derive the posterior density of the model parameters $p(\boldsymbol{\pi} \mid \hat{\boldsymbol{\xi}})$ by means of equations 14.1 and 14.2. A noisy precision matrix estimate influences this inference in two ways:

- it adds noise to the width of likelihood contours derived from inserting the precision matrix estimate into the figure of merit χ^2 (Eq. 14.2).
- it adds noise to the location of likelihood contours. E.g. the maximum likelihood estimator for the parameters $\boldsymbol{\pi}$ would be

$$\hat{\boldsymbol{\pi}}_{\text{ML}} = \min_{\boldsymbol{\pi}} \left\{ (\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}[\boldsymbol{\pi}])^T \hat{\boldsymbol{\Psi}} (\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}[\boldsymbol{\pi}]) \right\}. \quad (14.5)$$

When using a noisy precision matrix the uncertainties of $\hat{\boldsymbol{\pi}}_{\text{ML}}$ have contributions from both the noise in $\hat{\boldsymbol{\xi}}$ and the noise in $\hat{\boldsymbol{\Psi}}$.

The astro-statistics literature has so far focused on the first effect, i.e. on the uncertainties on contour width due to noise in the estimate $\hat{\Psi}$ (Taylor et al., 2013; Taylor & Joachimi, 2014; Sellentin & Heavens, 2016, 2017). Sellentin & Heavens (2017) provide the most complete demonstration that $\hat{\Psi}$ yields a good estimate of the width of the posterior contours as long as $N_s - N_d \gg N_p$.

The more critical effect however is the additional noise of $\hat{\pi}_{\text{ML}}$. DS13 (also see appendix 14.A) showed that the uncertainty on the position of likelihood contours from noise in $\hat{\Psi}$ is only negligible if $N_s - N_d \gg N_d - N_p$ which is a much more demanding criterion for current cosmological data vectors. We demonstrate this in the left-hand panel of Fig. 14.1, where we show 100 randomly drawn realizations of a DES-like weak lensing data vector with $N_d = 450$ and a halo model covariance matrix (see Sec. 14.4 for further details). For each of the 100 data vectors we have also generated Wishart realizations of our covariance matrix corresponding to an estimate from $N_s = 650$ simulations. Using either the true covariance or the estimated one, we then determine the best fitting parameters Ω_m and σ_8 (after marginalizing over equation-of-state parameters of dark energy, w_0 and w_a). The best-fits obtained from a noisy covariance (green points) clearly display a much larger scatter than those obtained from the true covariance (red points). Also shown are the best fits obtained by precision matrix expansion (PME, blue points) which we are going to introduce in the next section. Here we assumed that only $N_s = 200$ simulations are available to estimate the PME, which gives best fit values that are significantly closer to the ones obtained when knowing the true covariance matrix.

When reconstructing $p(\boldsymbol{\pi}|\hat{\xi})$ (e.g. from a Monte-Carlo-Markov-Chain) this can lead to significant offsets between likelihood contours inferred from the true covariance matrix and likelihood contours inferred from a covariance estimate – even if the overall width of the likelihood contours is captured well by the covariance estimate. We demonstrate this in the right-hand panel of figure 14.1. DS13 have derived a factor (see appendix 14.A) by which parameter contours obtained from a Wishart realization of the covariance should be expanded in order to account for this additional scatter. However, their derivation relies on the assumption of a Gaussian parameter likelihood and is only applicable to the extent that a Fisher analysis is accurate. The current state of the art for dealing with noisy covariance estimates is hence a combination of SH16a and DS13: expanding the contours derived from the SH16a likelihood by the DS13 factor. We implement this idea for the cyan contours in Fig. 14.1 and show that this brings the contours derived from a standard covariance estimate into consistency with those derived from the true covariance.

Downsides of this approach are a large increase of the uncertainties on cosmological parameters and the fact that one still needs at least as many realizations as data points in the data vector to even derive a precision matrix estimate. We now want to introduce an alternative method to estimate the precision matrix which is able to drastically decrease the offset of contours seen for the standard precision matrix estimator.

14.3 Precision matrix expansion

Let us split the covariance matrix \mathbf{C} into two contributions

$$\mathbf{C} = \mathbf{A} + \mathbf{B} , \quad (14.6)$$

where for matrix \mathbf{A} we have an accurate model (e.g. the shape-noise contributions to the covariance of cosmic shear correlation functions) and for \mathbf{B} we have a model \mathbf{B}_m which we know to be imperfect. We want to include this prior knowledge of the covariance matrix when estimating the precision matrix. Starting from

$$\mathbf{C} = \mathbf{M} + (\mathbf{B} - \mathbf{B}_m) , \quad (14.7)$$

where $\mathbf{M} = \mathbf{A} + \mathbf{B}_m$ is our model for the complete covariance matrix, we rewrite

$$\mathbf{C} = (\mathbb{1} + \mathbf{X}) \mathbf{M} , \quad (14.8)$$

where $\mathbb{1}$ is the identity matrix and we have defined

$$\mathbf{X} := (\mathbf{B} - \mathbf{B}_m) \mathbf{M}^{-1} . \quad (14.9)$$

The precision matrix $\Psi = \mathbf{C}^{-1}$ can then be expressed as the following power series in \mathbf{X} :

$$\begin{aligned} \Psi &= \mathbf{M}^{-1} \left(\sum_{k=0}^{\infty} (-1)^k \mathbf{X}^k \right) \\ &= \mathbf{M}^{-1} (\mathbb{1} - \mathbf{X} + \mathbf{X}^2 + \mathcal{O} [\mathbf{X}^3]) . \end{aligned} \quad (14.10)$$

We will call this series the *precision matrix expansion* (PME). In appendix 14.C we show that it converges under a wide range of conditions. There we also demonstrate that the series yields at each order a symmetric approximation of Ψ and that at second order it is always positive definite (at each order if the series converges).

14.3.1 Estimating the expansion of Ψ

Suppose we have an estimate $\hat{\mathbf{B}}$ of the matrix \mathbf{B} from a number of N-body simulations. This especially assumes that all covariance contributions included in \mathbf{A} can be turned off in the simulations (i.e. for cosmic shear covariances \mathbf{A} could consist of shape-noise contributions which can be set to zero in simulations). We want to use $\hat{\mathbf{B}}$ to construct unbiased estimators for the first order and second order term of the series in Eq. 14.10.

Our assumptions state that $\hat{\mathbf{B}}$ is drawn from a Wishart distribution with expectation value \mathbf{B} . In this case also $\mathbf{M}^{-1} \hat{\mathbf{B}} \mathbf{M}^{-1}$ is Wishart distributed but with the expectation value $\mathbf{M}^{-1} \mathbf{B} \mathbf{M}^{-1}$. Hence an unbiased estimator for the first order PME is given by

$$\hat{\Psi}_{1st} = \mathbf{M}^{-1} - \mathbf{M}^{-1} (\hat{\mathbf{B}} - \mathbf{B}_m) \mathbf{M}^{-1} . \quad (14.11)$$

Note that this does not involve the inversion of an estimated matrix. According to Taylor et al. (2013) the standard deviation of diagonal elements of an inverse-Wishart distributed matrix is proportional to $1/\sqrt{N_s - N_d - 4}$ while for Wishart distributed matrices it is only proportional to $1/\sqrt{N_s - 1}$. Hence, avoiding the occurrence of an inverted matrix estimate greatly reduces the estimation noise.

The second order term involves squares of Wishart matrices. Using the results of Letac & Massam (2004) on invariant moments of the Wishart distribution (cf. appendix 14.B) it is still possible to construct an unbiased estimator for the second order PME as

$$\begin{aligned} \hat{\Psi}_{2\text{nd}} = & \mathbf{M}^{-1} + \mathbf{M}^{-1} \mathbf{B}_m \mathbf{M}^{-1} \mathbf{B}_m \mathbf{M}^{-1} \\ & - \mathbf{M}^{-1} (\hat{\mathbf{B}} - \mathbf{B}_m) \mathbf{M}^{-1} \\ & - \mathbf{M}^{-1} \hat{\mathbf{B}} \mathbf{M}^{-1} \mathbf{B}_m \mathbf{M}^{-1} \\ & - \mathbf{M}^{-1} \mathbf{B}_m \mathbf{M}^{-1} \hat{\mathbf{B}} \mathbf{M}^{-1} \\ & + \mathbf{M}^{-1} \frac{\nu^2 \hat{\mathbf{B}} \mathbf{M}^{-1} \hat{\mathbf{B}} - \nu \hat{\mathbf{B}} \text{ tr}(\mathbf{M}^{-1} \hat{\mathbf{B}})}{\nu^2 + \nu - 2} \mathbf{M}^{-1}. \end{aligned} \quad (14.12)$$

The estimator in Eq. 14.12 is the key result of our paper. It has two advantages over the Anderson-Hartlap corrected standard estimator. First, it only requires matrix multiplications. As a consequence, it can even be used if $N_s \leq N_d$. Second, it only needs an estimate of \mathbf{B} instead of the whole covariance \mathbf{C} , i.e. it allows to incorporate a priori knowledge on the covariance in the form of \mathbf{M} (and \mathbf{A}).

In the next section we demonstrate that this significantly eases the requirement of $N_s - N_d \gg N_d - N_p$. Hence, in a likelihood analysis the noise in $\hat{\Psi}_{2\text{nd}}$ becomes negligible for a much smaller number of N-body simulations than required by the standard precision matrix estimator. In appendix 14.C we also show that the bias in parameter constraints which arises from cutting the power series in Eq. 14.10 after a finite number of terms is negligible even for very strong deviations of our covariance model \mathbf{M} from the N-body covariance \mathbf{C} .

14.4 Examples: parameter errors for LSST weak lensing and DES weak lensing and multi-probe analyses

We investigate the performance of our method in the context of ongoing and future surveys using DES and LSST as specific examples. These surveys differ in terms of survey area, galaxy number density, and redshift distribution and have different demands on the precision matrix. For DES we consider summary statistics in real space, i.e. auto- and cross-correlation functions of galaxy shear and position, for LSST we consider the corresponding Fourier quantities of a shear-shear only data vector. A summary of the scenarios considered is given in Table 14.1 and a more detailed description of the considered data vectors is given in appendix 14.D.

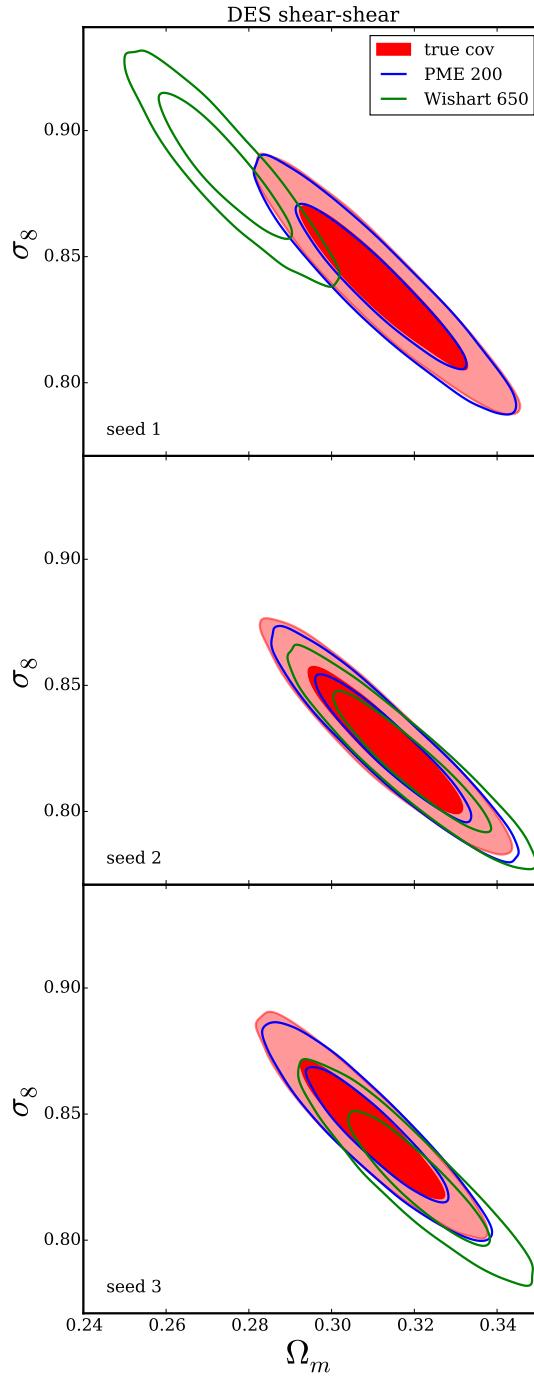


Figure 14.2: Contours in the Ω_m - σ_8 plane obtained from realizations of our DES-like weak lensing data vector after marginalizing over all other parameters. For each random seed also new Wishart realizations $\hat{\mathbf{B}}$ and $\hat{\mathbf{C}}$ of the matrices \mathbf{B} and \mathbf{C} were drawn in order to simulate new realizations of the second order PME estimator and the standard precision matrix estimator. $N_s = 200$ simulations were assumed for the estimation of the PME while $N_s = N_d + 200 = 650$ simulations were assumed for the standard estimator.

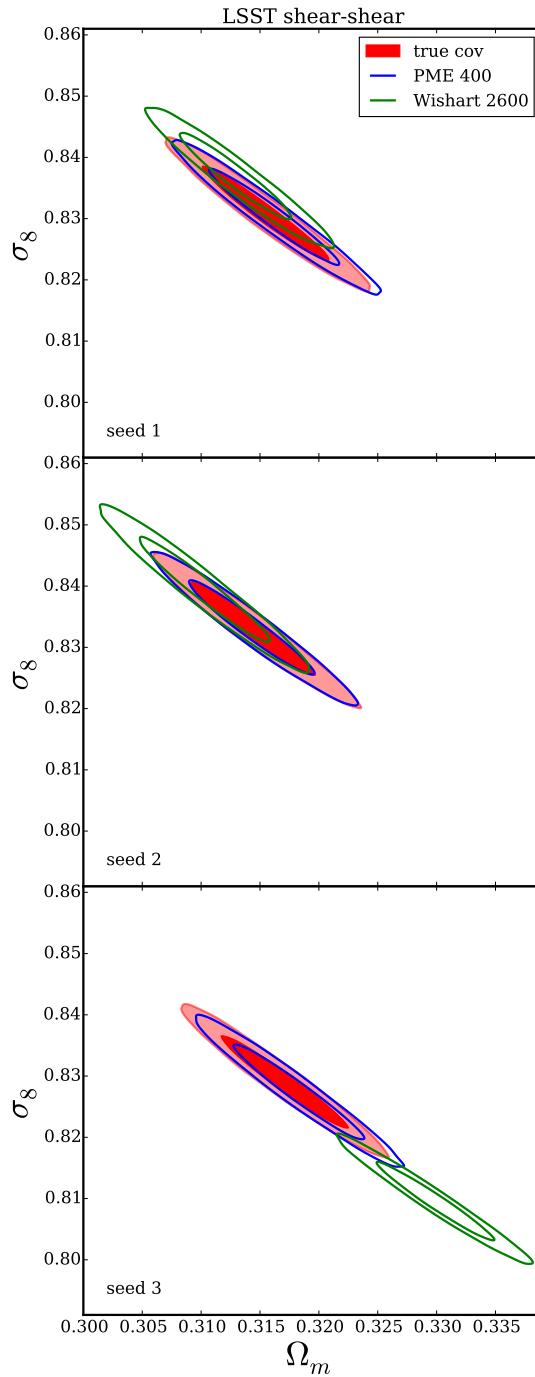


Figure 14.3: Same as Fig. 14.2 but for the LSST-like weak lensing data vector. $N_s = 400$ simulations where assumed for the estimation of the PME while $N_s = N_d + 400 = 2600$ simulations where assumed for the standard estimator.

| setup | survey | lens bins | source bins | N_{data} | data |
|-------|--------|-----------|-------------|-------------------|------------|
| Ia | DES | 0 | 5 | 450 | real space |
| Ib | DES | 3 | 5 | 630 | real space |
| II | LSST | 0 | 10 | 2200 | Fourier |

Table 14.1: Number of tomographic bins, total number of data points and type of data vector for the different setups used to test the performance of precision matrix expansion.

In order to test the performance of PME we set up mock experiments where we assume the true covariance matrix of each survey to be the analytic halo-model covariance described in Krause & Eifler (2016). This model divides the covariance into three contributions: a noise-only part that consists of shape- and shot-noise contributions, \mathbf{C}^{nn} , a contribution from the cosmic variance of the signal, $\mathbf{C}^{ss,\text{halo}}$, and a mixed term including noise and signal contributions, \mathbf{C}^{sn} . For shear-shear only covariances we set

$$\mathbf{A} = \mathbf{C}^{nn} + \mathbf{C}^{sn} \quad (14.13)$$

and

$$\mathbf{B} = \mathbf{C}^{ss,\text{halo}}. \quad (14.14)$$

The shape-noise contributions to the covariance can be modeled reliably since the ellipticity dispersion can be measured from the data itself and since the mixed term \mathbf{C}^{sn} involves only the modeling of two-point statistics of the shear field. The \mathbf{B} term comprises the more complex 4-point statistics of the shear field, which can be estimated from simulations by turning off shape-noise. This is more complicated for galaxy clustering where shot-noise is included in the covariance matrix (cf. 14.4.1).

In order to simulate a situation where our covariance model $\mathbf{M} = \mathbf{A} + \mathbf{B}_m$ deviates from the true covariance we degrade it as

$$\mathbf{B}_m = \alpha \mathbf{C}^{ss,\text{Gauss}} + \beta (\mathbf{C}^{ss,\text{halo}} - \mathbf{C}^{ss,\text{Gauss}}) \quad (14.15)$$

where $\mathbf{C}^{ss,\text{Gauss}}$ contains only the parts of the cosmic variance that are also present in a Gaussian covariance model. Hence, we allow the Gaussian and non-Gaussian cosmic variance parts to be over- or underestimated by a constant multiplicative factor. If not stated differently in this section we will use $\alpha = 1.0$ and $\beta = 0.5$. In appendix 14.C we explore a wider range of re-scalings and also consider more complex deformations of our fiducial covariance to show that the PME remains robust under more complicated deviations of \mathbf{M} from the true covariance matrix. All simulated likelihood analyses in this paper are computed using the `CosmoLike` cosmology package (Krause & Eifler, 2016; Eifler et al., 2014).

14.4.1 Performance for DES weak lensing data vector

We now carry out mock likelihood analyses for DES and LSST weak lensing data vectors, varying the parameters Ω_m , σ_8 , w_0 and w_a . Our fiducial values for these parameters are

$$(\Omega_m, \sigma_8, w_0, w_a) = (0.3156, 0.831, -1, 0). \quad (14.16)$$

We start by drawing random Gaussian realizations of our fiducial data vectors according to a covariance given by the halo model. For each realization we also draw new Wishart realizations $\hat{\mathbf{B}}$ and $\hat{\mathbf{C}}$ of cosmic variance and total covariance to compute the PME estimate $\hat{\Psi}_{2nd}$ and the standard estimator $\hat{\Psi}$. In practice, this is done by drawing additional realizations $\hat{\boldsymbol{\xi}}_i$, $i = 1 \dots N_s$, of our fiducial data vector from a multivariate Gaussian distribution whose covariance is \mathbf{B} respectively \mathbf{C} . These realizations represent measurements from N-body simulations and inserting them into Eq. 14.3 generates the desired Wishart realizations $\hat{\mathbf{B}}$ and $\hat{\mathbf{C}}$ of the two matrices.

Using `CosmoLike` we then run likelihood chains to infer a posterior distribution for our parameters using $\hat{\Psi}_{2nd}$, $\hat{\Psi}$ and the true precision matrix \mathbf{C}^{-1} . When computing the likelihood from \mathbf{C}^{-1} and $\hat{\Psi}_{2nd}$ we simply use standard ansatz given in Eq. 14.1. When deriving contours from the Wishart realization $\hat{\mathbf{C}}$ we furthermore compute the parameter likelihood as

$$p(\boldsymbol{\pi}|\hat{\boldsymbol{\xi}}) \sim \left[1 + \frac{(\hat{\boldsymbol{\xi}} - \hat{\boldsymbol{\xi}}[\boldsymbol{\pi}])^T \hat{\mathbf{C}}^{-1} (\hat{\boldsymbol{\xi}} - \hat{\boldsymbol{\xi}}[\boldsymbol{\pi}])}{N_s - 1} \right]^{-N_s/2} \quad (14.17)$$

which SH16a have shown to be a more accurate than using the Kaufman-Hartlap correction and the standard Gaussian likelihood. We however found only small differences to using the standard likelihood ansatz, which is due to the fact that in all cases considered in this paper $N_s - N_d \gg N_p$.

In Fig. 14.2 and 14.3 we show see the resulting 1σ and 2σ contours in the Ω_m - σ_8 plane (after marginalizing over the other parameters) for 3 different random draws of data vector and Wishart matrices. For each realization of the DES data vector we assumed that $N_s = 200$ simulations are available to estimate the PME and $N_s = N_d + 200 = 650$ simulations for the standard estimator. For each realization of the LSST data vector we assumed $N_s = 400$ simulations for the PME and $N_s = N_d + 400 = 2600$ simulations for the standard estimator.

Even though in each case we assumed many more simulations for the standard estimator than for the PME, the PME is significantly better in reconstructing the contours from the true precision matrix. In particular we find that deviations from the true contours are much smaller than the corresponding 1σ and 2σ uncertainties of the parameters.

Next we generalize the findings in Figs. 14.2 and 14.3. We generate 1000 Wishart realizations of the matrices $\hat{\mathbf{C}}$ and $\hat{\mathbf{B}}$ for different assumptions on the number of available N-body simulations N_s . For each of the 1000 sets of matrices we also generate 10 realizations $\hat{\boldsymbol{\xi}}$ of our fiducial data vector (i.e. overall 10 000 different realizations $\hat{\boldsymbol{\xi}}$). Hence for each type

of precision matrix estimate we perform overall 10 000 likelihood analyses. In each analysis we determine the best fit parameters $\hat{\boldsymbol{\pi}}_{\text{ML}}$ and check whether our fiducial cosmology is outside the 68.3% confidence contour around these parameters. In order to make this computationally feasible, we are now linearly approximating the calculations of `CosmoLike` around our fiducial cosmology $\boldsymbol{\pi}_0$, i.e. we use

$$\boldsymbol{\xi}_{\text{simple}}[\boldsymbol{\pi}] = \boldsymbol{\xi}_{\text{exact}}[\boldsymbol{\pi}_0] + \sum_{i=1}^{N_p} (\pi_i - \pi_{0,i}) \frac{\partial \boldsymbol{\xi}_{\text{exact}}}{\partial \pi_i}[\boldsymbol{\pi}_0]. \quad (14.18)$$

This allows us to analytically determine the maximum likelihood parameters and the 68.3% confidence contours in each likelihood analysis. It is also the situation where a Fisher-matrix formalism and hence the derivations of DS13 hold exactly.

We define $F_{>1\sigma}$ as the fraction of times that our fiducial cosmology is outside of the 68.3% confidence contour around the best fit parameters and we use it as a metric for comparing the different precision matrix estimators. In Fig. 14.4 we show this fraction for all different types of precision matrices introduced before. The solid, dashed, and dotted lines show the fractions achieved when using the noise-less matrices \mathbf{C}^{-1} , \mathbf{M}^{-1} and $\boldsymbol{\Psi}_{\text{2nd}}$. Especially, the noise-less matrix $\boldsymbol{\Psi}_{\text{2nd}}$ would be the PME-estimator in the limit of infinitely many simulations and \mathbf{C}^{-1} would be the standard estimator in the same limit. The red and blue dots show the fraction achieved when using the noisy precision matrix estimates $\hat{\boldsymbol{\Psi}}$ and $\hat{\boldsymbol{\Psi}}_{\text{2nd}}$.

As expected, $F_{>1\sigma}$ is very close to 32% when using the true covariance \mathbf{C} in the likelihood analyses. For the deformed halo model covariance \mathbf{M} we assumed the two cases $\alpha = 0.7$, $\beta = 0.5$ (left panel) and $\alpha = 1.0$ and $\beta = 0.5$ (right panel). For $\alpha = 0.7$ and $\beta = 0.5$ our fiducial cosmology is regarded as outside the 68.3% contour in more than 40% of the cases. For both choices of \mathbf{M} the noise-free PME significantly corrects that fraction towards the optimal value of $\sim 32\%$. Especially promising is that the PME estimate performs very similar to the noise-free PME. If 200 simulations are available to estimate the PME, it essentially converges to its best possible performance. And even if only 100 simulations are available to estimate the PME, its value of $F_{>1\sigma}$ comes closer to 32% than when using \mathbf{M} to derive the contours.

When inferring the likelihood from the standard precision matrix estimator $F_{>1\sigma}$ is greater than 50% even if we allow $N_s = N_d + 800$ simulations for the covariance estimation, which corresponds to 1250 simulations. This is due to the additional variance of $\hat{\boldsymbol{\pi}}_{\text{ML}}$ caused by the noise of the precision matrix (cf. Eq. 14.5). Using the results of DS13 we can derive predictions for this effect (cf. appendix 14.A). As can be seen from the red dashed lines in Fig. 14.4 these predictions agree well with what we find in our simulated likelihood analyses. Extrapolating the results of DS13 to higher values of N_s we can also estimate, how many simulations would be required for the standard precision matrix estimator in order to achieve the same value of $F_{>1\sigma}$ as the second order PME. For the left panel of Fig. 14.4 we find that it would take ~ 8000 simulations for the standard estimator to get as close to $F_{>1\sigma} = 32\%$ as the PME with only 200 simulations. This statement however depends on the model covariance \mathbf{M} since it determines how well the PME has converged

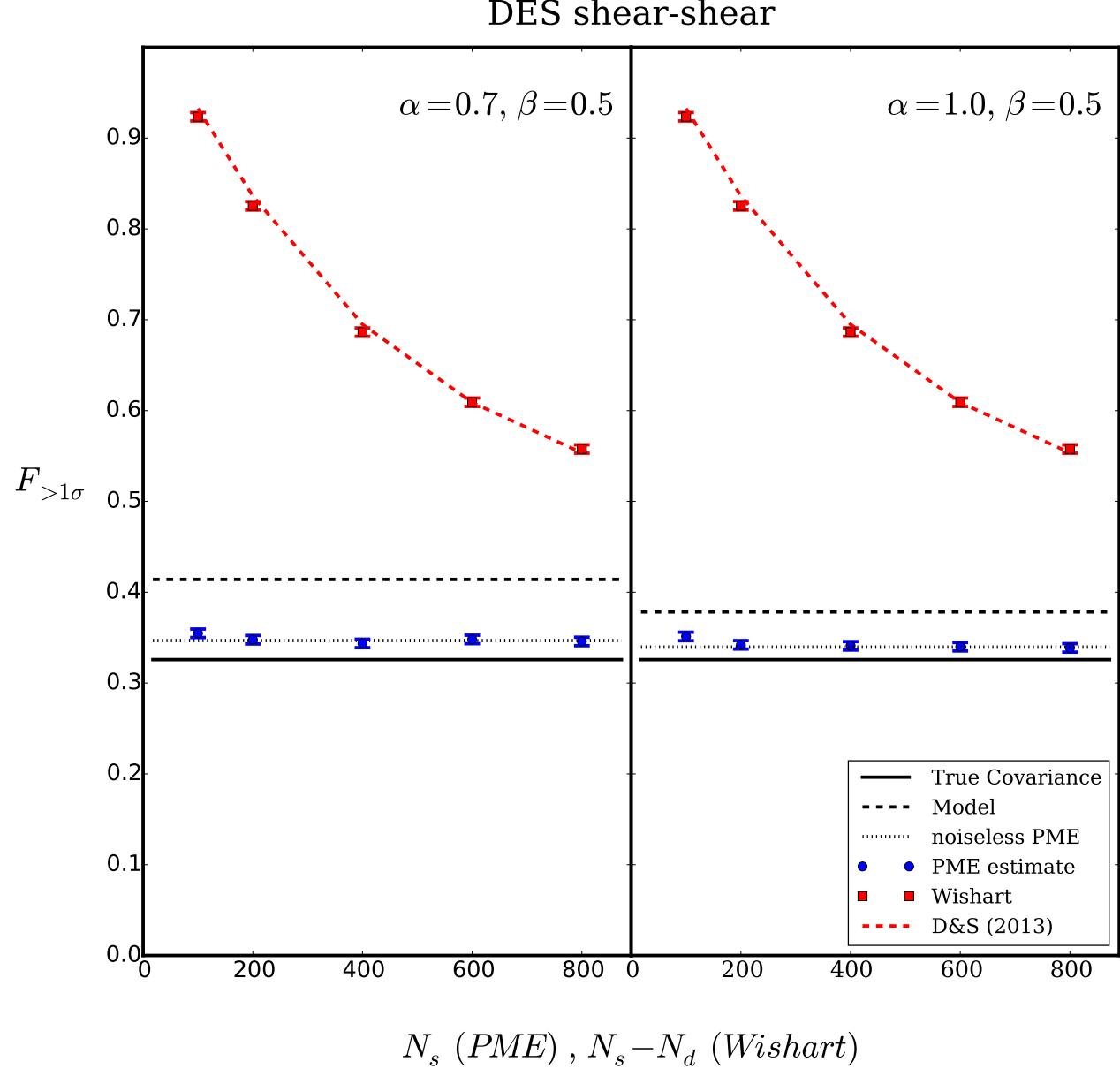


Figure 14.4: The figure compares $F_{>1\sigma}$, the number of times that our fiducial cosmology was considered outside the 68.3% confidence contour in our simulated likelihood analyses when using different precision matrix estimates for computing the posterior parameter likelihood. In order to carry a sufficient number of mock analyses, we simplified our modeling of the data vector by linearly approximating the full computation around our fiducial cosmology. For the DES-like weak lensing data vector we varied the four parameters $(\Omega_m, \sigma_8, w_0, w_a)$.

after its second order.

An \mathbf{M} -independent way of comparing standard estimator and PME estimator is to see how many simulations it takes each to have $F_{>1\sigma}$ within 1% of their best possible performance. It would take the standard estimator ~ 24000 simulations to be within 1% of $F_{>1\sigma} = 32\%$. The PME estimator is well within 1% of its best possible performance for only 200 simulations.

Note that with the results of DS13 one can in principle correct a likelihood analysis for the additional variance caused by the standard precision matrix estimator. This would result in a decreased constraining power of the analysis and it would hence be the main benefit of the PME to prevent this loss.

Larger covariance matrices: LSST weak lensing data vector

We repeat the above analysis for the LSST-like weak lensing data vector. Fig. 14.5 shows the fractions $F_{>1\sigma}$ obtained from PME and standard precision matrix estimator. The PME estimator now requires ~ 2400 simulations to be less than 1% away from its best possible performance. As before, this statement does not include any additional biases between PME and true precision matrix that might arise from the biased model matrix \mathbf{M} used to carry out the matrix expansion. The standard precision matrix estimator would need $N_s > 115\,000$ simulations to be less than 1% away from its best possible performance.

Defining \mathbf{A} and \mathbf{B} for multi-probe covariances

We now repeat the analysis of Fig. 14.4 for a DES-like multi-probe data vector. This vector includes contributions from galaxy clustering and galaxy-galaxy lensing, which introduces shot-noise terms to the covariance. These shot-noise contributions are in principle well understood theoretically and include, similar to the cosmic shear case, at most two-point statistics of the cosmic density field. Hence, one could absorb them into the matrix \mathbf{A} (cf. Eq. 14.6) and use N-body simulations only for the remaining part of the covariance - i.e. to define \mathbf{B} as only the cosmic variance. This is however difficult since most N-body simulations provide only simulated galaxy catalogs that are affected by shot noise themselves, which makes it impossible to independently estimate the cosmic variance. If however all shot-noise contributions are included in \mathbf{B} when defining and estimating the PME, then the estimator $\hat{\Psi}_{2\text{nd}}$ will have a higher variance in many of its elements. Hence, the additional scatter of best fit parameters due to a noisy precision matrix might not be negligible anymore.

In Fig. 14.6 we compare the fractions $F_{>1\sigma}$ obtained from different estimates of the precision matrix in our simulated likelihood analyses - this time for the DES multi-probe data vector. In each likelihood analysis we now vary 7 parameters, since for each lens bin we include a galaxy bias parameter in our model. The fiducial bias values are

$$(b_1, b_2, b_3) = (1.35, 1.50, 1.65) \quad (14.19)$$

in order of increasing redshift. Fig. 14.6 shows the results obtained for each of the mentioned options of defining \mathbf{B} . It is clear that the noisy PME approaches its best possible

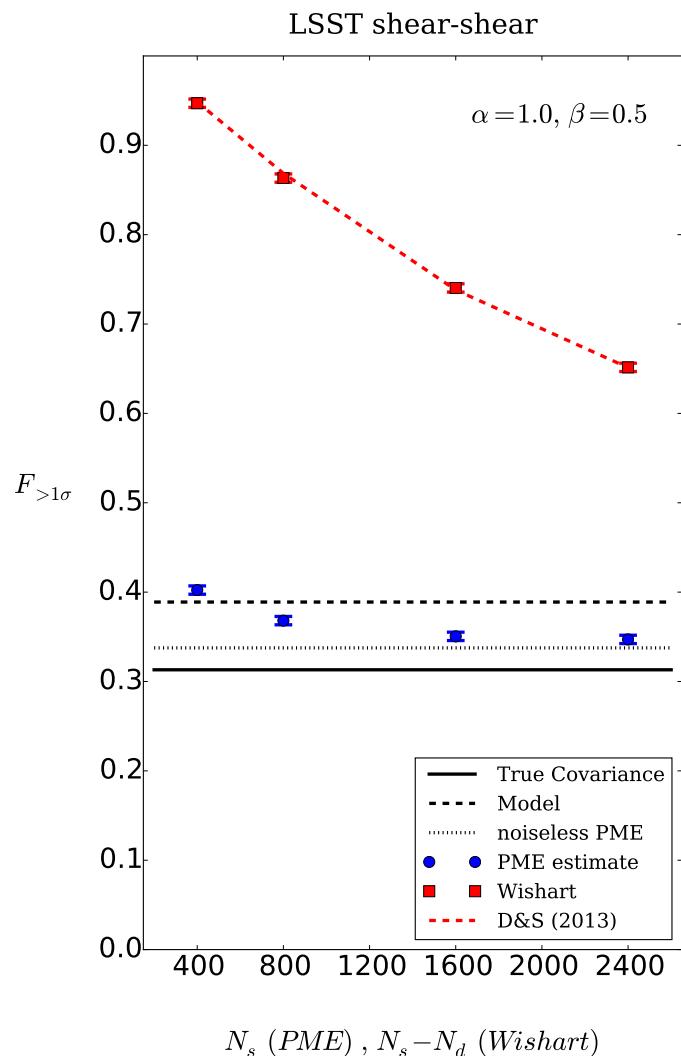


Figure 14.5: Same as Fig. 14.4 but for the LSST-like weak lensing data vector.

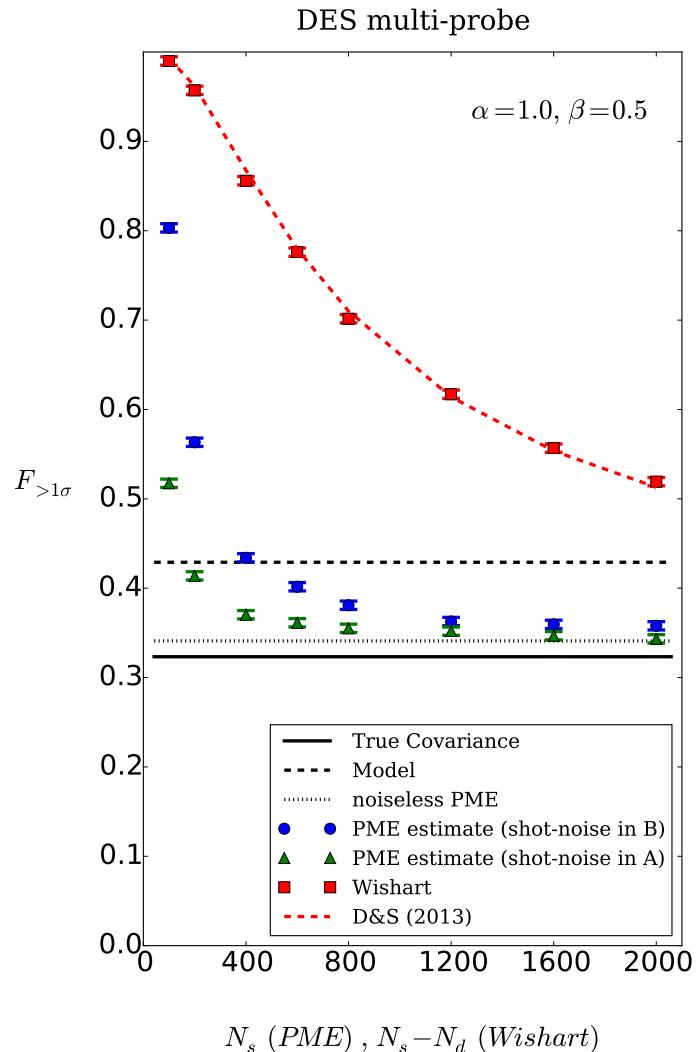


Figure 14.6: Same as Fig. 14.4 but for the DES multi-probe data vector. For this case the 7 parameters ($\Omega_m, \sigma_8, w_0, w_a, b_1, b_2, b_3$) were varied in each likelihood analysis. Green points assume that cosmic variance can be estimated from simulations without shot-noise. This would significantly improve the performance of PME for low numbers of available simulations.

performance already for a smaller number of simulations if the cosmic variance can be estimated directly. In practice this would however require density maps in thin redshift slices for each simulation in order to measure the correlation functions of the projected density fields without shot-noise.

Assuming one can directly measure the cosmic variance from simulations we again want to asses how many simulations are required for the standard precision matrix estimator and the PME estimator to be within 1% of their best possible performance. Extrapolating the results of DS13 we find that it would take the standard estimator $\sim 44\,000$ simulations to be within within 1% of $F_{>1\sigma} = 32\%$. The PME estimator is within 1% of its best possible performance for 1600 simulations. For $N_s = 2000$ the performance of the PME becomes almost solely restricted by the deviation between **M** and **C** in our mock experiment. However, below $N_s = 1600$ there seems to be significant additional scatter of the best fitting parameters due to the noise of the PME estimate. We demonstrate this in Fig. 14.7 for $N_s = 400$. Regardless of how **B** is defined, we can nevertheless conclude that also for multi-probe covariances the PME poses a vast improvement over the standard precision matrix estimator.

14.5 Conclusions

It was the starting point of our analysis to find a method for using a priori knowledge about the covariance matrix when estimating the precision matrix from simulations. This requires finding an equivalent of the Kaufman-Hartlap correction when only parts of the covariance are estimated. Using the results of Letac & Massam (2004) we partly solved this task by calculating an expansion of the precision matrix and showing how the leading terms of this expansion can be estimated from simulations. Our method enables the use of preexisting knowledge on the covariance structure to improve the convergence of the PME and to reduce the noise in its estimation. It also has the advantage that the relative uncertainties of the elements of the PME estimate scale with the number of available simulations N_s as $\sim 1/\sqrt{N_s - 1}$, which is typically much smaller than the uncertainties of the standard precision matrix estimator. The latter also depend on the number of data points N_d and scale as $\sim 1/\sqrt{N_s - N_d - 4}$.

We demonstrated that the PME converges even for drastic deviations between the model covariance and the N-body covariance and we also showed that it provides a much less noisy estimate of the parameter likelihood compared to estimating the precision matrix in the standard way. For a DES weak lensing data vector $N_s \gtrsim 8000$ simulations would be required for the standard estimator to reconstruct the likelihood similarly well as the PME with only $N_s = 200$ - even if the model covariance heavily underestimates Gaussian and non-Gaussian covariance parts. If we assume more realistic deviations between model and N-body covariance, up to 24 000 simulations would be needed for the standard estimator to reconstruct the 1σ quantile of the parameter distribution at the same precision as the PME with only 200 simulations. For an LSST-like weak lensing data vector with $N_d = 2200$ we found that up to 115 000 simulations would be required for the standard estimator to

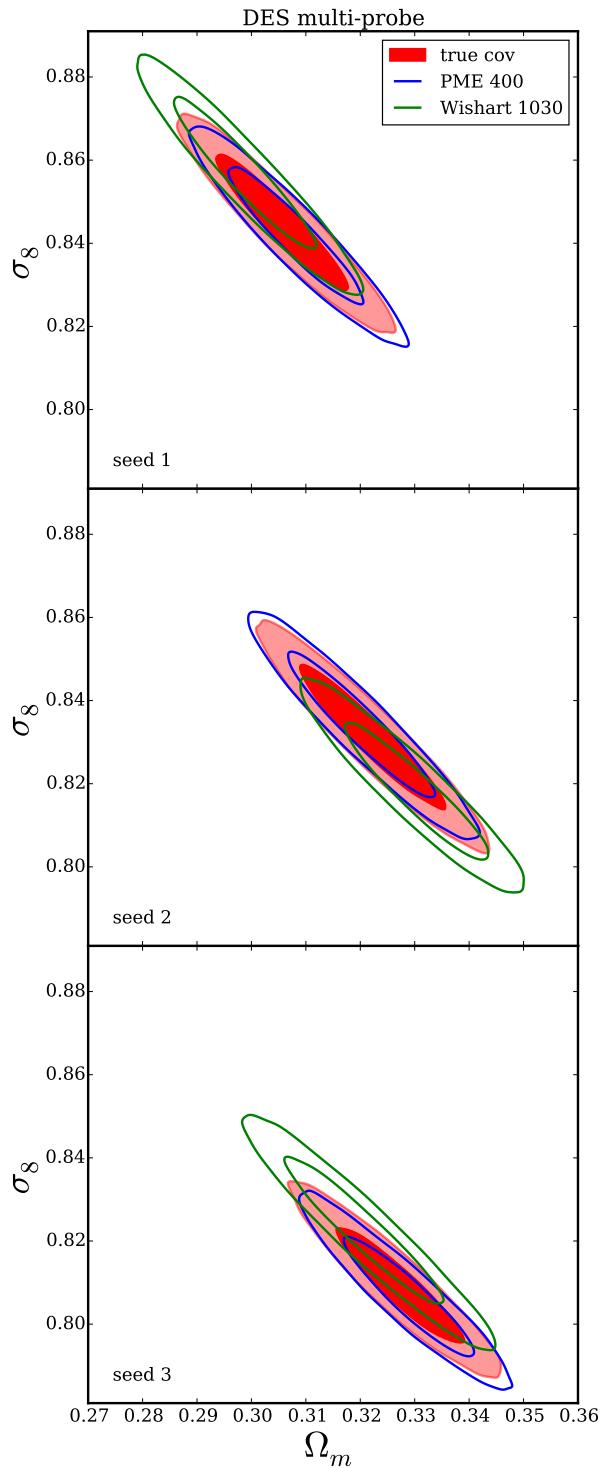


Figure 14.7: Same as Fig. 14.2 but for the multi-probe data vector. $N_s = 400$ simulations where assumed for the estimation of the PME while $N_s = N_d + 400 = 1030$ simulations where assumed for the standard estimator. Even with fewer simulations the PME is much better in reconstructing the contours from the true precision matrix. However, below $N_s = 1600$ a significant offset of the contours persists.

reconstruct the 1σ quantile as well as the PME with only 2400 simulations. It should however be stressed that these statements depend on the quality of the model covariance \mathbf{M} that was used to compute the PME.

Additional complications arise when galaxy clustering correlation functions are included in the data vector. A performance similar to the weak lensing case can still be achieved if one manages to estimate the cosmic variance of the correlation functions directly, i.e. without shot-noise. For this case, we find that a DES-like multi-probe data vector requires up to 44 000 simulations for the standard precision matrix estimator to reconstruct the 1σ quantile of the parameter distribution as well as the PME with 1600 simulations.

One aspect that should be addressed in future work, is to find a priori criteria for the convergence of the PME. In appendix 14.C we demonstrate that it converges for very strong deformations of the halo-model covariance, but one can not be certain whether and how fast it will converge for all possible data vectors and covariance models. As we show in appendix 14.C, situations where the PME does not converge can at least be identified a posteriori by a comparison of the first order and second order expansion. A strong oscillation of likelihood contours derived from the first and second order PME indicate a significant deviation of model and N-body covariance. This way, the PME provides a clear criterion for testing covariance models with simulations – even when the number of available simulations is small.

Acknowledgments

This work was supported by SFB-Transregio 33 ‘The Dark Universe’ by the Deutsche Forschungsgemeinschaft (DFG). We also acknowledge the support by the DFG Cluster of Excellence ”Origin and Structure of the Universe”. The simulations have been carried out on the computing facilities of the Computational Center for Particle and Astrophysics (C2PAP). This research is partially supported by NASA ROSES ATP 16-ATP16-0084 grant. Part of the research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

This paper has gone through internal review by the DES collaboration. We want to thank Eric Baxter, Gary Bernstein, Scott Dodelson, Franz Elsner, Juan Garcia-Bellido, John Peacock, Stella Seitz and the DES multi-probe pipelines group for very helpful comments and discussions on this project.

Funding for the DES Projects has been provided by the U.S. Department of Energy, the U.S. National Science Foundation, the Ministry of Science and Education of Spain, the Science and Technology Facilities Council of the United Kingdom, the Higher Education Funding Council for England, the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign, the Kavli Institute of Cosmological Physics at the University of Chicago, the Center for Cosmology and Astro-Particle Physics at the Ohio State University, the Mitchell Institute for Fundamental Physics and Astronomy at Texas A&M University, Financiadora de Estudos e Projetos, Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro, Conselho Nacional de Desenvolvimento Científico e Tecnológico and the Ministério da Ciência, Tecnologia e Inovação, the Deutsche Forschungsgemeinschaft and the Collaborating Institutions in the Dark Energy Survey. The DES data management system is supported by the National Science Foundation under Grant Number AST-1138766.

The Collaborating Institutions are Argonne National Laboratory, the University of California at Santa Cruz, the University of Cambridge, Centro de Investigaciones Enérgéticas, Medioambientales y Tecnológicas-Madrid, the University of Chicago, University College London, the DES-Brazil Consortium, the University of Edinburgh, the Eidgenössische Technische Hochschule (ETH) Zürich, Fermi National Accelerator Laboratory, the University of Illinois at Urbana-Champaign, the Institut de Ciències de l’Espai (IEEC/CSIC), the Institut de Física d’Altes Energies, Lawrence Berkeley National Laboratory, the Ludwig-Maximilians Universität München and the associated Excellence Cluster Universe, the University of Michigan, the National Optical Astronomy Observatory, the University of Nottingham, The Ohio State University, the University of Pennsylvania, the University of Portsmouth, SLAC National Accelerator Laboratory, Stanford University, the University of Sussex, and Texas A&M University.

The DES participants from Spanish institutions are partially supported by MINECO under grants AYA2012-39559, ESP2013-48274, FPA2013-47986, and Centro de Excelencia Severo Ochoa SEV-2012-0234. Research leading to these results has received funding from the European Research Council under the European Unions Seventh Framework Pro-

gramme (FP7/2007-2013) including ERC grant agreements 240672, 291329, and 306478.

14.A Influence of noisy covariance estimates on the scatter of best fitting cosmological parameters

Using a noisy precision matrix estimate to determine the best fitting cosmological parameters

$$\hat{\boldsymbol{\pi}}_{\text{ML}} = \min_{\boldsymbol{\pi}} \left\{ \left(\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}[\boldsymbol{\pi}] \right)^T \boldsymbol{\Psi} \left(\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}[\boldsymbol{\pi}] \right) \right\} . \quad (14.20)$$

leads to an additional scatter in these parameters. Especially, this additional noise is not accounted for by the width of contours generated from the precision matrix estimate. This effect has e.g. been described by Dodelson & Schneider (2013) who also derived a prediction for the additional noise assuming a Gaussian parameter likelihood. They find that the actual parameter covariance when using an inverse-Wishart realization of the precision matrix is given by

$$\mathbf{C}_{\hat{\boldsymbol{\pi}}_{\text{ML}}} = \mathbf{F}^{-1} \left(1 + \frac{(N_d - N_p)(N_s - N_d - 2)}{(N_s - N_d - 1)(N_s - N_d - 4)} \right) , \quad (14.21)$$

where N_p is the number of considered parameters and \mathbf{F} is the Fisher matrix computed from the true precision matrix. Hence, in the case of a Gaussian parameter likelihood, best fit parameters $\hat{\boldsymbol{\pi}}_{\text{ML}}$ that are computed from a Wishart realization of the covariance have also a Gaussian distribution but with a rescaled parameter covariance.

14.B Unbiased estimator of the square of a Wishart matrix

Let $\hat{\mathbf{C}}$ be distributed according to a Wishart distribution with ν degrees of freedom and expectation value \mathbf{C} . Then

$$\langle \hat{\mathbf{C}}^2 \rangle \neq \mathbf{C}^2 . \quad (14.22)$$

However, using the results of Letac & Massam (2004) it is possible to devise an unbiased estimator of \mathbf{C}^2 . It is given by

$$\widehat{(\mathbf{C}^2)} = \frac{\nu^2 \hat{\mathbf{C}}^2 - \nu \hat{\mathbf{C}} \text{tr} \hat{\mathbf{C}}}{\nu^2 + \nu - 2} , \quad (14.23)$$

where $\text{tr} \hat{\mathbf{C}}$ denotes the trace of $\hat{\mathbf{C}}$. Using this formula, it is straight forward to derive the estimator of the second order PME given in Eq. 14.12.

14.C General properties and convergence of the power series

14.C.1 General properties

In order to derive some general properties of the PME series, let us slightly change the notation of Sec. 14.3. First, let $\mathbf{M}^{1/2}$ be the unique symmetric and positive definite matrix such that

$$\mathbf{M}^{1/2}\mathbf{M}^{1/2} = \mathbf{M} . \quad (14.24)$$

This matrix exists as long as our covariance model \mathbf{M} is positive definite. Let us then re-define

$$\mathbf{X} = \mathbf{M}^{-1/2} (\mathbf{B} - \mathbf{B}_m) \mathbf{M}^{-1/2} \quad (14.25)$$

where $\mathbf{M}^{-1/2}$ is the inverse of $\mathbf{M}^{1/2}$, and \mathbf{B} and \mathbf{B}_m are the same as in Sec. 14.3. The complete covariance can then be written as

$$\mathbf{C} = \mathbf{M}^{1/2} (\mathbb{1} + \mathbf{X}) \mathbf{M}^{1/2} \quad (14.26)$$

and the precision matrix expansion now reads

$$\begin{aligned} \boldsymbol{\Psi} &= \mathbf{M}^{-1/2} \left(\sum_{k=0}^{\infty} (-1)^k \mathbf{X}^k \right) \mathbf{M}^{-1/2} \\ &= \mathbf{M}^{-1/2} (\mathbb{1} - \mathbf{X} + \mathbf{X}^2 + \mathcal{O}[\mathbf{X}^3]) \mathbf{M}^{-1/2} . \end{aligned} \quad (14.27)$$

Since both $\mathbf{M}^{-1/2}$ and \mathbf{X} are symmetric matrices, it is immediately clear that this gives a symmetric approximation of $\boldsymbol{\Psi}$ at each order of the power series. The series converges if and only if all eigenvalues of \mathbf{X} fulfill

$$|\lambda_i| < 1 , \quad i = 1 , \dots , N_d . \quad (14.28)$$

In each eigendimension of \mathbf{X} the series $(\mathbb{1} - \mathbf{X} + \mathbf{X}^2 + \mathcal{O}[\mathbf{X}^3])$ is simply the geometric series. For $|\lambda_i| < 1$ the value of this series is > 0 at each finite order. At second order, the value of this series is > 0 regardless of the values of λ_i . Hence, the second order PME is always positive definite.

14.C.2 Special cases

Rescaling of the covariance

Let us investigate the convergence properties of the power series in Eq. 14.10 in a couple of special cases. We start by assuming that our model for the covariance matrix, \mathbf{M} , under- or overestimates the true covariance matrix by a constant factor α , i.e.

$$\mathbf{M} = \alpha \mathbf{C} . \quad (14.29)$$

In this case we have

$$\begin{aligned}\mathbf{X} &= \mathbf{M}^{-1/2} (\mathbf{C} - \alpha \mathbf{C}) \mathbf{M}^{-1/2} \\ &= \frac{1 - \alpha}{\alpha} \mathbf{C}^{-1/2} \mathbf{C} \mathbf{C}^{-1/2} \\ &= \frac{1 - \alpha}{\alpha} \mathbf{1}.\end{aligned}\quad (14.30)$$

Hence, all eigenvalues of \mathbf{X} are given by $\lambda = \frac{1-\alpha}{\alpha}$. This has absolute value smaller than 1 for all $\alpha > 0.5$. This especially means that the series used to define the PME converges even if the model covariance overestimates the true covariance by an arbitrarily high overall factor. Since we cut Eq. 14.10 after the second order we must however look at how well the series is converged after that order. The relative error on each element of the precision matrix is given by

$$\frac{\Psi_{ij} - \Psi_{2nd,ij}}{\Psi_{ij}} = \lambda^3 = \frac{(1 - \alpha)^3}{\alpha^3}.\quad (14.31)$$

This is $< 10\%$ for $\alpha \in [0.69, 1.86]$ and $< 1\%$ for $\alpha \in [0.83, 1.27]$.

Partial rescaling of the covariance

Now let us assume that \mathbf{C} falls into two contributions \mathbf{A} and \mathbf{B} and that only \mathbf{B} is mischaracterized by a constant factor in our model,

$$\mathbf{M} = \mathbf{A} + \alpha \mathbf{B}.\quad (14.32)$$

Let us furthermore assume that \mathbf{B} has a dominant eigenvalue λ and that \mathbf{v} is an eigenvector to it. If

$$|\lambda \mathbf{v} - \mathbf{C} \mathbf{v}| \ll |\lambda \mathbf{v}| \quad (14.33)$$

then the matrix \mathbf{C} and – for values of α that are not too small – also the matrix \mathbf{M} will have an eigendimension close the that of \mathbf{B} with eigenvalues $\lambda_C \approx \lambda$ and $\lambda_M \approx \alpha \lambda$. As a consequence, the matrix \mathbf{X} will have an eigendimension with eigenvalue close to $\lambda_X \approx \frac{1-\alpha}{\alpha}$ which allows the same conclusion in 14.C.2.

In section 14.4 we considered a deformation of the halo model covariance of the form $\mathbf{M} = \mathbf{A} + \mathbf{B}_m$ with

$$\mathbf{B}_m = \alpha \mathbf{C}^{ss, \text{Gauss}} + \beta (\mathbf{C}^{ss, \text{halo}} - \mathbf{C}^{ss, \text{Gauss}}).\quad (14.34)$$

This is similar to the situation described above. To illustrate how the rescaling factors α and β impact the convergence of the PME we can e.g. compare the Fisher contours derived from \mathbf{C}^{-1} , \mathbf{M}^{-1} , Ψ^{1st} and Ψ^{2nd} . In Fig. 14.8 we show the 1σ and 2σ Fisher-contours for the parameter pair Ω_m - σ_8 derived for the DES multi-probe data vector using different values of α and β . The figure shows that the PME manages to correct the bias between contours derived from \mathbf{C}^{-1} and contours derived from \mathbf{M}^{-1} even for rather drastic choices of the rescaling factors. Especially for $\alpha, \beta > 1.0$ the convergence is very robust. As predicted

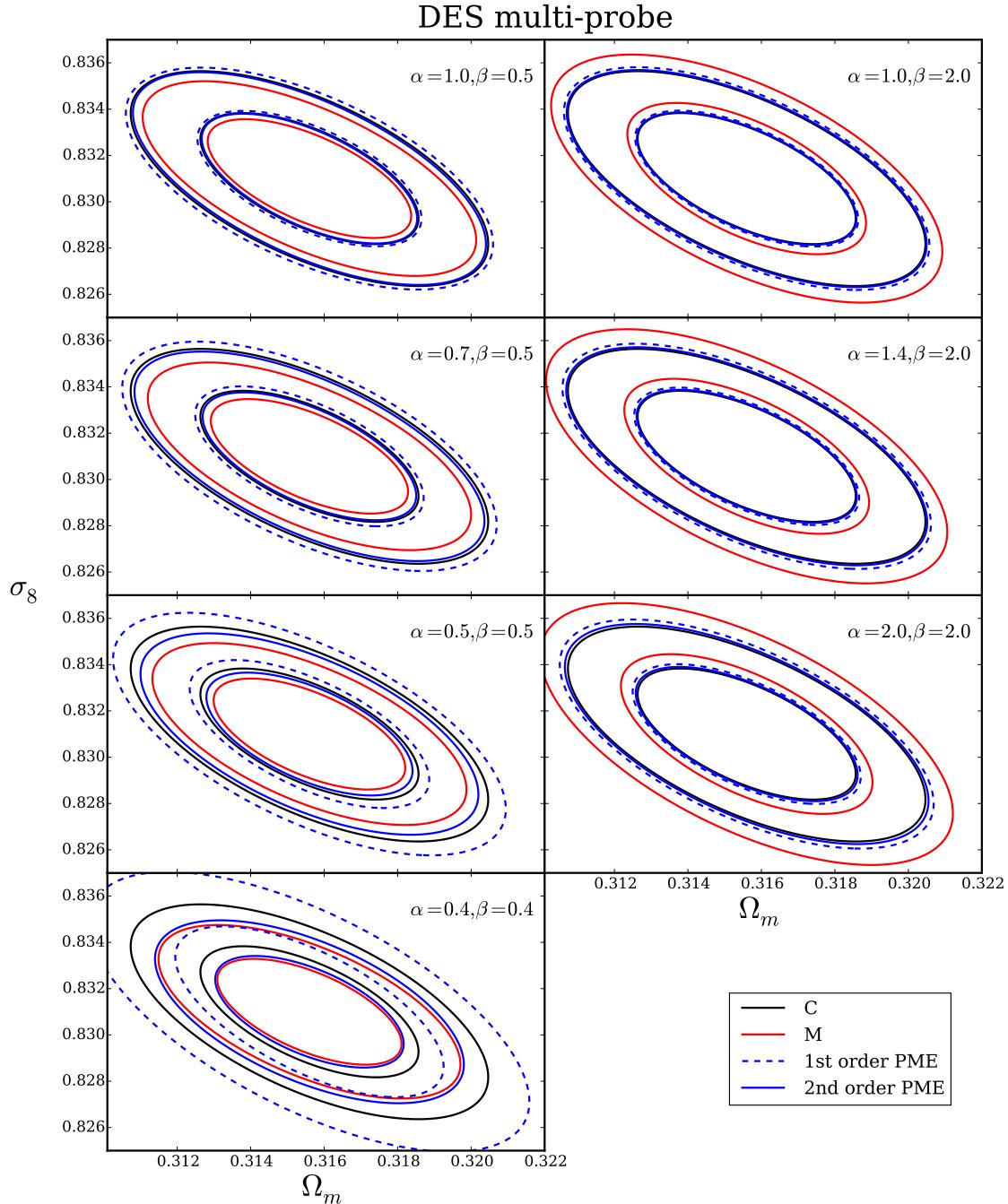


Figure 14.8: We show the 1σ and 2σ Fisher contours in the Ω_m - σ_8 plane around our fiducial cosmology using the DES multi-probe data vector and keeping all other cosmological parameters fixed. For the black contours the Fisher matrix was derived from the fiducial covariance matrix \mathbf{C} of our experiment – the halo-model covariance. For the red contours we rescaled the Gaussian and non-Gaussian parts of the cosmic variance in \mathbf{C} by constant factors α and β to create our model covariance matrix \mathbf{M} (cf. Eq. 14.15). The blue contours show the constraints derived from the 1st order PME (dashed lines) and 2nd order PME (solid lines) of \mathbf{C} around \mathbf{M} . The PME manages to significantly correct the miss estimation of the Fisher matrix by the model precision matrix for most values of the rescaling factors. Only for $\alpha, \beta < 0.5$ the convergence of the PME seems to break down and a strong oscillation between 1st order and 2nd order correction occurs. We discuss this behavior in detail in appendix 14.C where we also study examples of more complicated deviations between \mathbf{M} and \mathbf{C} .

by our considerations above, it however breaks down for $\alpha, \beta < 0.5$ where one can see strong oscillations between $\Psi_{1\text{st}}$ and $\Psi_{2\text{nd}}$. The convergence of the contours in Fig. 14.8 is very similar when other parameter combinations are considered or when the contours are derived for the other data vectors considered in this paper.

Log-normal motivated approximation to the halo-model covariance

Motivated by the work of Hilbert et al. (2011) on approximating the shear-shear covariance matrix with a log-normal approach (cf. their equation 26) we approximate the non-Gaussian parts of the covariance of shear correlation functions as

$$\langle \Delta\xi_{\pm}^A(\theta_i) \Delta\xi_{\pm}^B(\theta_j) \rangle_{\text{non Gauss.}} = \xi_{\pm}^A(\theta_i) \xi_{\pm}^B(\theta_j) R_{AB} \quad (14.35)$$

where θ_i labels the different angular bins, A and B label the different auto- and cross-correlation functions and R_{AB} is just a constant factor (depending only on the pair A, B and not on whether ξ_+ or ξ_- are involved). We fix the values of R_{AB} by demanding that our approximation coincides with the halo-model for $\langle \Delta\xi_{+}^A(\theta) \Delta\xi_{+}^B(\theta) \rangle_{\text{non Gauss.}}$ where θ is a certain angular scale which we chose to be either our smallest angular bin ($\theta \approx 3'$) or a slightly larger scale ($\theta \approx 20'$). Note that this is a very crude approximation – even to the log-normal model by Hilbert et al. (2011) since they have not even considered cross-correlations between redshift bins.

We nevertheless use the above matrix as our model covariance \mathbf{M} for the DES shear-shear data vector and compare it to the halo-model covariance \mathbf{C} and the PME. All eigenvalues of matrix \mathbf{X} have in that case $|\lambda_i| < 1$. The three most dominant eigenvalues are

$$\begin{aligned} \lambda_1 &= 0.776 \\ \lambda_2 &= -0.675 \\ \lambda_3 &= 0.197 \end{aligned} \quad (14.36)$$

in the case where we match the amplitudes of \mathbf{M} and \mathbf{C} at $\theta \approx 20'$ and

$$\begin{aligned} \lambda_1 &= 0.966 \\ \lambda_2 &= -0.442 \\ \lambda_3 &= 0.203 \end{aligned} \quad (14.37)$$

when we match the amplitudes at $\theta \approx 3'$. In both of these cases the PME in principle converges. However, in the second case at least one eigenvalue comes dangerously close to 1. In Fig. 14.9 we show that in terms of the Fisher contours in the Ω_m - σ_8 -plane the PME nevertheless converges and significantly corrects for the deviations between halo-model and log-normal motivated covariance. We have also checked other parameter combinations and find similar results. The reason that a matching at larger scales gives smaller eigenvalues (i.e. better agreement between halo-model and log-normal motivated covariance) is probably that the scaling of Eq. 14.35 fails at small scales.

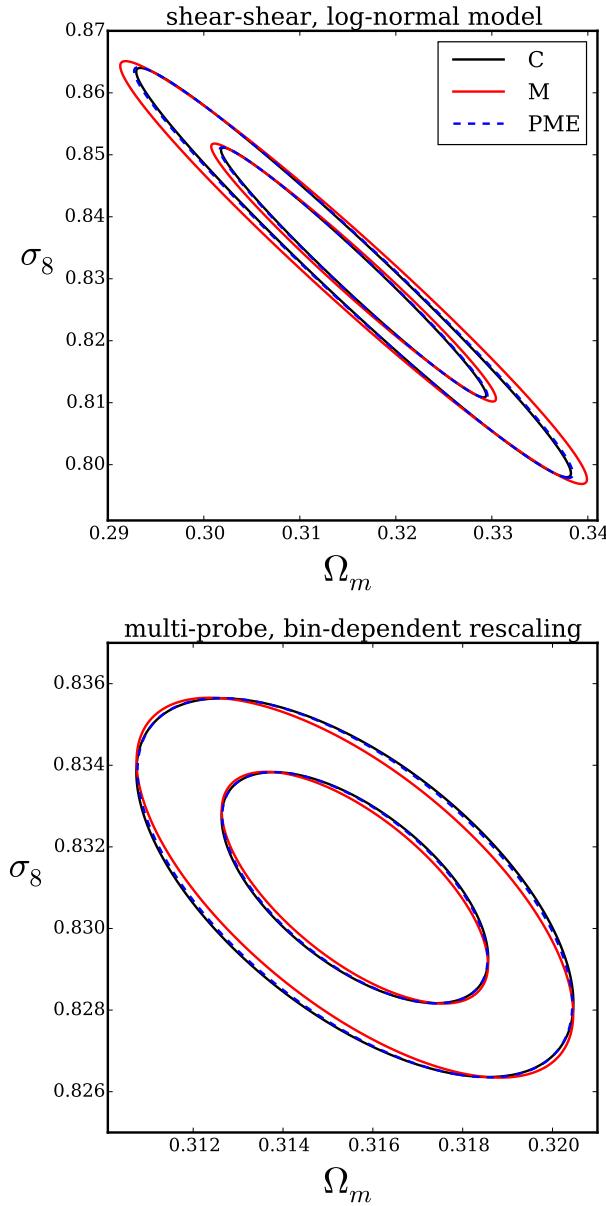


Figure 14.9: Top: 1σ and 2σ Fisher contours in the Ω_m - σ_8 plane for the DES weak lensing data vector. The black contours are derived from our fiducial halo-model covariance \mathbf{C} . For the red contours we used a model covariance \mathbf{M} that was motivated from the general structure of the log-normal covariance model for shear-shear correlation functions by Hilbert et al. (2011, see main text). The PME (blue dashed contours) still manages to correct for the deviation between the two models. It should however be noted that in this case one eigenvalue of the deviation matrix \mathbf{X} comes dangerously close to one ($\lambda_{\max} = 0.966$). As we discuss in the main text, this situation stabilizes if we match the amplitudes of the halo-model and the log-normal motivated covariance at intermediate angular scales ($\theta \sim 20'$) instead of the smallest scale of our data vector ($\theta \sim 3'$). Bottom: we applied a scale dependent rescaling of the halo-model covariance for the multi-probe data vector motivated by findings of Friedrich et al. (2016). The PME converges also in this case.

Scale dependent rescaling of the cosmic variance of the multi-probe data vector

Another alternative way to deform the halo-model covariance is to apply different rescaling factors α and β for the Gaussian and non-Gaussian cosmic variance parts for different angular scales (cf. Eq. 14.15). If e.g. the finite area of a survey is not correctly accounted for in a covariance model, the results of Friedrich et al. (2016) indicate that this leads to a scale dependent miss-characterization of the Gaussian cosmic variance and to an almost scale independent over- or underestimation of the non-Gaussian parts. Covariance parts involving shape- or shot-noise on the other hand are less sensitive to the survey area (only to the product of area and galaxy density which is the total number of galaxies).

Motivated by this we replace Eq. 14.15 by

$$B_{m,ij} = \alpha_{ij} C_{ij}^{ss,\text{Gauss}} + \beta \left(C_{ij}^{ss,\text{halo}} - C_{ij}^{ss,\text{Gauss}} \right) \quad (14.38)$$

where we choose $\beta = 0.5$ and $\alpha_{ij} = \sqrt{a_i a_j}$ setting a_i to 1.0 at the smallest scales and to 0.5 at the largest scales of the data vector and linearly interpolating for intermediate bins (interpolating in terms of the bin-index).

The most dominant eigenvalues of the matrix \mathbf{X} for this choice of the matrix \mathbf{M} are

$$\begin{aligned} \lambda_1 &= 0.709 \\ \lambda_2 &= -0.440 \\ \lambda_3 &= 0.242, \end{aligned} \quad (14.39)$$

i.e. the PME converges. The bottom panel of Fig. 14.9 also shows that the Fisher contours derived from the 2nd order PME around this model almost coincide with the ones derived from the halo-model covariance again.

14.C.3 Convergence in the General Case

Let us now consider the general case. We want invert the equation

$$\mathbf{C} = \mathbf{M}^{1/2} (\mathbf{1} + \mathbf{X}) \mathbf{M}^{1/2} \quad (14.40)$$

where

$$\mathbf{X} := \mathbf{M}^{-1/2} (\mathbf{B} - \mathbf{B}_m) \mathbf{M}^{-1/2}. \quad (14.41)$$

Since both \mathbf{M} and \mathbf{C} are positive definite matrices we can immediately infer that also the matrix $\mathbf{1} + \mathbf{X}$ must be positive definite, i.e. all its eigenvalues must be greater than 0. As a consequence, all eigenvalues λ_i of \mathbf{X} must fulfill

$$\lambda_i > -1 \quad \forall i. \quad (14.42)$$

In order to invert $\mathbf{1} + \mathbf{X}$ let us change into the eigenbasis of \mathbf{X} by means of an orthogonal matrix \mathbf{U} , i.e.

$$\mathbf{1} + \mathbf{X} = \mathbf{U}^T \text{diag}(1 + \lambda_i) \mathbf{U}. \quad (14.43)$$

It is not a priori clear whether we can invert this by means of the geometric series, since we do not know a priori that $|\lambda_i| < 1$. As discussed in section 14.5 in the case that $|\lambda_i| > 1$ the PME can at least help to identify differences between a covariance model and covariance from (possibly very few) simulations since in that case the 1st order and 2nd order PME will display a divergent behavior. However, since we know a priori that $\lambda_i > -1$ we can in principle apply a trick to let the PME series converge in any case. This trick is to expand $1/(1 + \lambda)$ not around $\lambda_0 = 0$ but around some other point $\lambda_0 = a > 0$:

$$\frac{1}{1 + \lambda} = \frac{1}{1 + a} \left[1 - \left(\frac{x - a}{1 + a} \right) + \left(\frac{x - a}{1 + a} \right)^2 - \dots \right]. \quad (14.44)$$

In terms of the PME series this is in fact equivalent to replacing the model covariance \mathbf{M} by $(1 + a)\mathbf{M}$. This way, one can in principle always ensure convergence of the series. This however comes at the expense of the series converging very slowly for eigenvalues of \mathbf{X} that are already close to or smaller than 0. Since in a real case scenario \mathbf{M} is assumed to be our best guess for the true covariance we hence recommend to stay with $a = 0$ and interpret a divergent PME as a significant difference between model and N-body covariance.

14.D Data vectors

14.D.1 Weak lensing data vectors

The redshift distribution and tomographic binning used for our LSST-like weak lensing data vector was chosen to be exactly that of Krause & Eifler (2016, see section 3). This means we assumed an overall source density of 26arcmin^{-2} and a source distribution with a median redshift of ≈ 0.7 that extends out to $z \gtrsim 3.0$. The tomographic bins were defined by first splitting the redshift distribution into 10 non-overlapping bins of equal source density and then assuming a Gaussian photoz uncertainty of $\sigma_z = 0.05$. The intrinsic ellipticity dispersion of the sources was assumed to be $\sigma_\epsilon = 0.26$ per ellipticity component.

The redshift distribution for the DES-like data vector was chosen to be shallower as for the LSST case reflecting the smaller depth of DES. Here our source distribution has a median redshift of ≈ 0.5 and extends out to $z = 2.0$. The overall source density was taken to be $10/\text{arcmin}^2$ and the 5 tomographic bins were defined assuming a photoz uncertainty of $\sigma_z = 0.08$. The intrinsic ellipticity dispersion was chosen to be the same as for the LSST-like case.

14.D.2 Lens galaxies

For the DES multi-probe data vector we also considered galaxy clustering and galaxy-galaxy lensing correlation functions. For this we were assuming a sample of foreground galaxies with a constant co-moving density motivated by the DES REDMAGiC sample (Rozo et al., 2016) divided into 3 tomographic bins whose redshift ranges are $(0.20, 0.35)$, $(0.35, 0.50)$ and $(0.50, 0.65)$. For these galaxies we assumed zero redshift uncertainties

motivated by the fact that the REDMAGIC redshift errors are small compared to the values for our source samples. The overall density of foreground galaxies was taken to be $0.15/\text{arcmin}^2$.

14.D.3 Binning and scales

The real space data vectors use 15 logarithmic angular bin from $\theta = 2.5'$ to $\theta = 250'$ for each correlation function and the Fourier space data vector uses 40 logarithmic bin from $\ell = 20$ to $\ell = 5000$ for each power spectrum. Data vector I contains the correlation functions ξ_+ and ξ_- for each possible combination of source bins. Data vector Ia also contains the auto-correlation of the lens bins and all possible combinations of lens-source correlations (i.e. only those combinations where the sources are at higher redshifts than the lenses). Data vector II contains the auto- and cross-power spectra of all possible combinations of source bins.

Part IV

Density Split Statistics

Chapter 15

Idea and formalism

The topic of this thesis are the higher order moments of the cosmic density field (where higher order is any order > 2). In the previous part (chapters 12 to 14) we only needed these moments to describe the statistical uncertainties of 2-point statistics. In this part (chapters 15 to 17) I investigate their potential as a cosmological probe in its own. In particular, I will describe *density split statistics* (DSS) which is a new cosmological probe that I helped develop and that is sensitive to the local probability density function (PDF) of the matter density contrast. As we saw in chapter 7 the PDF depends on the moments of the density field at all orders. Previous chapters that are required to understand this part of my thesis are

- Chapter 6 - Theory of Cosmic Structure Formation
(*Especially section 6.4 about non-linear perturbation theory.*)
- Chapter 7 - The Cosmic Density PDF
- Chapter 9 - 2D Power Spectra and Correlation Functions
(*Especially section 9.2 about galaxy bias.*)
- Chapter 10 - Gravitational Lensing

Unfortunately, the total matter density field in the universe is not a direct observable. As a consequence, there is no direct way to measure the PDF of density fluctuations. One can instead consider the galaxy density field and measure the PDF of fluctuations in that field. But as was explained in section 9.2, galaxies are both a biased and stochastic tracer of the total density field. I will examine this issue in more detail in chapter 16 but for now it suffices to acknowledge: galaxy density and matter density are correlated, but not identical.

Gravitational lensing on the other hand is directly caused by fluctuations of the total matter density (cf. chapter 10). This can be used indirectly to translate galaxy density into matter density. Recall that the lensing convergence κ is a line-of-sight projection of the density contrast (weighted by a geometric kernel; cf. section 10.4). Also recall that the

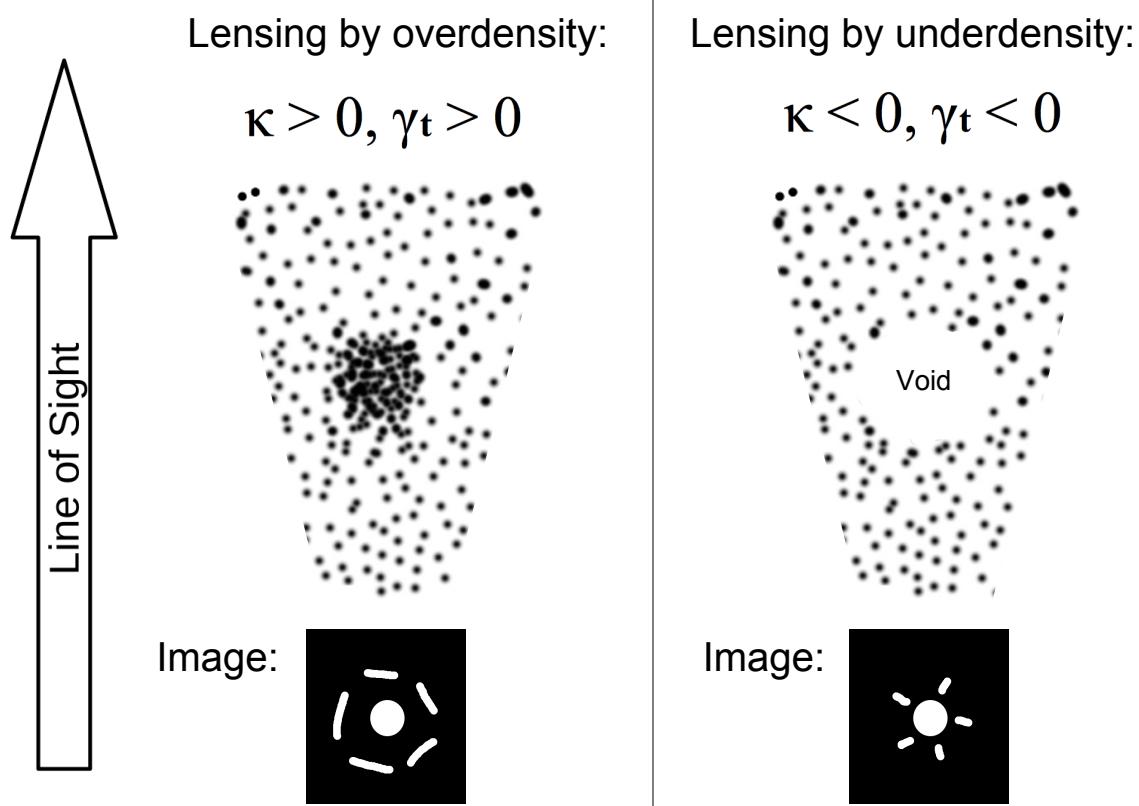


Figure 15.1: Left part of the sketch: The images of light sources behind a massive foreground object will appear tangentially stretched around this object due to gravitational lensing. Such a situation is typically described by a positive lensing convergence κ and a positive tangential component γ_t of the shear field (but not always; a region with $\kappa < 0$ can still cause a positive γ_t if it is surrounded by regions of even smaller κ). Right part of the sketch: The images of light sources behind an underdense region will appear radially stretched around this object due to gravitational lensing. Such a situation is typically described by a negative lensing convergence κ and a negative tangential component γ_t of the shear field).

tangential shear γ_t around a point on the sky quantifies how much images of light sources appear tangentially stretched around that point (cf. section 10.4.4). Then the tangential shear averaged along a circle of angular radius θ is related to the lensing convergence within that circle via

$$\langle \gamma_t \rangle_\theta = \langle \kappa \rangle_{<\theta} - \langle \kappa \rangle_\theta. \quad (15.1)$$

Here $\langle \kappa \rangle_{<\theta}$ is the convergence averaged *within* the circle and $\langle \kappa \rangle_\theta$ is the average convergence *at* the circle. In figure 15.1 I sketch the typical tangential shear signals that one expects around overdense and underdense signals in the sky. Around an overdense region, the line-of-sight density will typically drop and approach the average density of the universe from above. This means that the right hand-side of equation 15.1 will be positive, leading to a positive tangential shear signal. Around underdense regions, the line-of-sight density will typically rise and approach the average density of the universe from below. This will lead to a negative tangential shear signal (with signifies radial alignment of images).

This behavior of the tangential shear signal can be used to relate a map of galaxy density to the underlying matter density. If galaxy density and matter density are indeed correlated, then γ_t around points located in galaxy underdensities is expected to be negative while γ_t around points located in galaxy overdensities is expected to be positive. I sketch this in figure 15.2.

As I explain in chapter 16, the exact amplitude of γ_t around a region of a given galaxy density depends on two things: the matter density PDF and the relation between galaxies and matter. And as I also explain in chapter 16, one can disentangle these two influences by measuring in addition to γ_t also the PDF of galaxy density. This is the program of DSS. It can be summarized in the following steps:

1. Consider a sample of low redshift galaxies. Based on the positions of these galaxies, construct as map of projected galaxy density on the sky. This requires to smooth the discrete field of galaxy positions by some angular aperture -, e.g. , a circular top-hat filter of radius θ_T .
2. Measure the PDF of galaxy density (averaged over the radius θ_T). **This will be one part of the data vector analyzed by DSS.**
3. Split the map into quantiles of galaxy density. This could, e.g. , be the 20% of the area that has the lowest galaxy density or the 20% of the area that has the highest galaxy density. (The choice of 20% is arbitrary at this point; but I argue in chapter 16 that it is indeed a good choice.)
4. Around a uniform set of points in each quantile, measure the tangential shear $\gamma_t(\theta)$ at a number of radii θ . **This will be the second part of the data vector analyzed by DSS.**

As I demonstrate in chapter 16 these measurements can be used to simultaneously determine

- the relation between galaxy and total matter density,

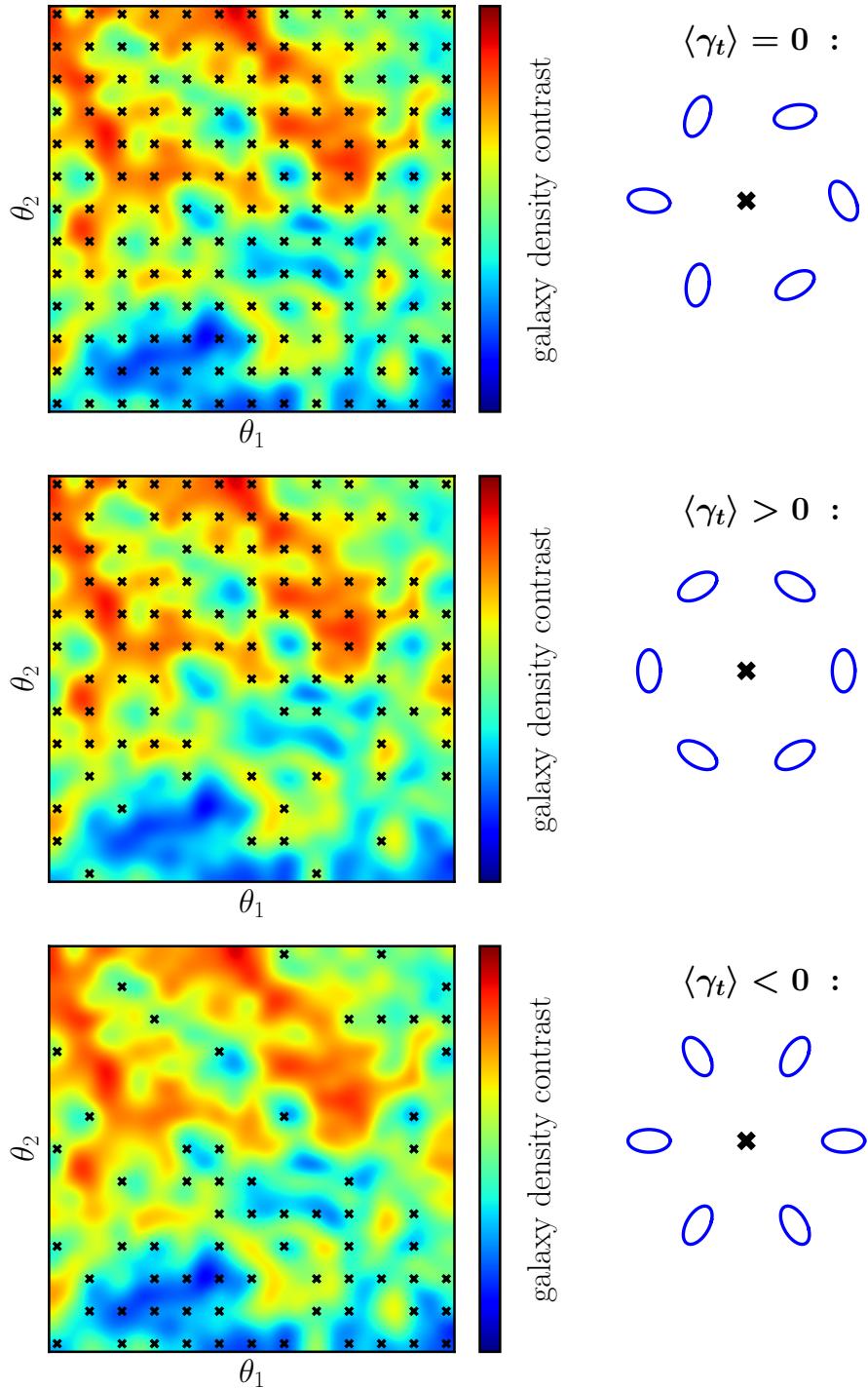


Figure 15.2: The tangential shear γ_t (of the images of background sources) around a uniform set of points on the sky is zero on average (upper panel). Around points in overdense regions γ_t will on average be > 0 (middle panel). And around points in underdense regions γ_t will on average be < 0 (lower panel; $\gamma_t < 0$ indicates radial alignment).

- the properties of the matter density PDF,
- parameters describing the background Λ CDM model.

This is significantly more qualitative information than obtained in analyses of 2-point statistics (see, e.g. , Kilbinger et al., 2013; van Uitert et al., 2018; DES Collaboration et al., 2017). The reason why DSS can provide that information is the following: it is not only sensitive to the amplitude of the galaxy density and matter density fields but also to their higher order moments (mostly their skewness). If one can predict the properties of the density PDF from within the Λ CDM model, then DSS also quantitatively becomes a competitive probe of that model (cf. chapter 17 where I present the results of Gruen, Friedrich et al., 2018).

However, the more interesting aspect of DSS is that it can be used to test predictions about the behavior of higher order density moments in a way that is almost independent of the exact parameters of the Λ CDM model. In particular we can measure the value of the scaling coefficient

$$S_3 = \frac{\langle \delta_R^3 \rangle_c}{\{\langle \delta_R^2 \rangle_c\}^2} \quad (15.2)$$

that relates the skewness $\langle \delta_R^3 \rangle_c$ of matter density to its variance $\langle \delta_R^2 \rangle_c$ (cf. section 7.3 for the definition of the coefficients S_n). The Λ CDM predictions for this parameter are almost independent of the parameters Ω_m and σ_8 that are typically a focus of 2-point statistical analyses. Another interesting aspect of DSS is that it probes the relation of galaxies and their environment in much more detail than can be achieved with 2-point statistics (see the next chapter for details).

I have contributed to the following 3 journal articles related to DSS:

- Gruen, Friedrich et al., 2016, MNRAS:

In this paper we tested for the first time Daniel Gruen's idea of measuring the tangential shear profile around quantiles of different galaxy density. My part of this work was to derive a model for these profiles based on the assumptions that

- a) both the matter density contrast and the lensing convergence are Gaussian random fields,
- b) the relations between the matter density contrast and the density contrast of luminous red galaxies can be described with a linear bias model + Poissonian shot-noise.

With this model we could show that the main features of density split statistics can be derived from simple assumptions within the Λ CDM model. In figure 15.3 I show the values of γ_t measured along different angular radii θ around the 20% most underdense and the 20% most overdense area of early DES data (Rozo et al., 2016, using the REDMAGIC sample of luminous red galaxies to determine overdensities and underdensities). As you can see in that figure, there is an asymmetry in the amplitudes

of the overdense and underdense signal that is not captured by the Gaussian model. This disagreement was not significant in that early data but is let us suspect, that the skewness of the density field or density dependence in the relation of galaxies and matter might indeed be detectable with DSS.

- Friedrich, Gruen et al., 2018, PRD:

In this paper I developed a more realistic model for DSS that uses cosmological perturbation theory to predict the behavior of the density PDF and that allows for a density dependence in the behavior of galaxies and matter. We then test the accuracy of this model using N-body simulations. This paper is presented in chapter 16.

- Gruen, Friedrich et al., 2018, PRD:

In this paper we analyze year 1 data of DES based on the improved model. I summarize the results of this paper in chapter 17.

In figure 15.4 I compare plots of the lensing signals obtained in each of the papers (see chapters 16 and 17 for details).

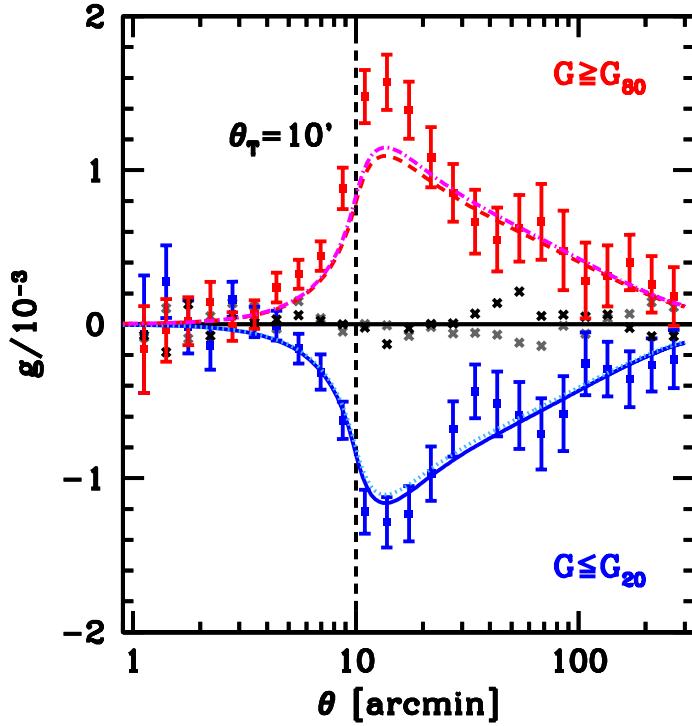
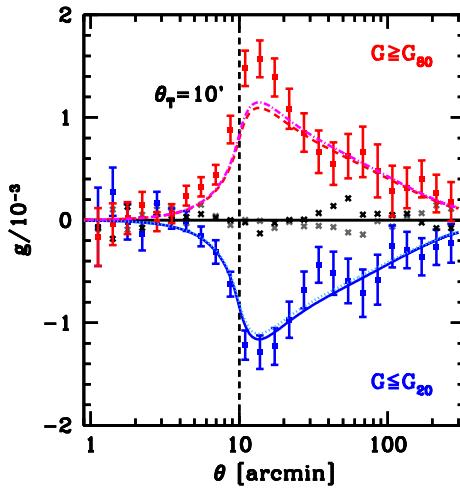


Figure 15.3: Figure from Gruen, Friedrich et al., 2016. The blue points show measurements of γ_t around the 20% of the area of early DES data that was most underdense in terms of REDMAGIC galaxies (Rozo et al., 2016). The red points show the corresponding measurements around the 20% of the area of early DES data that was most overdense. Our map of galaxy density was constructed using a circular top-hat filter of radius $\theta_T = 10'$. This explains why signals drop off at $\theta < 10'$. The blue solid and red dashed lines show the model that I developed, assuming that the density field is a Gaussian random field. As you can see, there is an asymmetry in the amplitudes of the overdense and underdense signal that is not captured by the Gaussian model (though this is not yet significant in this early data).

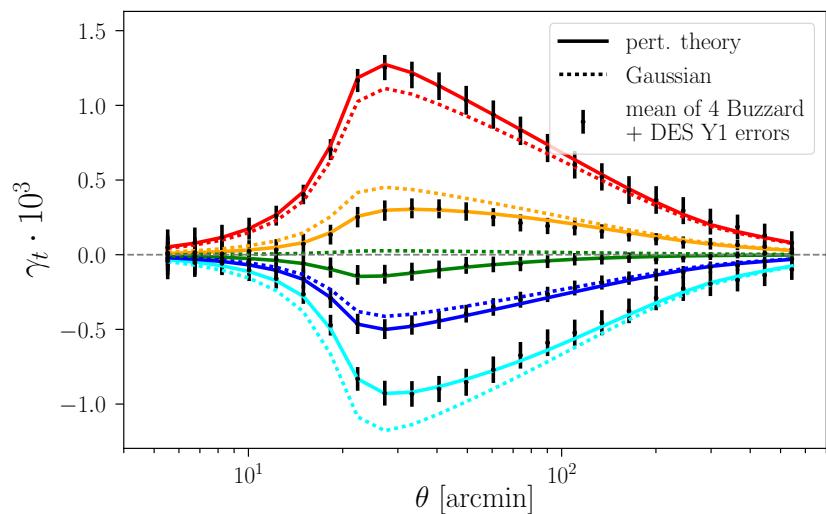
Gruen, Friedrich et al., 2016:

Measurement in DES science verification data compared to a Gaussian model. This early model does not capture the skewness of the cosmic density field.



Friedrich, Gruen et al., 2018:

Our updated model compared to simulated data. The error bars represent uncertainties of DES year 1 data.



Gruen, Friedrich et al., 2018:

Measurement in DES year 1 data and our best fit model.

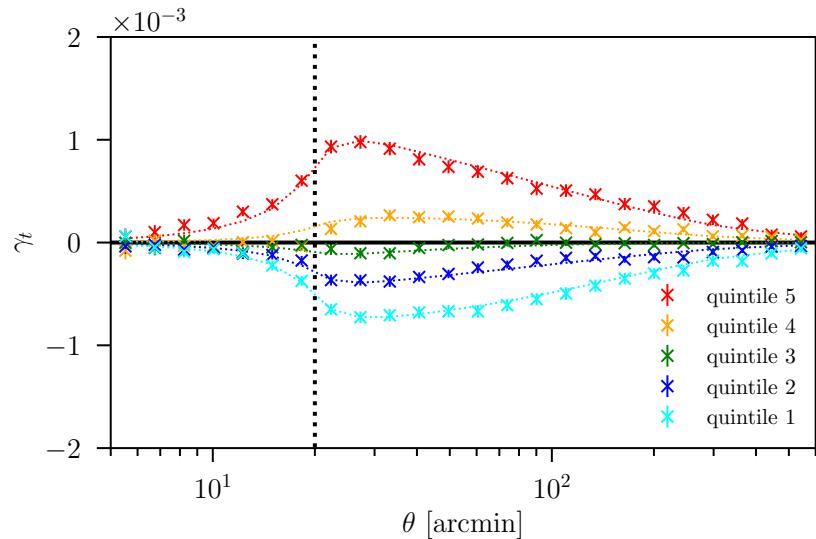


Figure 15.4: Description see main text.

Chapter 16

Density split statistics: joint model of counts and lensing in cells

In Gruen, Friedrich et al., 2016 we presented an exploratory study of density split statistics (DSS) in early DES data (the so-called science verification data set). We did not perform a cosmological analysis of DSS in that data. The model we used in Gruen, Friedrich et al., 2016 was too unrealistic for that, because it assumed that the matter density contrast is a Gaussian random field.

In the following chapter I will present an improved model that uses cosmological perturbation theory to predict the behavior of the density PDF and that allows for a density dependence in the behavior of galaxies and matter. This model was the basis of our cosmological analysis of DES year 1 data and SDSS data presented in Gruen, Friedrich et al., 2018. (The results of that paper are summarized in chapter 17.)

The present chapter has been published as Friedrich, Gruen, DeRose, Kirk, Krause, McClintonck, Rykoff, Seitz, Wechsler, Bernstein, Blazek, Chang, Hilbert, Jain, Kovacs, Lahav et al. (2018) in PRD. I exclusively developed the model for density split statistics presented in that article (building on earlier work done by Francis Bernardeau, Patrick Valageas and others that is referenced in the paper). I also exclusively designed the numerical code used to calculate predictions from that model and programmed most of that code by myself. Furthermore, I carried out the majority of the analysis presented in this paper (exceptions are mentioned below).

Daniel Gruen and I have in cooperation developed the concept of *density split statistics* (DSS) on which this article is based. DSS is a generalization of *trough lensing*, which was first proposed by Mr. Gruen. He also contributed to this article in particular in several ways: He prepared some of the data vectors, the covariance matrix and in part also the projected density maps used in this work from simulated data. He has programmed the likelihood pipeline used in this study as well as the `C++-to-python` interface that links my modeling code with this pipeline. Together with Donnacha Kirk he has also carried out the measurements of galaxy bias and galaxy stochasticity that are used to test individual

ingredients of the presented model. Finally, Mr. Gruen has in part written the sections 16.3.2 and 16.4.3.

Joe DeRose, Eli Rykoff and Risa Wechsler have developed and produced the mock data from the Buzzard N-body simulations that is used in this work. From these simulations Joe DeRose has also derived particular data products needed for this analysis and he is the author of section 16.3.1.

Elisabeth Krause and Tom McClintock helped to significantly increase the speed of my modeling code. Elisabeth Krause also contributed through discussions about our galaxy bias and stochasticity model and her insight into large scale structure cosmology helped us to put the project into adequate context with other cosmological probes.

Stella Seitz contributed through frequent, detailed discussions about the project (especially emphasizing the galaxy evolution aspect of it and helping to define future perspectives). All authors mentioned above have contributed to discussions and proofreading. Authors that are not mentioned have contributed to the overall infrastructure of the Dark Energy Survey.

Permission for non-commercial re-use of the material included in this thesis has been granted by the PRD editorial office. The American Physical Society holds the copyright on the paper.

ABSTRACT

We present density split statistics, a framework that studies lensing and counts-in-cells as a function of foreground galaxy density, thereby providing a large-scale measurement of both 2-point and 3-point statistics. Our method extends our earlier work on trough lensing and is summarized as follows: given a foreground (low redshift) population of galaxies, we divide the sky into subareas of equal size but distinct galaxy density. We then measure lensing around uniformly spaced points separately in each of these subareas, as well as counts-in-cells statistics (CiC). The lensing signals trace the matter density contrast around regions of fixed galaxy density. Through the CiC measurements this can be related to the density profile around regions of fixed matter density. Together, these measurements constitute a powerful probe of cosmology, the skewness of the density field and the connection of galaxies and matter.

In this paper we show how to model both the density split lensing signal and CiC from basic ingredients: a non-linear power spectrum, clustering hierarchy coefficients from perturbation theory and a parametric model for galaxy bias and shot-noise. Using N-body simulations, we demonstrate that this model is sufficiently accurate for a cosmological analysis on year 1 data from the Dark Energy Survey.

16.1 Introduction

The large-scale structure (LSS) observed today is thought to originate from almost perfectly Gaussian density perturbations in the early Universe. This means that there was a complete symmetry in the abundance and amplitude of underdense and overdense regions in very early times. Gravitational attraction then caused initial overdensities to collapse to small but highly overdense structures such as galaxy clusters, while initial underdensities expanded but stayed moderately underdense and e.g. became voids. As a consequence the majority of the volume in the late-time Universe is underdense, compensated by the presence of few highly overdense spots. Or, in other words, a positive skewness in the distribution of density fluctuations emerges due to gravitational collapse.

A variety of probes have been used to study the statistical properties of the late-time density field and to thereby understand the physics of gravitational collapse as well as the processes responsible for the properties of the initial density fluctuations. So far, the most extensive studies have been carried out on the 2-point statistics of density fluctuations, i.e. on measuring the variance of density fluctuations as a function of scale. This has e.g. been done through measurements of cosmic shear 2-point correlation functions (e.g. Wittman et al., 2000; Van Waerbeke et al., 2000; Benjamin et al., 2007; Fu et al., 2008; Schrabback et al., 2010; Kilbinger et al., 2013; Becker et al., 2016; Jee et al., 2016; Hildebrandt et al., 2017; Troxel et al., 2017), galaxy clustering (e.g. Crocce et al., 2016; Elvin-Poole et al., 2017; Alam et al., 2017) and galaxy-galaxy lensing (e.g. Brainerd et al., 1996; Hudson et al., 1998; Wilson et al., 2001; van Uitert et al., 2011; Brimioule et al., 2013; Clampitt et al.,

2017; Prat et al., 2017) as well as combined measurements thereof (e.g. Mandelbaum et al., 2013; van Uitert et al., 2018; DES Collaboration et al., 2017).

While 2-point statistics are only sensitive to the overall amplitude of density fluctuations, higher-order statistics also know about the skewness arising from the different behavior of underdense and overdense regions. This does not necessarily mean that higher-order statistics are better than 2-point statistics in discriminating between particular choices of cosmological parameters (Barreira et al., 2017). But they scale differently with parameters such as Ω_m , σ_8 , galaxy bias and galaxy stochasticity than their 2-point counterparts. Hence, in a cosmological analysis that varies a large number of parameters, probes that are sensitive to both 2-point and higher order statistics have the power to break degeneracies between these parameters (Bernardeau et al., 1997; Takada & Jain, 2002; Pires et al., 2012; Uhlemann et al., 2018a).

Observations of higher-order statistical features of the density field include measurements of three point correlation functions (Semboloni et al., 2011), shear peak statistics (Lin & Kilbinger, 2015; Liu et al., 2015; Kacprzak et al., 2016) and the cluster mass function (Mantz et al., 2016). Also, a number of probes have been suggested (and in some cases measured in data) that study the correlation of 2-point statistics and background density. Chiang et al. (2015) have measured this by means of the integrated bispectrum. Simpson et al. (2011, 2013, 2016) have proposed a clipped power spectrum approach, where 2-point statistics are measured on the sky after excluding high density regions. They have shown that these measurements contain information complementary to the corresponding measurements on the full sky.

A similar direction was investigated by Gruen et al. (2016) who separately measured the lensing power spectrum in underdense and overdense lines of sight. The framework presented in this paper is based on their concept of trough lensing. We will call it *density split lensing* when only lensing measurements are involved and *density split statistics* when it is combined with counts-in-cells measurements. This method can be summarized as follows: we consider a foreground (low redshift) population of galaxies and smooth their position field with a circular top-hat aperture. This smoothed density field is then used to divide the sky into sub-areas of equal size but distinct galaxy density. In this paper we consider in particular 5 sub-areas and call them quintiles of galaxy density. As a next step, we use a background (high redshift) population of galaxies to measure the tangential shear of these galaxies around a set of uniformly spaced points within the area of each density quintile. The resulting lensing signals trace the matter density contrast around regions of fixed foreground galaxy density. This data vector is then complemented by the histogram of counts-in-cells of the foreground galaxies to pin down their bias and stochasticity. As we show in this paper, a cosmological analysis based on this density split data vector has a number of desirable features:

- it allows for an accurate analytic modeling with the help of cosmological perturbation theory and a non-linear power spectrum,
- it yields high signal-to-noise measurement,

- it avoids systematics common to cosmic shear such as additive shear biases or intrinsic alignments (as long as tracer sample and source sample do not overlap in redshift),
- it has a very intuitive interpretation.

This paper is a companion paper of Gruen et al. (2018), where we present our actual data analysis, including tests for systematic effects as well as a description of how we estimate the covariance of our signal. This paper is presenting the modeling framework used in that analysis. Our section 16.2 gives a general overview of density split statistics: we describe our data vector, explain how it can be modeled and also present forecasts on cosmological parameter constraints, both for a Λ CDM model and an extended model that allows gravitational collapse to behave differently than within general relativity. In section 16.3 we describe the simulated data used in this work. Section 16.4 explains details of the model presented in section 16.2. There we also compare individual components of this model directly to measurements in N-body simulations. In section 16.5 we show that our model for a data vector combining density split lensing and counts-in-cells statistics is accurate enough to recover the cosmology underlying our N-body simulations. Any possible deviation between our model and the simulations is shown to be well within statistical uncertainties of year 1 data of the Dark Energy Survey (DES Y1).

In appendix 16.A, we review a number of differential equations that govern gravitational collapse. In appendix 16.B, we review the leading order perturbative calculation of the 3-point statistics of the cosmic density field for a general Λ CDM model. Appendix 16.C qualitatively compares our model of the cosmic density PDF to a second set of N-body simulations as a complement to the comparison carried out in the main text. Appendix 16.D derives properties of joint log-normal random fields and appendix 16.E repeats the validation of our model for an alternative shot-noise parametrization.

16.2 Density split statistics: data vector, modeling and forecasts

This section provides an introduction to the program of density split statistics. In section 16.2.1 we describe how we obtain the density split lensing signal and how this signal can be further complemented with information on galaxy bias and stochasticity from counts-in-cells. In section 16.2.2 we outline our modeling of this signal (but postponing technical details of this model to section 16.4). In section 16.2.3 we provide forecasts on the cosmological information that can be obtained with a measurement of density split statistics in year-1 data of the Dark Energy Survey (DES Y1).

16.2.1 Measuring density split statistics

Density split lensing is a generalization of trough lensing (Gruen et al., 2016) and can be described in three steps:

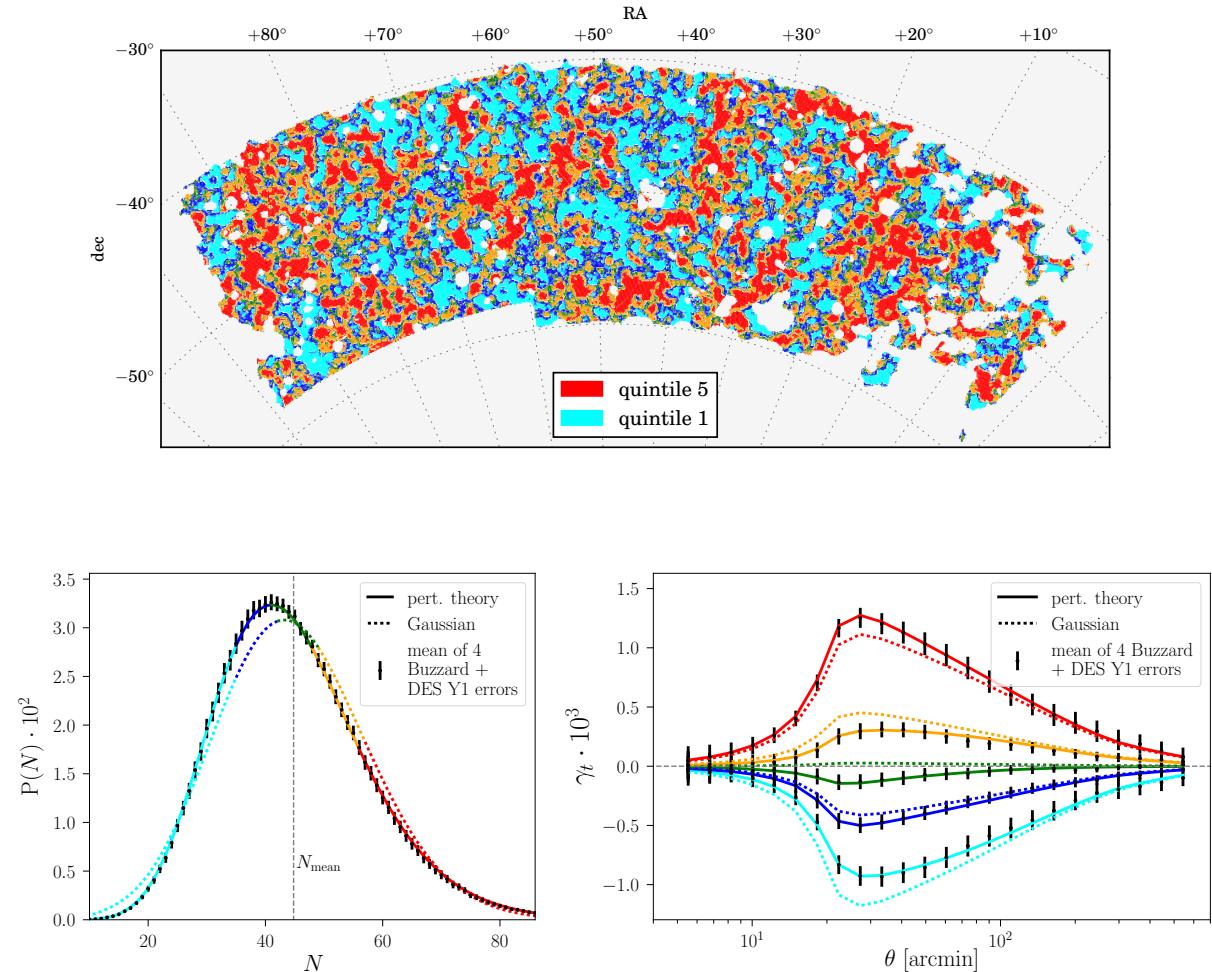


Figure 16.1: Top panel: splitting the lines of sight in one DES-Y1 like Buzzard simulation into 5 quantiles of galaxy density (color coding from cyan, most underdense, to red, most overdense). The map uses a 20 arcmin top-hat radius and REDMAGIC galaxies with a redshift range of $0.2 \lesssim z \lesssim 0.45$. Bottom left: histogram of REDMAGIC galaxy counts in 20 arcmin radii (counts-in-cells). We show the mean histogram from 4 Buzzard realizations of DES-Y1 (black points), our model based on perturbation theory and cylindrical collapse (solid line) and a model that assumes the projected density contrast to be a Gaussian random field (dotted line). The color coding corresponds exactly to the density quantiles in the top panel. Bottom right: Lensing signals around random points split by the density quantile in which these points are located. We show the mean measurement from 4 Buzzard realizations (black points), our perturbation theory model (solid line) and a model that assumes projected density contrast and lensing convergence to be joint Gaussian random variables (dotted line). Color coding is the same as in the other panels. The asymmetry between the lensing signals around the most underdense and most overdense lines-of-sight indicates the skewness of the cosmic density PDF.

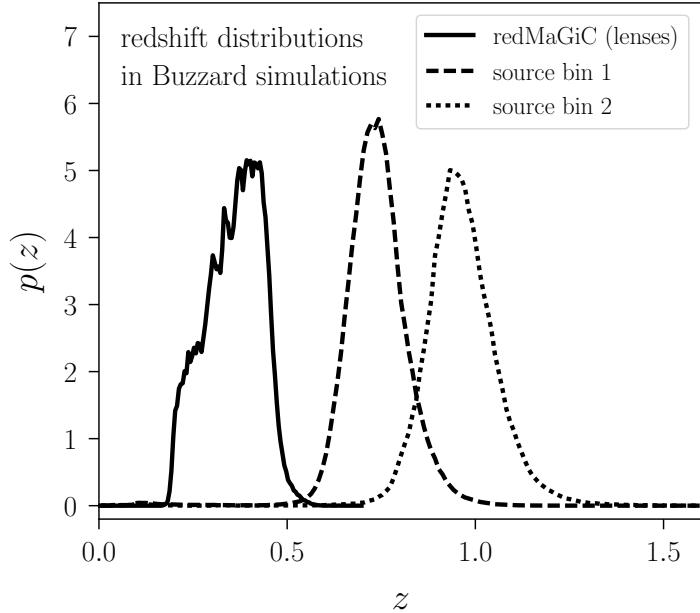


Figure 16.2: Redshift distributions of the tracer galaxy sample and the source samples of our N-body realizations of DES-Y1.

1.) *Splitting the sky into quantiles of different foreground galaxy density*

Consider a sample of low-redshift galaxies that are tracing the line-of-sight density of matter with some redshift distribution $n_l(z)$. We will call these galaxies the foreground sample. For an angular radius θ_T , which we will call the *top-hat aperture radius*, we define $N_T(\hat{\mathbf{n}})$ to be the number of galaxies found within a radius θ_T around the point on the sky specified by the vector $\hat{\mathbf{n}}$ on the unit sphere. The field $N_T(\hat{\mathbf{n}})$ can be used to divide the sky into regions of different galaxy density. Gruen et al. (2016) have done this by discretizing the sky with a HEALPIX¹ grid and sorting the pixels according to their value of $N_T(\hat{\mathbf{n}})$. Then they considered the 20% of the pixels with the lowest values of N_T , calling them troughs. In the limit of a fine pixelization these pixels can be considered the most underdense quintile of the sky area. This can be generalized to the second most underdense quintile, the third most underdense quintile etc. or even to finer splits using more than just 5 quantiles.

We stick to a division into 5 quantiles (quintiles) throughout this paper. The upper panel of figure 16.1 illustrates such a subdivision on a patch of a simulated sky (from the *Buzzard flock*, see section 16.3 and especially DeRose et al. (2018) for details). There we use a top-hat aperture radius $\theta_T = 20'$ and the tracer galaxies have the redshift distribution that is displayed by the solid line in Figure 16.2. Figure 16.1 shows the most underdense quintile of the simulated patch in cyan, the most overdense quintile in red and the three

¹See Górski et al. (2005) for details on HEALPIX.

intermediate quintiles in blue, green and orange.

Note that the sum of the 5 lensing signals will vanish on average (we subtracted 1/5 times the shear around random points from each signal, but due to boundary effects their sum will not vanish exactly). This means that roughly 4 of the 5 signals contain independent information. We have not investigated, whether our choice of 5 quantiles is in any way optimal. Choosing 3 quantiles would leave us with 2 independent signals and would hence suffice to be sensitive to both the variance and skewness of the density field. 5 quantiles enable a sensitivity beyond the 3rd moment of the density field. And they also allow us to explicitly show, that the median universe is underdense (which we could not do with 4 quantiles). In section 16.4 we investigate different radii of our top-hat aperture and find that $\theta_T = 20'$ is the smallest radius at which our model is reliable (given the redshift distribution we use in Gruen et al. (2018)).

2.) Tracing the mean dark matter density in each sky quintile with gravitational lensing

Now consider a second sample of galaxies at higher redshifts than the foreground sample (the *source sample*, see e.g. the dashed and dotted redshift distributions in Figure 16.2). As the light of these galaxies passes the large-scale structure of the foreground density distribution it undergoes gravitational lensing effects such as gravitational shear (see e.g. Bartelmann & Schneider, 2001). The density split lensing signal around each quintile of the sky is obtained by measuring the stacked radial profile of tangential shear around random points located within that quintile. These points are constrained to lie within the part of the sky covered by a certain quintile of galaxy density but are otherwise random in their location. Because these random points are split according to the density quintile they are located in, their stacked shear signals trace the average profile of density contrast around each quintile.

In the lower right panel of Figure 16.1 we show the signals measured for each density quintile in our mock data. The points show the average measurement from 4 Buzzard realizations of DES year-1 data (using the highest redshift source population shown in Figure 16.2) and the solid lines show predictions by the model presented in this paper. The error bars are derived from a set of log-normal realizations (using the **FLASK** tool by Xavier et al. (2016); in Gruen et al. (2018) we describe in detail how we configured **FLASK** to generate our mock catalogs). Two main features of the density split lensing signals are apparent: first, the amplitude of the radial shear around the 20% most underdense pixels is lower than the amplitude of the tangential shear around the 20% most overdense pixels. This is reflecting the skewness of the cosmic density PDF. Secondly, the signal around points in the third quintile is still significantly negative, which reflects the fact that the median universe is underdense. A more subtle feature is the fact that the underdense signals fall off less rapidly with increasing scale than the overdense signals. This is because on large scales the density field becomes Gaussian and hence recovers its initial symmetry between overdensities and underdensities.

3.) Measuring the average counts-in-cells in each density quintile to obtain additional in-

formation on galaxy bias and stochasticity

If galaxy counts and the matter density field were perfectly correlated, then a split of the sky by galaxy density would be identical to a split by matter density. Hence, in this limit the density split lensing signals would be independent of the bias of the tracer galaxies. In a realistic scenario however, shot-noise of the galaxies smears out our attempts to divide the sky into areas of different matter density. Hence the density split lensing signals obtain a dependence on galaxy bias, but also on galaxy stochasticity. Increasing the linear bias of galaxy clustering will sharpen the tracers' ability to distinguish between overdensities and underdensities. Thus, increasing this bias will increase the amplitudes of the signals. This means that linear bias is to some degree degenerate with the amplitude of density fluctuations, σ_8 . But σ_8 and bias influence the third moments of the density field in different ways and their degeneracy is not complete. As a consequence, it is possible to obtain constraints on cosmological parameters from the lensing signals alone (cf. section 16.2.3 and the blue contour in the left panel of figure 16.3).

But additional information on bias and stochasticity nevertheless helps to tighten these constraints. In this paper we decided to add that information in the form of normalized quantiles of the counts-in-cells (CiC) histogram of the tracer galaxies: we measure the histogram of tracer counts within the same aperture that was used to identify our density quintiles. Then we identify the parts of this histogram that correspond to these quintiles (cf. lower left panel of Figure 16.1). For each quintile q we then determine the mean galaxy count in that quintile, N_q , and normalize it by the overall mean galaxy count in our aperture, \bar{N} . i.e. for each quintile we add N_q/\bar{N} to our data vector. This indeed helps to tighten constraints on cosmological parameters (cf. section 16.2.3 and the green contour in the left panel of figure 16.3).

16.2.2 Modeling density split statistics

We now outline a general framework for modeling the data vector described above, leaving details of this framework to section 16.4. Unless stated differently, we will assume a flat Λ CDM universe throughout this paper.

Let us start by introducing the quantities whose relations need to be modeled. First, we denote with $\delta_{m,2D}$ the line-of-sight projection of the 3D density contrast according to the redshift distribution $n_l(z)$ of our foreground galaxy sample, i.e.

$$\delta_{m,2D}(\hat{\mathbf{n}}) = \int dw q_l(w) \delta_{m,3D}(w\hat{\mathbf{n}}, w) \quad (16.1)$$

where $\hat{\mathbf{n}}$ denotes a unit vector on the sky, w is co-moving distance and the projection kernel $q_l(w)$ is given in terms of $n_l(z)$ as

$$q_l(w) = n_l(z[w]) \frac{dz[w]}{dw} . \quad (16.2)$$

We furthermore define $\delta_{m,T}$ to be the average of $\delta_{m,2D}$ over top-hat filters with aperture radius θ_T , i.e.

$$\delta_{m,T}(\hat{\mathbf{n}}) = \int_{|\hat{\mathbf{n}}, \hat{\mathbf{n}}'| < \theta_T} d\Omega' \frac{\delta_{m,2D}(\hat{\mathbf{n}}')}{2\pi(1 - \cos \theta_T)} . \quad (16.3)$$

Here $|\cdot, \cdot|$ denotes the angular distance between two points on the sky.

We identify regions of different density by means of our foreground galaxy sample. When smoothed with a top-hat filter of radius θ_T , these galaxies are biased and possibly stochastic tracers of $\delta_{m,T}(\hat{\mathbf{n}})$. Hence our model also needs to include a description of how $N_T(\hat{\mathbf{n}})$, the number of tracer galaxies found within an angular radius θ_T around the line-of-sight $\hat{\mathbf{n}}$, relates to $\delta_{m,T}(\hat{\mathbf{n}})$.

Finally, in order to describe the density split lensing signal, we need to consider the lensing convergence field for our population of source galaxies. Given the source redshift distribution $n_s(z)$, the convergence κ is given by the line-of-sight projection

$$\kappa(\hat{\mathbf{n}}) = \int dw W_s(w) \delta_{m,3D}(w\hat{\mathbf{n}}, w) , \quad (16.4)$$

where W_s is the lensing efficiency, which is defined by

$$W_s(w) = \frac{3\Omega_m H_0^2}{2c^2} \int_w^\infty dw' \frac{w(w' - w)}{w' a(w)} q_s(w') , \quad (16.5)$$

and $q_s(w) = n_s(z[w]) \frac{dz[w]}{dw}$ is the line-of-sight density of the sources. Smoothing the convergence field with a circular aperture of radius θ results in a field which we will denote by $\kappa_{<\theta}(\hat{\mathbf{n}})$.

Because of the isotropy of the universe, we will now omit the dependence of the above quantities on $\hat{\mathbf{n}}$. To model the density split lensing signal one needs to answer the following questions:

- Given the number of galaxies N_T found around a line-of-sight $\hat{\mathbf{n}}$, what distribution can be inferred for the matter density contrast $\delta_{m,T}$ in that line-of-sight? i.e. what is the expectation value $\langle \delta_{m,T} | N_T \rangle$?
- Given the matter density contrast $\delta_{m,T}$ in the line-of-sight $\hat{\mathbf{n}}$, what lensing convergence $\kappa_{<\theta}$ is expected inside an angular distance θ from that line-of-sight? i.e. what is the expectation value $\langle \kappa_{<\theta} | \delta_{m,T} \rangle$? The tangential shear profile around that line-of-sight can then be inferred from the convergence profile as

$$\begin{aligned} \langle \gamma_t(\theta) | \delta_{m,T} \rangle &= \langle \kappa_{<\theta} | \delta_{m,T} \rangle - \langle \kappa_\theta | \delta_{m,T} \rangle \\ &= \frac{\cos \theta - 1}{\sin \theta} \frac{d}{d\theta} \langle \kappa_{<\theta} | \delta_{m,T} \rangle , \end{aligned} \quad (16.6)$$

where κ_θ is the average convergence at the radius θ .

The first of the above questions can be answered in the form of a conditional PDF of $\delta_{m,T}$ given a certain value of N_T , i.e. $p(\delta_{m,T}|N_T)$. Using Bayes' theorem this can be written as

$$p(\delta_{m,T}|N_T) = \frac{P(N_T|\delta_{m,T}) p(\delta_{m,T})}{P(N_T)} , \quad (16.7)$$

where $P(N_T|\delta_{m,T})$ is the probability of finding a number of galaxies N_T given that the density contrast is $\delta_{m,T}$ and where $p(\delta_{m,T})$ and $P(N_T)$ are the total PDF of $\delta_{m,T}$ and the total probability of finding N_T tracer galaxies. The average convergence profile around a circle with N_T galaxies is then given by

$$\begin{aligned} \langle \kappa_{<\theta} | N_T \rangle &= \int d\delta_{m,T} \langle \kappa_{<\theta} | \delta_{m,T}, N_T \rangle p(\delta_{m,T}|N_T) \\ &\approx \int d\delta_{m,T} \langle \kappa_{<\theta} | \delta_{m,T} \rangle p(\delta_{m,T}|N_T) , \end{aligned} \quad (16.8)$$

where in the second step we have assumed that the expected convergence within θ only depends on the total matter density contrast within θ_T .

We now divide the sky into different quintiles of tracer galaxy density. Let us denote with $Q[0.0, 0.2]$ the 20% of the lines-of-sight on the sky that have the lowest value of N_T . There will be a maximal value $N_T \leq N_{\max}$ in this quintile and the stacked convergence profile around these lines-of-sight is given by

$$\begin{aligned} \langle \kappa_{<\theta} | Q[0.0, 0.2] \rangle &= \\ \frac{1}{0.2} \left(\sum_{N < N_{\max}} P(N) \langle \kappa_{<\theta} | N_T = N \rangle + \alpha \langle \kappa_{<\theta} | N_T = N_{\max} \rangle \right) . \end{aligned} \quad (16.9)$$

Here the factor α in the second term accounts for the fact that the lines-of-sight with exactly $N_T = N_{\max}$ might have to be split between the quintile $Q[0.0, 0.2]$ and the quintile $Q[0.2, 0.4]$. It is given by

$$\alpha = 0.2 - \sum_{N < N_{\max}} P(N) . \quad (16.10)$$

This can be easily generalized to the other quintiles $Q[q_{\min}, q_{\max}]$ and also to the case of dividing the sky into more than 5 density regimes. Finally, we also add the average of the counts-in-cells in each quintile normalized by the mean galaxy count \bar{N} to our data vector. For the quintile $Q[0.0, 0.2]$ this is given by

$$\frac{\langle N_T | Q[0.0, 0.2] \rangle}{\bar{N}} = \frac{1}{0.2\bar{N}} \left(\sum_{N < N_{\max}} P(N) N + \alpha N_{\max} \right) , \quad (16.11)$$

which is also straightforward to generalize to other quantiles $Q[q_{\min}, q_{\max}]$.

The probabilities $P(N)$ can be computed from the normalization of equation 16.7. Hence, our model for the density split lensing signal needs the following three ingredients:

(i) The PDF of matter density contrast, smoothed with a top-hat filter of radius θ_T ,

$$p(\delta_{m,T}) . \quad (16.12)$$

(ii) The expectation value of convergence inside a distance θ given the density contrast inside θ_T ,

$$\langle \kappa_{<\theta} | \delta_{m,T} \rangle . \quad (16.13)$$

(iii) The distribution of galaxy counts inside the top-hat radius θ_T given the density contrast within that radius,

$$P(N_T | \delta_{m,T}) . \quad (16.14)$$

Gruen et al. (2016) assumed $\delta_{m,T}$ and $\kappa_{<\theta}$ to have a joint Gaussian distribution. This allowed them to compute (i) and (ii) solely from the dark matter clustering power spectrum. To compute (iii) they assumed a linear galaxy bias and Poissonian shot-noise of the tracer galaxies. These assumptions allowed a sufficient model for their measurements made on DES Science Verification data. But as can be seen from the dotted lines in the lower panels of Figure 16.1, a Gaussian model for the density PDF is not sufficient within the much smaller uncertainty of DES-Y1. Also, in section 16.4 we demonstrate that the shot-noise of the tracer galaxies can not necessarily be assumed to be Poissonian. In this work we hence want to revise their model.

For each of the model components (i) and (ii) we investigate two different modeling approaches - a baseline approach and an approach with increased complexity used to check the validity of the baseline model. In the following we briefly outline each approach. The reader interested in details of each modeling ansatz is referred to section 16.4. Readers who are not interested in this technical part of the paper should feel free to skip section 16.4.

(i) Baseline model for $p(\delta_{m,T})$:

In our fiducial model we assume $\delta_{m,T}$ to be a log-normal random field (Hilbert et al., 2011). The PDF of such a variable can e.g. be fixed by specifying the variance $\langle \delta_{m,T}^2 \rangle$ and skewness $\langle \delta_{m,T}^3 \rangle$. We predict the variance of $\delta_{m,T}$ from the non-linear matter power spectrum (cf. Gruen et al., 2016). The latter is computed using *halofit* (Smith et al., 2003; Takahashi et al., 2012) and an analytic transfer function (Eisenstein & Hu, 1998). We then use leading order perturbation theory to compute a scaling relation between the bispectrum and the power spectrum of the density field. Together with our power spectrum this fixes the skewness of $\delta_{m,T}$.

Alternative model for $p(\delta_{m,T})$:

As an alternative we compute the PDF $p(\delta_{m,T})$ from its cumulant generating function (see section 16.4 for a definition and further details). To model this function we use a cylindrical collapse approach based on the work of Bernardeau (1994), Bernardeau & Valageas (2000) and Valageas (2002a).

(ii) Baseline model for $\langle \kappa_{<\theta} | \delta_{m,T} \rangle$:

In our fiducial model we assume that $\kappa_{<\theta}$ can be expressed as the sum of two random variables,

$$\kappa_{<\theta} = \kappa_{<\theta, \text{corr.}} + \kappa_{<\theta, \text{uncorr.}} , \quad (16.15)$$

where $\kappa_{<\theta, \text{uncorr.}}$ is assumed to be completely uncorrelated to $\delta_{m,T}$ and hence doesn't contribute to the density split lensing signal. As a result we have

$$\langle \kappa_{<\theta} | \delta_{m,T} \rangle \equiv \langle \kappa_{<\theta, \text{corr.}} | \delta_{m,T} \rangle . \quad (16.16)$$

We assume a joint log-normal PDF for the two random variables $\delta_{m,T}$ and $\kappa_{<\theta, \text{corr.}}$. The expectation value $\langle \kappa_{<\theta, \text{corr.}} | \delta_{m,T} \rangle$ is then fixed by specifying the moments

$$\langle \delta_{m,T}^2 \rangle , \langle \delta_{m,T}^3 \rangle \quad (16.17)$$

as well as

$$\langle \kappa_{<\theta} \delta_{m,T} \rangle \equiv \langle \kappa_{<\theta, \text{corr.}} \delta_{m,T} \rangle \quad (16.18)$$

and

$$\langle \kappa_{<\theta} \delta_{m,T}^2 \rangle \equiv \langle \kappa_{<\theta, \text{corr.}} \delta_{m,T}^2 \rangle . \quad (16.19)$$

Second order moments are again computed from our non-linear power spectrum while third order moments are inferred from perturbation theory. (The introduction of $\kappa_{<\theta, \text{uncorr.}}$ is only necessary for consistency reasons: a joint log-normal PDF of $\delta_{m,T}$ and $\kappa_{<\theta}$ characterized by the above moments would be inconsistent with the variance $\langle \kappa_{<\theta}^2 \rangle$ derived from our power spectrum.)

Alternative model for $\langle \kappa_{<\theta} | \delta_{m,T} \rangle$:

As an alternative we compute $\langle \kappa_{<\theta} | \delta_{m,T} \rangle$ from the joint cumulant generating function of the variables $\delta_{m,T}$ and $\kappa_{<\theta}$. This function can also be modelled in a cylindrical collapse approach.

For model component (iii) we also investigate two different modeling approaches - one ansatz introducing 2 free parameters and one ansatz introducing 3 free parameters. We find that our simulated tracer catalogs are well described by the 2-parametric model. But - anticipating real galaxies to behave more complicated than simulated ones - we do not consider either of these models as our baseline model and instead apply both approaches to DES data in Gruen et al. (2018). We summarize both ansatzes here, but the interested reader is again referred to section 16.4 for details of each ansatz.

(iii) Model 1 for $P(N_T | \delta_{m,T})$:

In our fiducial model we introduce an auxiliary field $\delta_{g,T}$ such that our foreground galaxies are Poissonian tracers of that field, i.e.

$$P(N_T | \delta_{g,T}) = \frac{[\bar{N}(1 + \delta_{g,T})]^{N_T}}{N_T!} e^{-\bar{N}(1 + \delta_{g,T})} . \quad (16.20)$$

$\delta_{g,T}$ can be thought of as a smooth (shot noise free) galaxy density contrast. We then assume that $\delta_{g,T}$ and $\delta_{m,T}$ are joint log-normal random variables with

$$\langle \delta_{g,T}^2 \rangle = b^2 \langle \delta_{m,T}^2 \rangle, \quad \langle \delta_{g,T}^3 \rangle = b^3 \langle \delta_{m,T}^3 \rangle \quad (16.21)$$

and

$$\langle \delta_{g,T} \delta_{m,T} \rangle = br \langle \delta_{m,T}^2 \rangle. \quad (16.22)$$

The parameters b and r will be called *galaxy bias* and *galaxy stochasticity* and are free parameters of the model.

Model 2 for $P(N_T|\delta_{m,T})$:

As an alternative we assume $P(N_T|\delta_{m,T})$ to be a generalization of the Poisson distribution, that allows for

$$\langle N_T^2 | \delta_{m,T} \rangle \neq \langle N_T | \delta_{m,T} \rangle + \langle N_T | \delta_{m,T} \rangle^2, \quad (16.23)$$

i.e. for a shot-noise that is either enhanced or suppressed wrt. the Poisson case. The enhancement of shot-noise is also allowed to be a function of $\delta_{m,T}$ of (approximately) the form

$$\frac{\langle N_T^2 | \delta_{m,T} \rangle - \langle N_T | \delta_{m,T} \rangle^2}{\langle N_T | \delta_{m,T} \rangle} \approx \alpha_0 + \alpha_1 \delta_{m,T}. \quad (16.24)$$

This model introduces an alternative bias parameter \tilde{b} such that

$$\langle N_T | \delta_{m,T} \rangle = \bar{N} [1 + \tilde{b} \delta_{m,T}]. \quad (16.25)$$

For the model components (i) and (ii) and on the scales considered in this paper, the baseline and alternative modeling approaches yield almost indistinguishable predictions (cf. section 16.4). For component (iii) the modeling ansatzes 1 and 2 are not necessarily identical, because they introduce a different number of degrees of freedom. Figure 16.1 as well as all parameter contours shown in this paper are using the baseline model for components (i) and (ii) and ansatz 1 for component (iii). The predictions derived from different modeling approaches are however almost indistinguishable (cf. figure 16.5).

16.2.3 Data vector and forecasts on parameter constraints

Binning and scales

Throughout this paper, we assume that one sample of tracer galaxies is used to identify density quintiles and that the lensing profiles around these quintiles are measured with two source redshift bins (cf. redshift distributions in Figure 16.2). To identify the different density quintiles, we use a top-hat filter with fiducial smoothing radius of $\theta_T = 20'$.

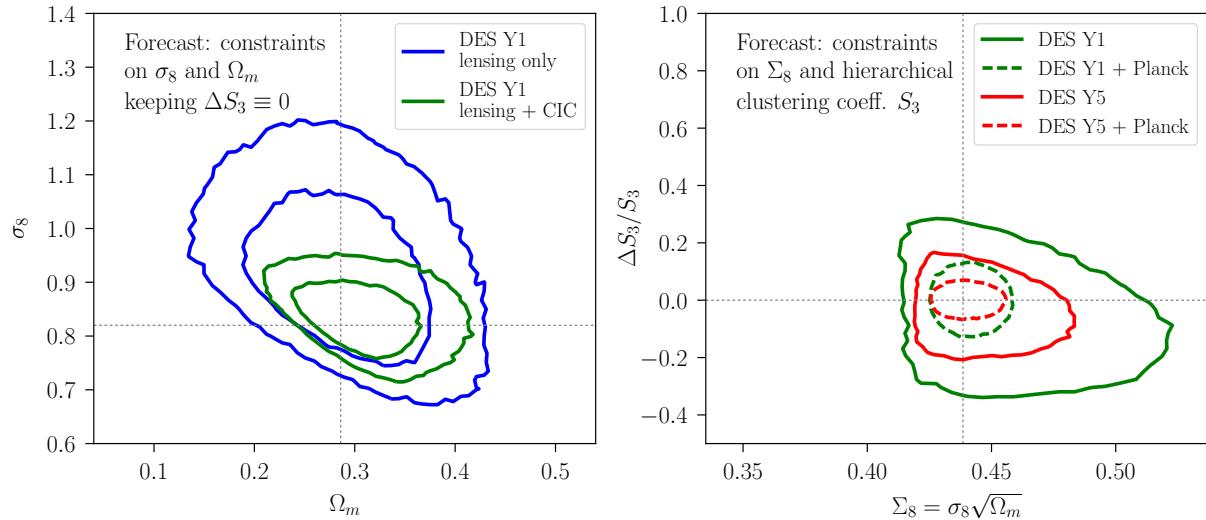


Figure 16.3: Left panel: Forecast of 1σ and 2σ constraints on Ω_m and σ_8 achievable with density split lensing and counts-in-cells in DES Y1 data. The constraints are marginalized over Ω_b , n_s , h_{100} , REDMAGIC galaxy bias b and galaxy-matter correlation coefficient r . For the parameters Ω_b , n_s , h_{100} we have assumed the same flat priors as used in the DES Y1 combined probes analysis presented in DES Collaboration et al. (2017). Right panel: $\Delta S_3 / S_3$ measures relative deviations from our fiducial perturbation theory prediction of the scaling coefficient $S_3 \equiv \frac{\langle \delta^3 \rangle}{\langle \delta^2 \rangle^2}$ (cf. equation 16.30 and section 16.4 for details). It can hence be thought of as the bispectrum amplitude. We show 1σ constraints on this parameter achievable with density split lensing and counts-in-cells in DES data alone (solid lines) and using additional information on cosmology from Planck (dashed lines, no lensing). The sharp cut-off of the contours for low values of Σ_8 is caused by the requirement that matter density and a shot-noise free galaxy density field must have a correlation coefficient $r \leq 1$. All likelihoods are centered around our fiducial cosmology, i.e. the parameters describing the Buzzard simulations.

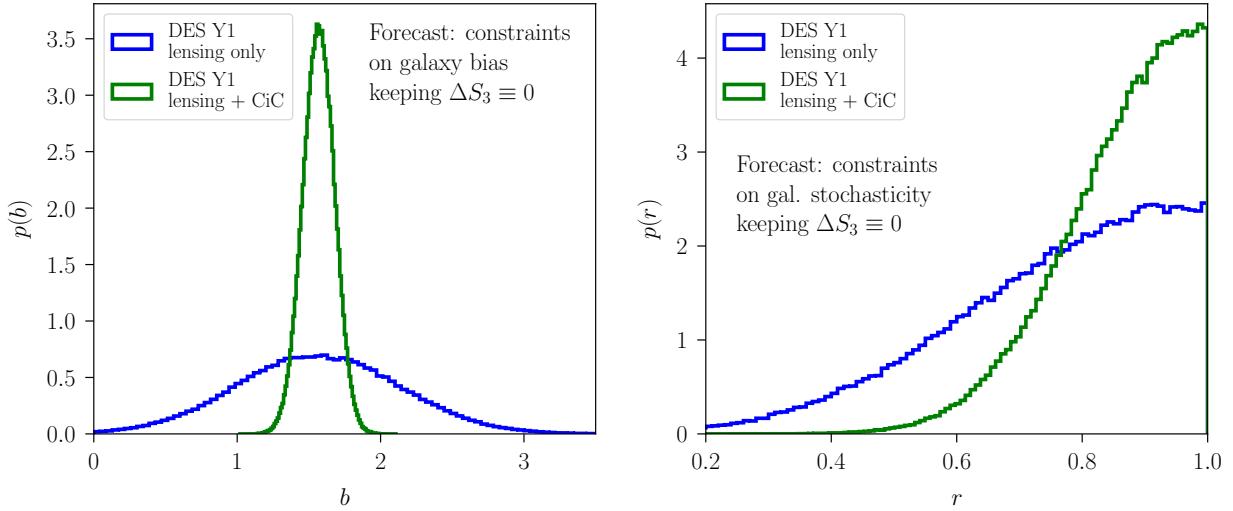


Figure 16.4: Forecast of posterior constraints on galaxy bias b and galaxy-matter correlation coefficient r achievable with density split lensing and counts-in-cells in DES Y1 data. The constraints are marginalized over Ω_m , σ_8 , Ω_b , n_s and h_{100} , where for the last three parameters we have assumed the same flat priors as used in the DES Y1 combined probes analysis presented in DES Collaboration et al. (2017).

We measure the density split lensing signal in 24 log-spaced angular bins with

$$5' < \theta < 600' . \quad (16.26)$$

But in the parameter forecasts and likelihood contours shown in the following, we exclude bins with scales $\leq \theta_T$. This reduces all lensing signals to 17 log-spaced angular bins with

$$20' \lesssim \theta < 600' . \quad (16.27)$$

The scales with $\theta < \theta_T$ are excluded from our analysis for a number of reasons: first, the signal-to-noise ratio of our lensing profiles drops quickly below θ_T (cf. Figure 16.1). Second, we chose our fiducial value of $\theta_T = 20'$ because we still trust the modeling of the density PDF described in section 16.4 on this scale, and we do not want smaller angular scales to contribute to our fiducial data vector. Third, the approximation made in equation 16.8 might fail at scales smaller than our aperture.

The sum of all 5 lensing signals will be very close to zero (though not exactly zero because of masking effects), so that they are not an independent set of observations. Hence, we only include the 2 most overdense and the 2 most underdense quintiles in our analysis, i.e. the cyan, blue, orange and red lines in Figure 16.1. For the same reason, we also use the normalized mean galaxy count only in four of the five quintiles of the counts-in-cells histogram to complement the lensing measurement. The total number of data points in

| parameter | fiducial value (in Buzzard) | prior in likelihood analysis | DES-Y1 constraints without $\Delta S_3/S_3$ | DES-Y1 constraints with $\Delta S_3/S_3$ | DES-Y1 constraints without $\Delta S_3/S_3$ (lensing only) |
|------------------|--------------------------------|---|--|---|--|
| σ_8 | 0.82 | [0.2, 1.6] | ± 0.05 | ± 0.10 | ± 0.11 |
| Ω_m | 0.286 | [0.1, 0.9] | ± 0.04 | ± 0.06 | ± 0.06 |
| Ω_b | 0.047 | [0.01, 0.07] | - | - | - |
| n_s | 0.96 | [0.7, 1.3] | - | - | - |
| h_{100} | 0.7 | [0.55, 0.91] | - | - | - |
| b | 1.618 | [0.8, 2.5] <i>(lensing only: [0.0, 4.5])</i> | ± 0.11 | ± 0.27 | ± 0.57 |
| r | 0.956 | [0.0, 1.0] | ± 0.10 | ± 0.11 | ± 0.18 |
| $\Delta S_3/S_3$ | 0.0 | [-1.0, 2.0] | - | ± 0.20 | - |

Table 16.1: Model parameters of the forecast described in section 16.2.3. The second column shows our fiducial values (the values describing the Buzzard simulations). The third column shows the parameter priors used to cut our prediction for the posterior distribution of best-fit parameters. The priors on Ω_b , n_s and h_{100} are informative and chosen to be the same as used by the DES Collaboration et al. (2017). The prior on r is needed for mathematical consistency. And the other priors only have to be introduced formally since we are approximating our analytic posterior by an MCMC. The 4th column shows the standard deviation of each parameter (as computed from the MCMC) after marginalization over all other parameters. The 5th column shows the same standard deviations but for the case where also $\Delta S_3/S_3$ is introduced as a free parameter of our model. In column 6, we again fix $\Delta S_3/S_3$ but assume that only the lensing part of the data vector is used. These forecasts can be compared to the actual errors we find in Gruen et al. (2018, their tables 2 and 3) which are close despite marginalizing over systematic uncertainties.

our data vector is thus

$$\begin{aligned}
 N_{\text{data}} &= N_{\text{lens}} + N_{\text{CiC}} \\
 &= [N_{\text{quant.}} - 1] \cdot N_{\text{source}} \cdot N_{\text{ang}} + [N_{\text{quant.}} - 1] \\
 &= 4 \cdot 2 \cdot 17 + 4 \\
 &= 140 .
 \end{aligned} \tag{16.28}$$

Model parameters and forecast on constraining power

We now investigate what constraints on model parameters can be expected from the above data vector when measured in DES-Y1 data. For this, we assume the optimistic case that component (iii) of our model is sufficiently described by two parameters, i.e. by modeling approach 1 in the previous section.² This means that our model is determined by the

²Gruen et al. (2018) will also consider modeling approach 2.

following 7 parameters:

- 1.) Ω_m : present total matter density in units the critical density of the universe,
- 2.) σ_8 : amplitude of present day density fluctuations in spheres of $8\text{Mpc}/h$ radius as predicted by the linear power spectrum,
- 3.) Ω_b : present density of baryonic matter in units of the critical density of the universe,
- 4.) n_s : the spectral index of the primordial power spectrum,
- 5.) h_{100} : the present day Hubble parameter in units of $100\text{Mpc}/(\text{km/s})$,
- 6.) b : linear bias of the tracers w.r.t. matter density (cf. equation 16.21),
- 7.) r : correlation coefficient between $\delta_{m,T}$ and $\delta_{g,T}$ (cf. equation 16.22).

We summarize these parameters and our fiducial values for them in Table 16.1. Throughout this paper, we assume the universe to be flat.

If π_α and π_β are any two of the above parameters then let $\hat{\pi}_{\alpha,\text{ML}}$ and $\hat{\pi}_{\beta,\text{ML}}$ be maximum likelihood estimates of these parameters based on a measurement of density split statistics. The covariance of $\hat{\pi}_{\alpha,\text{ML}}$ and $\hat{\pi}_{\beta,\text{ML}}$ can be estimated by

$$\text{Cov}[\hat{\pi}_{\alpha,\text{ML}}, \hat{\pi}_{\beta,\text{ML}}]^{-1} \approx \frac{\partial \mathbf{d}_{\text{th}}^T}{\partial \pi_\alpha} \cdot \mathbf{C}_d^{-1} \cdot \frac{\partial \mathbf{d}_{\text{th}}}{\partial \pi_\beta}, \quad (16.29)$$

where \mathbf{d}_{th} is our model of the density split data vector and \mathbf{C}_d is the covariance matrix of a measurement of this signal in DES-Y1. We will in the following use an estimate of \mathbf{C}_d from log-normal mocks and real DES Y1 shape noise (see section 16.3.2 for a brief summary of our covariance estimation and Gruen et al. 2018 for details). The parameter covariance computed with equation 16.29 can then be used to approximate the expected distribution of our best-fit parameter estimates with a multivariate Gaussian distribution.

Since the three parameters Ω_b , n_s and h_{100} are only poorly constrained by our data vector we are forced to assume prior knowledge on them. To do so, we cut the Gaussian posterior predicted from the parameter covariance with flat informative priors. These priors are chosen to be the same used by the DES Collaboration et al. (2017). For reasons of mathematical consistency we furthermore have to demand that $r \in [0, 1]$. These hard cuts of our originally Gaussian approximation to the posterior distribution of best-fit parameters make it difficult to marginalize over individual parameters. We hence numerically evaluate our analytic posterior with a Monte-Carlo Markov-Chain (MCMC). This chain is used in the following visualizations. Since we are using an MCMC to trace our analytic posterior, we have to formally define priors for all other model parameters. These are chosen to be flat and to extend well beyond the single-parameter standard deviations of the posterior. In the third column of table 16.1, we summarize the priors chosen for each model parameter.

The left panel of Figure 16.3 shows the 1σ and 2σ constraints achievable in the Ω_m - σ_8 plane. These contours are marginalized over the other model parameters, using the priors mentioned above. The blue contours assume that only the density split lensing signal has been used while the green contours allow for complementary information from the

tracer counts-in-cells histogram. In the fourth column of table 16.1 we show the standard deviation of each parameter as found in our approximation to the posterior (and assuming the full data vector, including lensing and counts-in-cell).

Density split statistics is complementary to an analysis based on 2-point statistics not just because it has a different dependence on the connection of galaxies and matter, but also since it is sensitive to higher order moments of the density field. We demonstrate this by introducing an additional degree of freedom in our model, described by an additional parameter:

8.) $\Delta S_3/S_3$: a factor multiplying all third order statistics in our predictions. The notation for this parameter is based on the ratio

$$S_3 \equiv \frac{\langle \delta^3 \rangle}{\langle \delta^2 \rangle^2} \quad (16.30)$$

which connects third and second moments of the density contrast and hence characterizes the amplitude of the density bispectrum (see section 16.4.1 for details). In our fiducial setup we compute it from leading order perturbation theory and $\Delta S_3/S_3$ hence describes a relative deviation from that result.

Within the Λ CDM model and at leading order in perturbation theory, the scaling between 2-point and 3-point statistics of the density field is almost independent of the cosmological parameters Ω_m and σ_8 (Bernardeau et al., 2002). Hence, a value of $\Delta S_3/S_3 \neq 0$ would allow for deviations from the leading order result that cannot be compensated by changing Ω_m or σ_8 . Such deviations could be caused non-standard physics of dark matter and dark energy that affect overdensities and underdensities differently (see e.g. Multamäki et al. (2003); Lue et al. (2004); though $f(R)$ modified gravity theories have been shown to largely preserve the Λ CDM scaling, cf. Jain & Zhang (2008); Borisov & Jain (2009)). Alternatively, they could indicate a break down the perturbative scaling relations due to highly non-linear evolution of the density field or any small scale Baryonic physics that do not follow the scaling relations of perturbation theory (cf. Bernardeau et al., 2002; Uhlemann et al., 2018a; Jeong & Komatsu, 2009).

In the right panel of Figure 16.3 we show how density split statistics including lensing and counts-in-cells can simultaneously constrain $\Delta S_3/S_3$ and the parameter

$$\Sigma_8 = \sigma_8 \sqrt{\Omega_m} , \quad (16.31)$$

even after marginalization over the other model parameters. We also project how these constraints will improve when moving to year 5 (Y5) data of DES or when adding cosmological information from the cosmic microwave background. For the latter we estimated the parameter covariance in a Planck chain³ and added this covariance as an additional Gaussian prior around our fiducial cosmology.

³plikHM_TT_lowTEB in COM_CosmoParams_base-plikHM_R2.00.tar.gz from the Planck legacy archive <https://pla.esac.esa.int/pla/>, no lensing, cf. Planck Collaboration et al. (2016)

With DES Y1 alone, we will be able to constrain the amplitude of third order statistics of the density field to about 20% accuracy (cf. last column of table 16.1). Combining DES Y5 and Planck, this improves to about 5%. And this is even underestimating the power of DES-Y5: to project our constraints onto year-5 we simply divided our covariance by a factor of 4 in order to account for the larger area of the final DES survey. But this does not take into account the fact that DES Y5 will also be deeper than DES-Y1, which reduces shape noise and opens up the possibility of analyzing a larger number of redshift bins.

16.3 Simulated data and covariance matrix

In this work we use two sets of simulated data: the Buzzard galaxy catalogs which are constructed from high-resolution N-body simulations (DeRose et al., 2018; Wechsler et al., 2018; MacCrann et al., 2018) and simulated random fields on the sky generated by the FLASK tool (Xavier et al., 2016). We briefly describe these data sets in the following sections.

16.3.1 Buzzard mock galaxy catalogs

Here we describe the key aspects of the Buzzard galaxy catalogs for the purposes of this work and refer the reader to more detailed descriptions elsewhere (DeRose et al., 2018; Wechsler et al., 2018; MacCrann et al., 2018). We use four independent realizations of a DES Y1-like survey in version Buzzard-v1.1. These catalogs were constructed from N-body simulations run using L-GADGET2 (Springel, 2005), a version of GADGET2 modified for memory efficiency. Second-order Lagrangian perturbation theory initial conditions (Crocce et al., 2006) were employed using 2LPTIC (Crocce et al., 2006), and light-cones were output on the fly. Each galaxy catalog is built from a set of three nested light-cones using progressively larger volume and lower resolution at higher redshifts. The force resolution in each box is 20, 35 and 53 kpc/h with the boundaries between the light-cones falling at redshifts $z = 0.34$ and $z = 0.9$.

The galaxy catalogs are constructed from the light-cones using the ADDGALS algorithm (Wechsler et al., 2018) which assigns galaxy luminosities and positions based on the relation between redshift, r-band absolute magnitude, and large-scale density, $p(\delta|M_r, z)$, found in a sub-halo abundance matching (SHAM) model (Conroy et al., 2006; Reddick et al., 2013, e.g.), in a high resolution N-body simulation. Spectral Energy Distributions (SEDs) are given to each simulated galaxy by finding a SDSS DR7 galaxy (Cooper et al., 2011) that has a close match in M_r and distance to its fifth nearest neighbor galaxy and assigning the SDSS galaxy's SED to the simulated galaxy. Galaxy sizes and ellipticities were assigned by drawing from distributions fit to high resolution SuprimeCam i -band data (Miyazaki et al., 2002). Observed magnitudes in $griz$ were generated by redshifting the SEDs to the observer frame and integrating over the DES passbands. Photometric errors were added using the DES Y1 Multi Object Fitting (MOF) depth maps.

The effects of weak gravitational lensing are calculated using the multiple-plane ray-tracing algorithm CALCLENS (Becker, 2013). The ray-tracing is done on a $n_{side} = 4096$ HEALPIX (Górski et al., 2005) grid yielding an effective angular resolution of $0.85'$. At each lens plane, the inverse magnification matrix of the ray closest to every galaxy is interpolated to the galaxy's 3D position and used to shear and magnify the galaxy.

With galaxy catalogs with magnitudes, sizes and lensing effects in hand, REDMAGiC and METACALIBRATION (Zuntz et al., 2017b; Sheldon & Huff, 2017; Huff & Mandelbaum, 2017) samples of galaxies are selected from the simulations in an effort to approximate the selections done in the data. In the case of REDMAGiC, the same algorithm which is used for selection in the data is applied to the simulations, resulting in a tracer galaxy catalog of equivalent volume density and at least comparable bias. For the METACALIBRATION sample, as we do not run full image simulations, we must make approximate cuts on signal to noise of the galaxies to create a facsimile of the source sample in the data. For in depth comparisons of these simulated samples to their data counterparts see DeRose et al. (2018).

For the density split analysis in this work, we use REDMAGiC high density tracer galaxies ($L > 0.5L_{star}$, $\rho = 10^{-3}\text{Mpc}^{-3}$ co-moving density) selected at a REDMAGiC photometric redshift estimate of $0.2 < z < 0.45$. For the source redshift split lensing signals, we bin source galaxies by the expectation value of their $p(z)$ as estimated with BPZ (Benítez, 2000; Hoyle et al., 2018) from the Buzzard mock photometry. Bin limits are chosen such that the true mean redshifts of the bins match the mean redshifts of the two highest redshift source samples defined in Hoyle et al. (2018).

16.3.2 Simulated density and convergence fields from FLASK and covariance matrix

For testing the numerical implementation of the model described in the following section and for estimating a covariance matrix of the density split lensing and counts-in-cells signals in the Buzzard simulations and the DES data, we use log-normal realizations of matter density and convergence fields. We summarize their properties here, with details given in appendix A of Gruen et al. (2018).

We generate these fields as all-sky HEALPIX maps of matter density and convergence using the FLASK software (Xavier et al., 2016). For the matter field, we choose the true redshift distribution of REDMAGiC galaxies in the Buzzard simulations, selected as described in subsection 16.3.1. The matter field is sampled by a tracer population with REDMAGiC density, bias of $b = 1.54$, and Poissonian noise, from which lines of sight of different density are identified by the same algorithm as in Buzzard or in the data. For the convergence fields, we choose the estimated redshift distributions of DES source galaxies in the two highest redshift bins of Hoyle et al. (2018). Log-normal parameters of the density and convergence fields are set by the perturbation theory formalism described in the following section.

960 of these realizations are used to estimate large-scale structure and shot noise contributions to the covariance matrix of the signals modeled herein. The contribution of

shape noise is estimated by measuring the lensing signals in actual DES Y1 data with 960 realizations of the METACALIBRATION shape catalog (Zuntz et al., 2017b) in which each galaxy ellipticity was rotated by a random angle.

16.4 Modeling details and comparison to simulations

In this section we present the approximations used to compute the ingredients (i), (ii) and (iii) of our model that are listed at the end of section 16.2.2. We also test the model ingredients (i) and (iii) directly against our N-body simulations.

Section 16.4.1 describes our model for the PDF of projected density contrast within the top-hat smoothing radius θ_T , $p(\delta_{m,T})$. Section 16.4.2 describes how we model the convergence profile $\langle \kappa_{<\theta} | \delta_{m,T} \rangle$ around apertures of fixed density contrast $\delta_{m,T}$. And in section 16.4.3 we describe our modeling of the probability $P(N_T | \delta_{m,T})$ of finding N_T tracer galaxies in an aperture with density contrast $\delta_{m,T}$.

Section 16.4.4 summarizes our fiducial model and the approximations used therein.

16.4.1 Projected density PDF

The computation of the PDF of the density field when smoothed by top-hat filters has e.g. been addressed by Bernardeau (1994), Bernardeau & Valageas (2000) and Valageas (2002a) (see also more recent developments in Bernardeau et al. (2014, 2015); Codis et al. (2016a); Codis et al. (2016b); Uhlemann et al. (2018a) which however do not yet enter our formalism). Bernardeau & Valageas (2000) demonstrated how to extend methods for the computation of the 3D density PDF to the weak lensing aperture mass which is a projected quantity. In the following we show how to modify their formalism in order to compute the line-of-sight density PDF in angular circles of radius θ_T . To do so, we have to consider the *cumulant generating function* (CGF) of the field $\delta_{m,T}(\hat{\mathbf{n}})$. The *moment generating function* (MGF) of $\delta_{m,T}(\hat{\mathbf{n}})$ is defined as

$$\psi(y) = \sum_k \frac{\langle \delta_{m,T}(\hat{\mathbf{n}})^k \rangle}{k!} y^k . \quad (16.32)$$

Due to the isotropy of the universe it does not depend on $\hat{\mathbf{n}}$. The CGF $\varphi(y)$ is given in terms of the MGF $\psi(y)$ as

$$\begin{aligned} \varphi(y) &= \ln \psi(y) \\ &\equiv \sum_k \frac{\langle \delta_{m,T}(\hat{\mathbf{n}})^k \rangle_c}{k!} y^k , \end{aligned} \quad (16.33)$$

where in the last line we have defined the *connected moments* or *cumulants* $\langle \delta_{m,T}(\hat{\mathbf{n}})^k \rangle_c$ of $\delta_{m,T}(\hat{\mathbf{n}})$. The CGF of $\delta_{m,T}(\hat{\mathbf{n}})$ is related to its PDF via

$$e^{\varphi(y)} = \int d\delta_{m,T} e^{y\delta_{m,T}} p(\delta_{m,T}) , \quad (16.34)$$

which is the Laplace transform of the PDF. Hence, if $\varphi(y)$ is a known analytic function, then $p(\delta_{m,T})$ can be computed by the inverse Laplace transform

$$p(\delta_{m,T}) = \int_{-\infty}^{\infty} \frac{dy}{2\pi} e^{-iy\delta_{m,T} + \varphi(iy)}. \quad (16.35)$$

This integral in the complex plain is most efficiently evaluated along the path of steepest descent (cf. Bernardeau et al. (2014), especially their appendix B). The cumulants of $\delta_{m,T}(\hat{\mathbf{n}})$ can be approximated as (cf. Bernardeau & Valageas (2000))

$$\langle \delta_{m,T}(\hat{\mathbf{n}})^k \rangle_c \approx \int dw \frac{\langle \{\delta_{\text{cy.},\theta w,L}(w\hat{\mathbf{n}},w)q_l(w)L\}^k \rangle_c}{L}. \quad (16.36)$$

Here q_l is the line-of-sight density of tracer galaxies defined in equation 16.2 and $\delta_{\text{cy.},R,L}$ is the average of $\delta_{m,3\text{D}}$ over a cylinder of length L and radius R , where L has to be chosen such that correlations of $\delta_{m,3\text{D}}$ over a distance L vanish for all practical purposes. Equation 16.36 employs a small angle approximation and a Limber-like approximation (Limber, 1953). This means it assumes that any n -point correlation function between density contrast at different redshifts z_i , $i = 1 \dots n$, varies much more quickly with the redshift differences Δz_{ij} than the projection kernel q_l . Comparing 16.33 and 16.36 we see that the CGF of $\delta_{m,T}$ can be computed in terms of the CGF of $\delta_{\text{cy.},R,L}$ as

$$\varphi(y) \approx \int dw \frac{\varphi_{\text{cy.},\theta w,L}(q_l(w)Ly,w)}{L}. \quad (16.37)$$

Hence, we have reduced the task of computing $p(\delta_{m,T})$ to the task of computing the connected moments of matter contrast in a long 3D cylinder.

To proceed we consider two different ansatzes. The first is to approximate $p(\delta_{m,T})$ by a log-normal distribution which is fixed by the first three connected moments of $\delta_{m,T}(\hat{\mathbf{n}})$. The second ansatz is to compute $\varphi_{\text{cy.},\theta w,L}(y)$ as a whole in a way similar to the one of Bernardeau (1994) and Valageas (2002a) for the matter contrast in a 3-dimensional sphere. Using 16.37 we can then attempt to solve the integral in 16.35 directly. We present details of both approaches in the following subsections.

Log-normal approximation for the PDF

Instead of computing the complete cumulant generating function of $\delta_{m,T}$ via equation 16.37 we start with an approach that only requires knowledge of the second and third cumulant, i.e. the moments $\langle \delta_{m,T}^2 \rangle_c$ and $\langle \delta_{m,T}^3 \rangle_c$ (implicitly we also assume $\langle \delta_{m,T} \rangle_c \equiv 0$ for the first cumulant). To do so, we approximate $\delta_{m,T}$ as a shifted log-normal random variable (Hilbert et al., 2011; Xavier et al., 2016). In this case the PDF of $\delta_{m,T}$ is given by

$$p(\delta_{m,T}) = \frac{\exp\left(-\frac{[\ln(\delta_{m,T}/\delta_0-1)+\sigma^2/2]^2}{2\sigma^2}\right)}{\sqrt{2\pi}\sigma(\delta_{m,T}-\delta_0)} \quad (16.38)$$

if $\delta_{m,T} > \delta_0$ and $p(\delta_{m,T}) = 0$ otherwise. The expectation value of this PDF is zero, as is appropriate. The variance and skewness of $\delta_{m,T}$ are given in terms of the parameters parameters δ_0 and σ by (Hilbert et al., 2011)

$$\begin{aligned}\langle \delta_{m,T}^2 \rangle_c &= \delta_0^2 (e^{\sigma^2} - 1) \\ \langle \delta_{m,T}^3 \rangle_c &= \frac{3}{\delta_0} \langle \delta_{m,T}^2 \rangle_c^2 + \frac{1}{\delta_0^3} \langle \delta_{m,T}^2 \rangle_c^3.\end{aligned}\quad (16.39)$$

The ansatz of a log-normal PDF has been shown to be consistent with early DES data by Clerkin et al. (2017). For the 3-dimensional density contrast $\delta_{m,3D}$ this can be reasonably motivated from theory by observing that at leading order in perturbation theory the skewness of $\delta_{m,3D}$ scales as

$$\langle \hat{\delta}_{m,3D}^3 \rangle_c \sim \langle \hat{\delta}_{m,3D}^2 \rangle_c^2 \quad (16.40)$$

with corrections of the order $\langle \hat{\delta}_{m,3D}^2 \rangle_c^3$. This is exactly the scaling obeyed by the log-normal distribution and choosing the parameter δ_0 appropriately allows one to exactly reproduce the scaling coefficients predicted by perturbation theory.

At least for a power law power spectrum, the same kind of scaling is observed also for 2-dimensional, projected versions of the density field. Hence, one of the ansatzes used in this paper to compute the PDF of $\delta_{m,T}$ is to derive its variance and skewness as described in appendix 16.B and fix δ_0 and σ by demanding that the PDF in 16.38 has the same second and third moments.

Tree level computation of $\varphi_{\text{cy},\theta w,L}(y, w)$ in the cylindrical collapse model

Let us first consider the cumulant generating function $\varphi_{3D}(y, \tau)$ of the 3-dimensional density contrast $\delta_{3D}(\mathbf{x}, \tau)$. To compute φ_{3D} at tree-level in perturbation theory it is sufficient to assume spherical symmetry around a particular point \mathbf{x} (see e.g. Valageas, 2002a). Doing so, we can directly compute $\delta_{3D}(\mathbf{x}, \tau)$ as a function of the linear density contrast $\delta_{3D,\text{lin.}}(\mathbf{x}, \tau)$ by means of the spherical collapse model (Fosalba & Gaztanaga, 1998; Valageas, 2002a), i.e.

$$\delta_{3D}(\mathbf{x}, \tau) = F[\delta_{3D,\text{lin.}}(\mathbf{x}, \tau), \tau] \quad (16.41)$$

where F is determined by one of the differential equations presented in appendix A. Hence, using the assumption that the linear density contrast has a Gaussian distribution with variance $\sigma_{3D,\text{lin.}}^2(\tau)$ we can express the cumulant generating function as (cf. equation 16.34)

$$\begin{aligned}& \exp \{ \varphi_{3D}(y, \tau) \} \\ &= \int d\delta_{3D} e^{y\delta_{3D}} p(\delta_{3D}, \tau) \\ &= \int \frac{d\delta_{3D,\text{lin.}}}{\sqrt{2\pi\sigma_{3D,\text{lin.}}^2(\tau)}} \exp \left(yF[\delta_{3D,\text{lin.}}, \tau] - \frac{\delta_{3D,\text{lin.}}^2}{2\sigma_{3D,\text{lin.}}^2(\tau)} \right),\end{aligned}\quad (16.42)$$

where in the last step we simply performed a change of variables from δ_{3D} to $\delta_{3D,\text{lin.}}$. We now employ Laplace's method, which states that a function $f(x)$ which strongly peaks around x_0 fulfills

$$\int dx e^{f(x)} \approx \sqrt{\frac{2\pi}{|f''(x_0)|}} e^{f(x_0)}. \quad (16.43)$$

This way we can approximate the last line of 16.42 as

$$\begin{aligned} e^{\varphi_{3D}(y, \tau)} &\approx \sqrt{\frac{1}{|1 - yF''[\delta^*, \tau] \sigma_{3D,\text{lin.}}^2(\tau)|}} \times \\ &\times \exp\left(yF[\delta^*, \tau] - \frac{\delta^{*2}}{2\sigma_{3D,\text{lin.}}^2(\tau)}\right), \end{aligned} \quad (16.44)$$

where δ^* is the linear density contrast that maximizes the exponent in 16.42 and $'$ denotes derivation wrt. δ . In the quasi-linear limit of $\sigma_{3D,\text{lin.}}^2 \rightarrow 0$ this gives

$$\varphi_{3D}(y, \tau) \approx yF[\delta^*, \tau] - \frac{\delta^{*2}}{2\sigma_{3D,\text{lin.}}^2(\tau)}, \quad (16.45)$$

where δ^* has to be determined by the implicit equation

$$\frac{\delta^*}{\sigma_{3D,\text{lin.}}^2(\tau)} = yF'[\delta^*, \tau]. \quad (16.46)$$

It should be noted that equations 16.45 and 16.46 reproduce *exactly* the tree-level results for the cumulant generating function (cf. Bernardeau, 1994; Valageas, 2002a; Bernardeau et al., 2002, 2015) ! As described in Bernardeau et al. (2002) the coefficients

$$S_n = \frac{\langle \delta^n \rangle_c}{\langle \delta^2 \rangle_c^{n-1}} \quad (16.47)$$

are given quite accurately by the lowest order of perturbation theory. Hence, using *halofit* (Smith et al., 2003; Takahashi et al., 2012) and an analytic transfer function (Eisenstein & Hu, 1998) to compute the non-linear matter power spectrum, we can compute the non-linear variance $\langle \delta^2 \rangle_{c,\text{non.lin}}$ and then rescale the leading order CGF to its non-linear version via

$$\varphi_{\text{non.lin}}(y, \tau) = \frac{\langle \delta^2 \rangle_{c,\text{lin}}}{\langle \delta^2 \rangle_{c,\text{non.lin}}} \varphi_{\text{lead}}\left(y \frac{\langle \delta^2 \rangle_{c,\text{non.lin}}}{\langle \delta^2 \rangle_{c,\text{lin}}}, \tau\right). \quad (16.48)$$

To perform the projection integral in equation 16.37 we need to know the cumulant generating function of the density contrast in a long cylinder of radius R and length $L \gg R, \delta_{\text{cy.}, R, L}$. Bernardeau (1994; see also Valageas, 2002a, and the other references above) have generalized equation 16.45 to the case of matter density in a spherical aperture. Their results can easily be transferred to cylindrical apertures, yielding

$$\varphi_{\text{cy.}, R, L}(y, \tau) \approx yF_{\text{cyl.}}[\delta^*, \tau] - \frac{\delta^{*2}}{2\sigma_{R\sqrt{1+F_{\text{cyl.}}[\delta^*]}, L, \text{lin.}}^2(\tau)}, \quad (16.49)$$

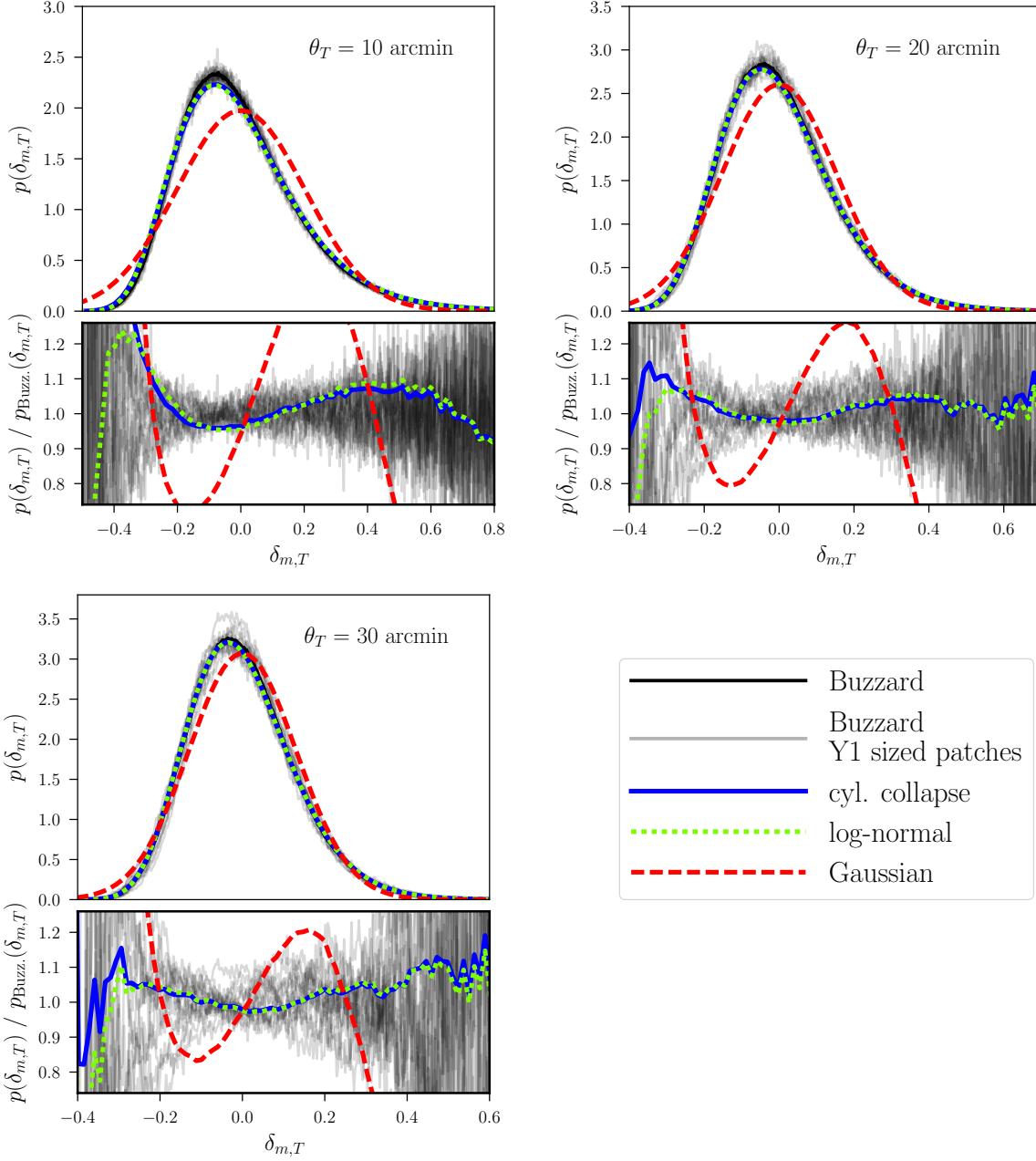


Figure 16.5: The PDF of projected density contrast in Buzzard compared to several models for various smoothing scales ($\delta_{m,T}$ is the projected density contrast δ_m smoothed by our top-hat radius θ_T). In the upper panel of each plot the black line shows a histogram of $\delta_{m,T}$ measured in an all-sky map from Buzzard. The gray lines show histograms measured in 14 DES year1 shaped patches of that all-sky map. The blue lines show the PDF predicted by our tree-level computation of the cumulant generating function, the green lines show the PT-motivated log-normal model and the red lines show a Gaussian PDF with the same variance as the other two models. The bottom panels of each plot are showing the ratio of each PDF to the one measured in the Buzzard all-sky. For all aperture radii our halofit power spectrum is predicting a standard deviation of $\delta_{m,T}$ that is $\gtrsim 2\%$ higher than what we find in Buzzard (cf. figure 16.6). For $\theta_T = 20'$ and $30'$ this is the dominant source of mismatch.

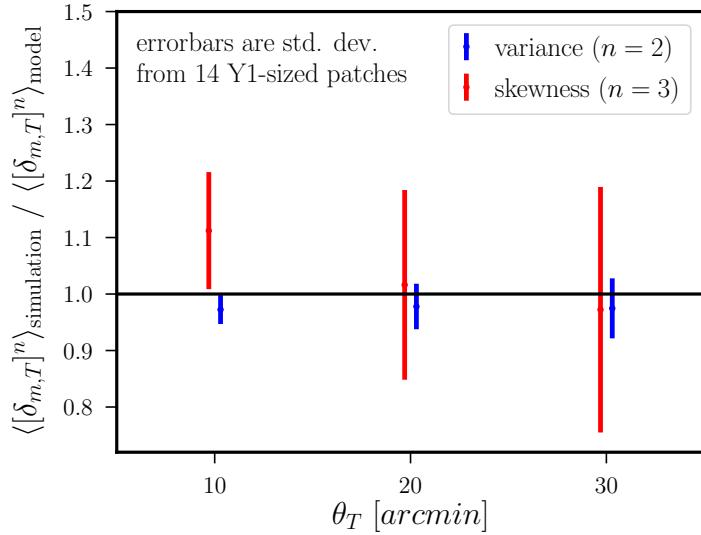


Figure 16.6: This figure shows the ratio of moments of $\delta_{m,T}$ measured in a Buzzard all-sky density map to the moments predicted by our model as a function of θ_T . The errorbars represent the standard deviations of the moments found in a set of 14 DES Y1-sized patches in the Buzzard map. At our fiducial radius $\theta_T = 20'$ we find a $\sim 2.4\%$ deviation between the variance of $\delta_{m,T}$ measured in Buzzard and that predicted by our model. This would correspond to a $\sim 1.2\%$ deviation in σ_8 .

where $\sigma_{R,L,\text{lin.}}^2$ is the variance of linear density contrast in a cylinder, $F_{\text{cyl.}}$ is now determined by cylindrical collapse and δ^* has to be determined by the implicit equation

$$\frac{1}{2} \frac{d}{d\delta^*} \frac{\delta^{*2}}{\sigma_{R\sqrt{1+F_{\text{cyl.}}[\delta^*]},L,\text{lin.}}^2(\tau)} = y F'[\delta^*, \tau]. \quad (16.50)$$

Using equation 16.48 we can again rescale this leading order result for the generating function to its non-linear counterpart.

The validity of equation 16.48 is ultimately limiting the accuracy of our model for the distribution of the density contrast inside the aperture radius θ_T , $p(\delta_{m,T})$. In figure 16.5 we compare the PDF measured in the Buzzard simulations to both the log-normal model and the full CGF computation of the PDF for aperture radii $\theta_T = 10', 20', 30'$. For both $\theta_T = 20'$ and $30'$ our model PDF's and the Buzzard simulations agree within DES-Y1 cosmic variance as can be seen in the residuals shown in the lower panels of each plot. Also, the difference between log-normal approximation and full CGF computation is completely negligible. To investigate the agreement of Buzzard and our models more quantitatively, we also compare the variance and skewness of each PDF. In figure 16.6 we show the ratio of these moments as found in Buzzard to our predictions. For $\theta_T = 10'$, the density field in Buzzard seems to have a significantly higher skewness than predicted by our model. We attribute this indeed to the failing of the Quasi-linear rescaling, Eq. 16.48. For the other radii the skewness values agree to within 2 – 3%. A similar relative agreement is

achieved for the variance of the distributions. At our fiducial radius $\theta_T = 20'$ the variances of Buzzard and our model differ by about 2.4%. This corresponds to a disagreement in the fluctuation amplitude σ_8 of about 1.2%.

For comparison, we also show a Gaussian model for the density PDF in figure 16.5, using the same variance as for the other PDF models. It can clearly be seen, that $p(\delta_{m,T})$ cannot be well described by a Gaussian distribution for any of the considered smoothing radii.

16.4.2 Convergence profile around lines-of-sight of fixed density contrast $\delta_{m,T}$

We now want to know the average lensing convergence $\kappa_{<\theta}$ inside a radius θ around a line-of-sight with a given value of $\delta_{m,T}$. By means of equation 16.6 this can be turned into a prediction of the density split lensing signal. We start by looking at the joint moment generating function of $\kappa_{<\theta}$ and $\delta_{m,T}$,

$$\psi_\theta(y, z) = \sum_{k,l} \frac{\langle \delta_{m,T}^k \kappa_{<\theta}^l \rangle}{k! l!} y^k z^l , \quad (16.51)$$

and their joint cumulant generating function defined by

$$\begin{aligned} \varphi_\theta(y, z) &= \ln \psi_\theta(y, z) \\ &\equiv \sum_{k,l} \frac{\langle \delta_{m,T}^k \kappa_{<\theta}^l \rangle_c}{k! l!} y^k z^l . \end{aligned} \quad (16.52)$$

Using a Limber-like approximation similar to the one employed in eq. 16.37, one can write $\varphi_\theta(y, z)$ as a line-of-sight projection of the CGF of matter density contrasts that are smoothed over concentric cylindrical apertures with length L and radii $R_1, R_2, \varphi_{\text{cyl}, R_1, R_2, L}(y, z, w)$:

$$\varphi_\theta(y, z) \approx \int dw \frac{\varphi_{\text{cyl}, \theta_T w, \theta w, L}(q_l(w) Ly, W_s(w) Lz, w)}{L} . \quad (16.53)$$

Here, $q_l(w)$ is again the line-of-sight density of lens galaxies and $W_s(w)$ is the lensing efficiency defined in Eq. 16.5.

The joint PDF of $\kappa_{<\theta}$ and $\delta_{m,T}$ is related to φ_θ via

$$p(\delta_{m,T} = s, \kappa_{<\theta} = r) = \int_{-\infty}^{\infty} \frac{dy dz}{(2\pi)^2} e^{-iys - izr + \varphi_\theta(iy, iz)} . \quad (16.54)$$

The expectation value of $\kappa_{<\theta}$ given a certain value of $\delta_{m,T}$ is then given by

$$\begin{aligned}
 \langle \kappa_{<\theta} | \delta_{m,T} = s \rangle &= \int dr p(\delta_{m,T} = s, \kappa_{<\theta} = r) \\
 &= \int_{-\infty}^{\infty} \frac{dy dz}{(2\pi)^2} e^{-iys + \varphi_{\theta}(iy, iz)} \int dr r e^{-izr} \\
 &= \int_{-\infty}^{\infty} \frac{dy dz}{2\pi} e^{-iys + \varphi_{\theta}(iy, iz)} i \frac{d}{dz} \delta_D(z) \\
 &= \int_{-\infty}^{\infty} \frac{dy}{2\pi} e^{-iys + \varphi_{m,T}(iy)} G_{\theta}(iy) ,
 \end{aligned} \tag{16.55}$$

with

$$\begin{aligned}
 G_{\theta}(y) &= \left. \frac{d}{dz} \varphi_{\theta}(y, z) \right|_{z=0} \\
 &= \sum_k \frac{\langle \delta_{m,T}^k \kappa_{<\theta} \rangle_c}{k!} y^k .
 \end{aligned} \tag{16.56}$$

Using equation 16.53 we can express $G_{\theta}(y)$ by the corresponding cylindrical, 3-dimensional quantity:

$$G_{\theta}(y) \approx \int dw W_s(w) G_{\text{cyl.}, \theta_T w, \theta w, L}(q_l(w) Ly, w) . \tag{16.57}$$

We again pursue two ansatzes for computing $G_{\theta}(y)$: one involving a log-normal model for the joint cumulants of $\kappa_{<\theta}$ and $\delta_{m,T}$ and one involving the model of cylindrical collapse to compute a leading order perturbation theory prediction for $G_{\text{cyl.}, \theta_T w, \theta w, L}(y, w)$. We are detailing these ansatzes in the following subsections.

Log-normal model for the joint cumulants of $\kappa_{<\theta}$ and $\delta_{m,T}$

From equation 16.56 one can see that only joint cumulants of the form

$$\langle \delta_{m,T}^k \kappa_{<\theta} \rangle_c \tag{16.58}$$

enter the convergence profile around a given density contrast. Hence, in a spirit similar to section 16.4.1 we only compute the leading order cumulants $\langle \delta_{m,T} \kappa_{<\theta} \rangle_c$ and $\langle \delta_{m,T}^2 \kappa_{<\theta} \rangle_c$ as described in appendix 16.B and use these moments to fix a joint log-normal PDF for $\kappa_{<\theta}$ and $\delta_{m,T}$. Note that this is indeed *not* assuming, that $\kappa_{<\theta}$ is a log-normal random variable. It rather assumes that

$$\kappa_{<\theta} = \kappa_{\text{log-normal}} + \kappa_{\text{uncorr.}} , \tag{16.59}$$

where $\kappa_{\text{log-normal}}$ is log-normal and $\kappa_{\text{uncorr.}}$ is an unspecified random variable that is uncorrelated with $\delta_{m,T}$. Only $\kappa_{\text{log-normal}}$ will actually contribute to the density split lensing signal and we can model the expectation value $\langle \kappa_{<\theta} | \delta_{m,T} = s \rangle$ by the following relation holding for two joint log-normal variables:

$$\frac{\langle \kappa_{<\theta} | \delta_{m,T} = s \rangle}{\kappa_0} = \exp \left(\frac{C(2 \log(1 + s/\delta_0) + V - C)}{2V} \right) - 1 \quad (16.60)$$

with

$$\begin{aligned} V &= \log \left(1.0 + \frac{\langle \delta_{m,T}^2 \rangle_c}{\delta_0^2} \right) \\ C &= \log \left(1.0 + \frac{\langle \delta_{m,T} \kappa_{<\theta} \rangle_c}{\delta_0 \kappa_0} \right) \\ \kappa_0 &= \frac{\langle \delta_{m,T} \kappa_{<\theta} \rangle_c^2 e^V}{\langle \delta_{m,T}^2 \kappa_{<\theta} \rangle_c - 2 \langle \delta_{m,T} \kappa_{<\theta} \rangle_c \langle \delta_{m,T}^2 \rangle_c / \delta_0} \end{aligned} \quad (16.61)$$

and δ_0 determined as described in section 16.4.1.

It should be noted that the log-normal parameter κ_0 which we use to approximate the contribution of $\kappa_{<\theta}$ to the lensing signal is dependent on the smoothing scale θ . This indicates even further, that we do indeed not approximate κ as a log-normal field.

Tree level computation of $G_{\text{cyl.}, R_1, R_2, L}(y, w)$ in the cylindrical collapse model

For convenience we will shorten the notation of section 16.4.1 by defining

$$\begin{aligned} G(y) &\equiv G_{\text{cyl.}, R_1, R_2, L}(y, w) \\ \varphi(y, z) &\equiv \varphi_{\text{cyl.}, R_1, R_2, L}(y, z, w) \\ F[\delta] &\equiv F_{\text{cyl.}}[\delta, \tau] \\ \sigma_R &\equiv \sigma_{R, L, \text{lin.}} \\ R_{i, L}(\delta) &\equiv R_i \sqrt{1 + F[\delta]}, \quad i = 1, 2. \end{aligned} \quad (16.62)$$

The joint cumulant generating function of density contrast in concentric cylinders is then (in complete analogy to equation 16.49; see also Bernardeau & Valageas, 2000, who present very similar calculations) given by

$$\varphi(y, z) \approx y F[\delta_1^*] + z F[\delta_2^*] - \frac{1}{2} \sum_{i,j} \delta_i^* \delta_j^* (\mathcal{C}^{-1})_{ij}, \quad (16.63)$$

where the elements of the matrix \mathcal{C} are given by $\mathcal{C}_{11} = \sigma_{R_{1,L}(\delta_1^*)}^2$, $\mathcal{C} = \sigma_{R_{2,L}(\delta_2^*)}^2$ and $\mathcal{C}_{12} = \mathcal{C}_{21}$ is the linear covariance of density contrasts in concentric cylinders of radii $R_{1,L}(\delta_1^*)$ and $R_{2,L}(\delta_2^*)$. This time the critical linear density contrasts δ_1^* and δ_2^* are given by the implicit

equations

$$\frac{1}{2} \frac{\partial}{\partial \delta_1^*} \sum_{i,j} \delta_i^* \delta_j^* (\mathcal{C}^{-1})_{ij} = y F'[\delta_1^*] \quad (16.64)$$

$$\frac{1}{2} \frac{\partial}{\partial \delta_2^*} \sum_{i,j} \delta_i^* \delta_j^* (\mathcal{C}^{-1})_{ij} = z F'[\delta_2^*] . \quad (16.65)$$

Note that these conditions force each δ_i^* to be a function of both y and z , i.e. $\delta_i^* = \delta_i^*(y, z)$.

To predict the convergence profile around apertures of a given density contrast $\delta_{m,T}$ by means of equations 16.55 and 16.57 we are interested in computing the function

$$G(y) = \left. \frac{\partial}{\partial z} \varphi(y, z) \right|_{z=0} . \quad (16.66)$$

Using the conditions 16.64 and 16.65 one can see right away that

$$G(y) = F[\delta_2^*(y, 0)] . \quad (16.67)$$

Furthermore, for $z = 0$ equations 16.64 and 16.65 can be simplified to

$$\frac{1}{2} \frac{d}{d \delta_1^*} \frac{\delta_1^{*2}}{\mathcal{C}_{11}} = y F'[\delta_1^*] \quad (16.68)$$

$$\delta_2^* = \frac{\mathcal{C}_{12}}{\mathcal{C}_{11}} \delta_1^* . \quad (16.69)$$

This way we obtain a solution for $G(y)$ at leading order in perturbation theory. In appendix 16.B we argue that for $R_2 \geq R_1$ the cumulants $\langle \delta_{R_1}^k \delta_{R_2} \rangle_c$ approximately follow the scaling relation

$$\langle \delta_{R_1}^k \delta_{R_2} \rangle_c \sim \langle \delta_{R_1} \delta_{R_2} \rangle_c \langle \delta_{R_1}^2 \rangle_c^{k-1} . \quad (16.70)$$

This can be used to correct the tree-level approximation of $G(y)$ for the non-linear evolution of the power spectrum. To do so, we first determine the proportionality factors of the relation 16.70 at leading order by fitting a polynomial in y to the function $G(y)$ and extracting the cumulants $\langle \delta_{R_1}^k \delta_{R_2} \rangle_c$ from the polynomial coefficients. In practice, we do this with a polynomial of degree 10, but already a polynomial of degree 5 gives almost identical results⁴. Then, we use the revised halofit of Takahashi et al. (2012) to compute a late-time version of the right-hand-side of 16.70. This, together with the tree-level proportionality factors determined before, yields a non-linear approximation of the polynomial coefficients representing $G(y)$. We use those to re-compute $G(y)$ and then carry out the projection integral in equation 16.57.

Our rescaling of the coefficient corresponding to the cumulant $\langle \delta_{R_1}^2 \delta_{R_2} \rangle_c$ is in fact more complicated than described here, cf. appendix 16.B.5. But we find our prediction of the density split lensing signal to be insensitive to the details of the rescaling procedure.

⁴The coefficients of linear and quadratic order in y are always obtained from the exact perturbation theory computation of appendix 16.B.

16.4.3 Shot-noise, stochasticity and Counts-in-Cells

We now want to model the conditional probability $P(N_T|\delta_{m,T})$ of finding N_T galaxies in an angular radius of θ_T , when the projected density contrast in that radius is $\delta_{m,T}$. This is the third ingredient of the framework described in section 16.2 and completes our modeling of the density split lensing signal as well as the counts-in-cells histogram.

To analyze the relation of N_T and $\delta_{m,T}$ in a systematic way, let us introduce the auxiliary field $\delta_{g,T}$. We assume that $\delta_{g,T}(\hat{\mathbf{n}})$ is a smooth field in the sky and that N_T is a Poissonian tracer of this field. This means we will assume that

$$P(N_T = N|\delta_{g,T}) = \frac{[\bar{N}(1 + \delta_{g,T})]^N}{N!} e^{-\bar{N}(1 + \delta_{g,T})}, \quad (16.71)$$

where $\bar{N} \equiv \langle N_T \rangle$. A consequence of this assumption is that the expectation value of N_T for fixed $\delta_{g,T}$ is given by

$$\langle N_T | \delta_{g,T} \rangle = \bar{N}(1 + \delta_{g,T}) \quad (16.72)$$

and that the variance of N_T for fixed $\delta_{g,T}$ fulfills

$$\frac{\text{Var}[N_T | \delta_{g,T}]}{\langle N_T | \delta_{g,T} \rangle} = 1. \quad (16.73)$$

To connect the galaxy field to the lensing convergence we however need to know the relation between N_T and $\delta_{m,T}$. Assuming a generic joint PDF of $\delta_{m,T}$ and $\delta_{g,T}$ we can write the expectation values of N_T for fixed $\delta_{m,T}$ as

$$\begin{aligned} \langle N_T | \delta_{m,T} \rangle &= \int d\delta_{g,T} p(\delta_{g,T} | \delta_{m,T}) \langle N_T | \delta_{g,T} \rangle \\ &= \bar{N}(1 + \langle \delta_{g,T} | \delta_{m,T} \rangle). \end{aligned} \quad (16.74)$$

Also, it can be shown that the variance of N_T for a fixed value of $\delta_{m,T}$ is given by

$$\text{Var}[N_T | \delta_{m,T}] = \langle N_T | \delta_{m,T} \rangle + \bar{N}^2 \text{Var}[\delta_{g,T} | \delta_{m,T}]. \quad (16.75)$$

From equation 16.75 we can see that the distribution of N_T given $\delta_{m,T}$ can only be a Poisson distribution if $\text{Var}[\delta_{g,T} | \delta_{m,T}] \equiv 0$. This is the simplest model for the connection of N_T and $\delta_{m,T}$ that we test in this work and in Gruen et al. (2018). If $\text{Var}[\delta_{g,T} | \delta_{m,T}] \neq 0$ we will say that there is a stochasticity between the galaxy field and the matter density field, and we cannot assume a Poisson distribution for $P(N_T | \delta_{m,T})$. We note that “stochasticity” in this context could arise from a nonlinear biasing relationship between $\delta_{g,T}$ and $\delta_{m,T}$, including, e.g., a dependence on higher powers of $\delta_{m,T}$ or effects from halo exclusion (Baldauf et al., 2013), or from physical stochasticity in galaxy formation.

We explore two ways to account for a possible stochasticity (see also Dekel & Lahav (1999), who have discussed similar concepts). In our first approach we introduce a free parameter to our model - a Pearson correlation coefficient $r \neq 1$ between the random fields

$\delta_{g,T}$ and $\delta_{m,T}$. Within our log-normal framework we show that this automatically leads to a $\delta_{m,T}$ -dependence of the ratio in Eq. 16.77. We explain the details of this in section 16.4.3.

In our second approach we employ a generalized Poisson distribution for $P(N_T|\delta_{m,T})$ that allows for

$$\frac{\text{Var}[N_T|\delta_{m,T}]}{\langle N_T|\delta_{m,T} \rangle} \neq 1. \quad (16.76)$$

In this approach we introduce 2 parameters, α_0 and α_1 , to our model such that

$$\frac{\text{Var}[N_T|\delta_{m,T}]}{\langle N_T|\delta_{m,T} \rangle} \approx \alpha_0 + \alpha_1 \delta_{m,T}. \quad (16.77)$$

The details of this are explained in section 16.4.3.

Both of our approaches match our simulated data equally well (cf. Figure 16.8). This means that, for the galaxies in these realizations, the model based on the correlation coefficient r is a sufficient description. It will thus be the fiducial model in this paper, used in all figures unless otherwise noted. In Gruen et al. (2018) we will nevertheless apply both this and the two-parametric model to account for the possibility that the shot-noise of real galaxies behaves in a more complicated way than that of our simulated galaxies.

Shot-noise model 1: correlation $r \neq 1$ between galaxy density and matter density

In our fiducial model of $P(N_T|\delta_{m,T})$ we approximate the joint distribution of both $\delta_{m,T}$ and $\delta_{g,T}$ with a joint log-normal distribution (cf. Eq. 16.38 and Hilbert et al. (2011) for properties of joint log-normal distributions). The joint PDF of two log-normal random variables is characterized by 5 parameters, e.g. by the variance and skewness of each variable and the covariance between the two variables.

In our case, we compute the variance and skewness of $\delta_{m,T}$ as described in section 16.4.1 and set the variance and skewness of $\delta_{g,T}$ to

$$\begin{aligned} \langle \delta_{g,T}^2 \rangle_c &= b^2 \langle \delta_{m,T}^2 \rangle_c \\ \langle \delta_{g,T}^3 \rangle_c &= b^3 \langle \delta_{m,T}^3 \rangle_c \end{aligned} \quad (16.78)$$

where the *galaxy bias* b is a free parameter. The covariance of $\delta_{m,T}$ and $\delta_{g,T}$ is parametrized by their correlation coefficient

$$\begin{aligned} r &= \frac{\langle \delta_{g,T} \delta_{m,T} \rangle_c}{\sqrt{\langle \delta_{g,T}^2 \rangle_c \langle \delta_{m,T}^2 \rangle_c}} \\ &= \frac{\langle \delta_{g,T} \delta_{m,T} \rangle_c}{b \langle \delta_{m,T}^2 \rangle_c}, \end{aligned} \quad (16.79)$$

i.e.

$$\langle \delta_{g,T} \delta_{m,T} \rangle_c = rb \langle \delta_{m,T}^2 \rangle_c. \quad (16.80)$$

The log-normal model for the joint PDF of $\delta_{m,T}$ and $\delta_{g,T}$ now allows us to compute the variance of galaxy counts as a function of $\delta_{m,T}$ and more generally to compute $P(N_T|\delta_{m,T})$. We present the necessary derivations in detail in appendix 16.D.

In our data analysis we consider b and r as free parameters. The only restrictions we impose on them are

$$0 < b, \quad 0 \leq r \leq 1. \quad (16.81)$$

To test how accurately this model describes the behavior of our mock REDMAGIC catalogs based on the Buzzard N-body simulations we nevertheless want to determine what values of b and r are underlying our simulations. To this end, we generate HEALPIX maps of $\delta_{m,T}$ with different top-hat aperture radii $\theta_T = 10, 20, 30$ arcmin, based on particle count maps at resolution $N_{\text{side}} = 8192$ in slices of co-moving $50h^{-1}\text{Mpc}$ thickness. We co-add these maps to reproduce a redshift range close to that of our fiducial analysis, $z = 0.2100 \dots 0.4453$. We then select REDMAGIC galaxies with true redshift in this range and determine their counts around the same HEALPIX pixel centers and within the same aperture radii. The REDMAGIC mock catalogs have a complex mask similar to that of real DES data, which adds complication because the fraction of masked area in each aperture must be equal in order to meaningfully sort lines of sight by galaxy count. To this end, we convert all counts to a masking fraction of 20 per-cent of area within the aperture radius using the stochastic method of Gruen et al. (2018, their section 2.1). This leaves us with simulated 2D maps of $\delta_{m,T}$ and N_T within a DES-Y1 shaped mask.

We can then measure the variances of these maps, $\text{Var}(\delta_{m,T})$ and $\text{Var}(N_T)$, as well as their covariance $\text{Cov}(N_T, \delta_{m,T})$. These fulfill the relations

$$\text{Var}(N_T) = \bar{N} + \bar{N}^2 b^2 \text{Var}(\delta_{m,T}). \quad (16.82)$$

and

$$\text{Cov}(N_T, \delta_{m,T}) = \bar{N} b r \text{Var}(\delta_{m,T}) \quad (16.83)$$

which then fixes b and r . The values determined in this way are shown in table 16.2.

We now need to check whether these values for b and r together with our assumption of a log-normal PDF for $\delta_{m,T}$ and $\delta_{g,T}$ describe the properties of our tracer galaxies well. Using our simulated maps of $\delta_{m,T}$ and N_T we can measure the expectation value

$$\langle \delta_{g,T} | \delta_{m,T} \rangle = \frac{\langle N_T | \delta_{m,T} \rangle}{\bar{N}} - 1 \quad (16.84)$$

as a function of $\delta_{m,T}$. Within the log-normal model (cf. appendix 16.D for the relevant formulas) this is very well approximated by

$$\langle \delta_{g,T} | \delta_{m,T} \rangle \approx r b \delta_{m,T} \quad (16.85)$$

which becomes exact for Gaussian random variables. In Figure 16.7 we compare measurements of $\langle \delta_{g,T} | \delta_{m,T} \rangle$ with the different smoothing radii $\theta_T = 10, 20, 30$ arcmin to the prediction of the log-normal model. We find that in our simulations $\langle \delta_{g,T} | \delta_{m,T} \rangle$ is consistent

| Smoothing Scale [arcmin] | b | r |
|-----------------------------|-------------------|-------------------|
| 10 | 1.644 ± 0.008 | 0.938 ± 0.001 |
| 20 | 1.618 ± 0.008 | 0.956 ± 0.001 |
| 30 | 1.603 ± 0.008 | 0.961 ± 0.001 |

Table 16.2: Best-fit values galaxy bias and correlation coefficients of our simulated tracer galaxies within the model presented in section 16.4.3. Error bars are estimated from a jackknife approach.

| Smoothing Scale [arcmin] | \tilde{b} | α_0 | α_1 |
|-----------------------------|------------------|------------------|------------------|
| 10 | 1.54 ± 0.001 | 1.15 ± 0.001 | 0.22 ± 0.003 |
| 20 | 1.54 ± 0.002 | 1.26 ± 0.002 | 0.29 ± 0.010 |
| 30 | 1.54 ± 0.002 | 1.39 ± 0.003 | 0.45 ± 0.020 |

Table 16.3: Best-fit values galaxy bias and shot-noise parameters of our simulated tracer galaxies within the model presented in section 16.4.3. Error bars are again estimated from a jackknife approach.

with a linear relation in $\delta_{m,T}$. Interestingly, the scale dependence of b and r we find in table 16.2 almost perfectly cancels to give a scale independent proportionality coefficient

$$rb \approx 1.54 . \quad (16.86)$$

Next, we also measure the variance of galaxy counts N_T as a function of $\delta_{m,T}$ in our simulated maps and compare to the prediction of the log-normal model. In Figure 16.8 we indeed find that

$$\frac{\text{Var}[N_T|\delta_{m,T}]}{\langle N_T|\delta_{m,T} \rangle} \neq 1 \quad (16.87)$$

and that the $\delta_{m,T}$ -dependence is very well described by the log-normal model and the values of b and r we determined before. Finally, in figure 16.9 we show the residuals between our baseline prediction of the counts-in-cells histogram with $\theta_T = 20\text{arcmin}^5$ and the average of measurements in 4 Buzzard realizations of DES Y1 data (cf. figure 16.1). The residuals are well contained within DES Y1 errorbars.

Shot-noise model 2: Parametric model for super-Poissonianity

Our model of shot-noise based on galaxy bias b and galaxy matter correlation coefficient r describes our simulated tracer catalog well. But it contains the arbitrary assumption that both the variance and the skewness of $\delta_{m,T}$ and $\delta_{g,T}$ are related through the bias parameter b (cf. equation 16.78). To account for the possibility that real galaxies might

⁵This is the smoothing radius used in our data analysis (Gruen et al., 2018).

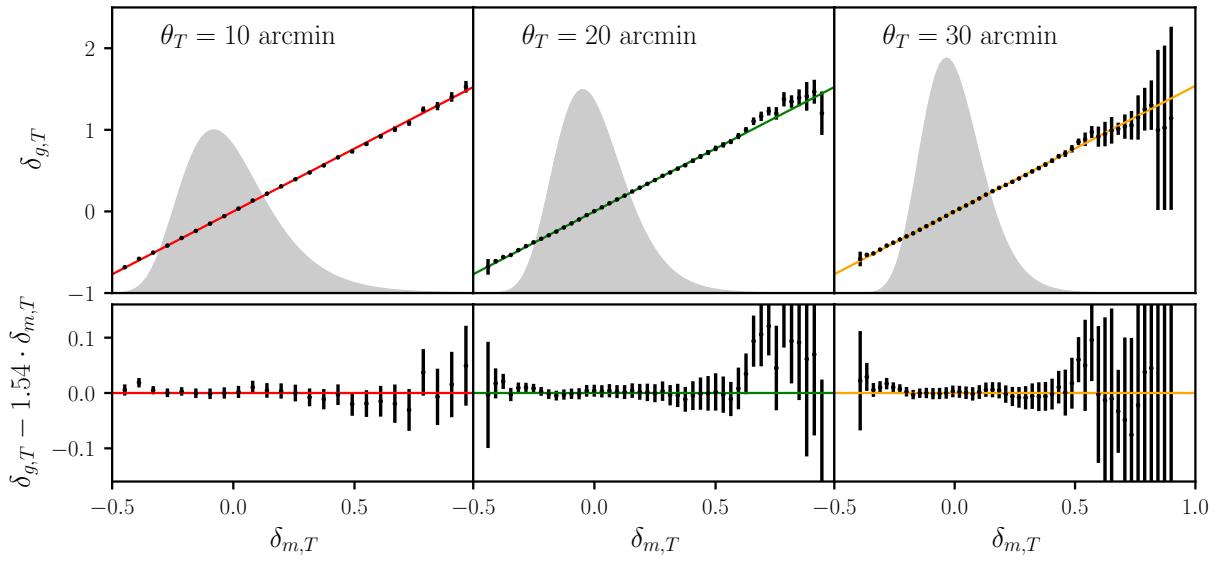


Figure 16.7: Average galaxy overdensity $\delta_{g,T}$ as a function of matter overdensity, $\delta_{m,T}$, for our simulated maps at different smoothing scales: 10 arcmin [left], 20 arcmin [middle] and 30 arcmin [right]. Solid lines show a linear bias model with the bias parameters obtained from maximizing the likelihood in equation 16.94 and the residual between the two are shown in the bottom panels. Note that the coefficient of linearity found with 16.94 (≈ 1.54) is almost identical to the value of the product $b \cdot r$ determined with equation 16.85. To indicate the range of $\delta_{m,T}$ that is relevant to our computation, we also show the density PDFs of figure 16.5 as shaded regions. The errorbars were derived from a jackknife approach.

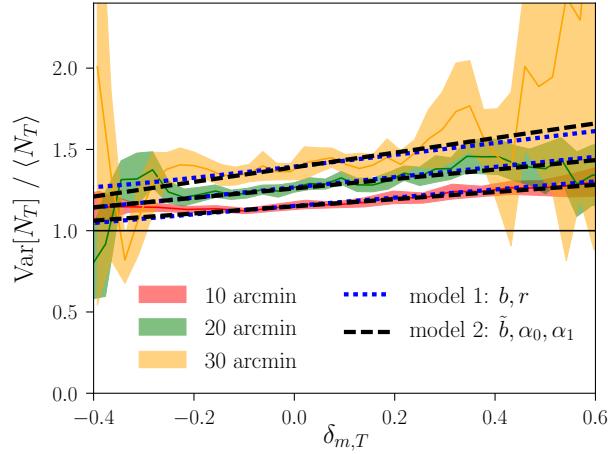


Figure 16.8: Ratio of the variance in the galaxy count distribution to average galaxy counts in our simulated maps as a function of matter overdensity. Differently colored solid lines show the result for each smoothing scale. The dashed black lines show the predictions of the 2-parametric shot-noise model described in section 16.4.3. The blue dotted lines show the corresponding prediction of the alternative, 1-parametric model described in section 16.4.3. The horizontal solid line shows the expectation if shot noise was purely Poissonian. The colored regions show the 95% confidence limits derived from jackknife re-sampling.

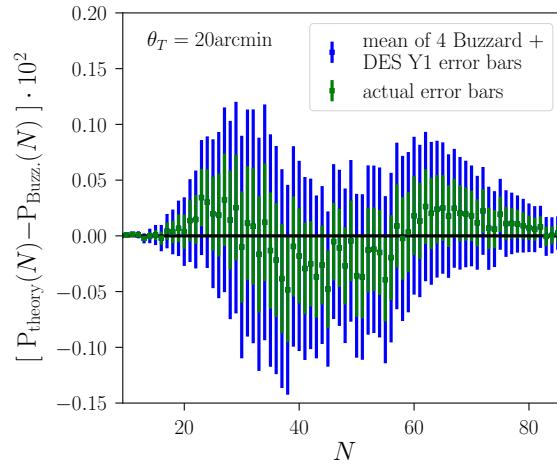


Figure 16.9: For the fiducial smoothing radius of our data analysis presented in Gruen et al. (2018), $\theta_T = 20\text{arcmin}$, we show the residuals between our baseline prediction of the counts-in-cells histogram and the average of measurements in 4 Buzzard realizations of DES Y1 data. Blue error bars represent the uncertainties we expect for DES Y1 while green error bars show the actual uncertainties of the mean measurement from our simulations.

behave in a more complicated way, we also consider a more flexible model of the conditional distribution $P(N_T|\delta_{m,T})$.

Gruen et al. (2016) assumed that there is no stochasticity in the relation of $\delta_{m,T}$ and $\delta_{g,T}$ and that galaxies trace the matter density with a linear bias and Poissonian shot noise. This means they set

$$P(N_T = N|\delta_{m,T} = s) = \exp \left(N \ln[\bar{N}(1 + \tilde{b}s)] - [\bar{N}(1 + \tilde{b}s)] - \ln \Gamma(N + 1) \right), \quad (16.88)$$

where \tilde{b} is the galaxy bias, and where we now use a generalizable definition of the Poisson distribution based on the Gamma function Γ , for reasons that will appear later. The galaxy bias \tilde{b} defined this way is not identical to the definition in our fiducial model. We now rather have

$$\delta_{g,T} \equiv \tilde{b} \delta_{m,T}. \quad (16.89)$$

We test this in our simulated maps of N_T and $\delta_{m,T}$ by fitting a linear relation to the mean smoothed galaxy contrast as a function of dark matter contrast that was shown in Figure 16.7. We indeed find that this linear biasing model describes the simulations very well and that $\tilde{b} \approx br$ as expected from our arguments in section 16.4.3.

The model used by Gruen et al. (2016) however predicts that

$$\frac{\text{Var}[N_T|\delta_{m,T}]}{\langle N_T|\delta_{m,T} \rangle} \equiv 1 \quad (16.90)$$

which is not what we find in Figure 16.8. To account for the deviations we observe from pure Poissonian shot-noise, we hence model the distribution of N_T given $\delta_{m,T}$ as

$$P(N_T = N|\delta_{m,T} = s) \sim \mathcal{N} \times \exp \left\{ \frac{N}{\alpha} \ln \left[\frac{\bar{N}_T}{\alpha} (1 + bs) \right] - \ln \Gamma \left(\frac{N + 1}{\alpha} \right) - \frac{\bar{N}_T}{\alpha} (1 + bs) \right\}, \quad (16.91)$$

where the parameter $\alpha > 0$ generalizes the distribution to one where groups of α galaxies appear with Poissonian noise and where the normalization coefficient \mathcal{N} is needed to ensure that $\int P(N_T = N) dN = 1$. We find \mathcal{N} to be very close to α^{-1} and identical to α^{-1} in the case where α is an integer value.

To account for the observed increase of super-Poissonianity with density, we allow α to depend on $\delta_{m,T}$,

$$\alpha(\delta_{m,T}) = \alpha_0 + \alpha_1 \times \delta_{m,T}. \quad (16.92)$$

This indeed leads to a $\delta_{m,T}$ -dependence of the variance of galaxy counts that is close to the relation mentioned in Eq. 16.77. In our analysis we treat α_0 and α_1 as free parameters

within the ranges

$$\begin{aligned}\alpha_0 &\in [0.1, 3.0] \\ \alpha_1 &\in [-1.0, 4.0].\end{aligned}\quad (16.93)$$

In principle we could allow any value $\alpha_0 > 0$ but we choose the boundary $0.1 < \alpha_0$ because it is numerically difficult (and slow) to predict the CiC histogram for values close to $\alpha_0 = 0.0$. The other boundaries roughly enclose the 2σ region of the posterior distribution of α_0 and α_1 we infer with DES Y1 like errors around the mean signal measured in Buzzard (after marginalizing over our other model parameters, cf. appendix 16.E). Also, the constraints on α_0 and α_1 we derive in Gruen et al. (2018) on DES data are well contained within our prior distributions.

Nevertheless, these priors must be considered mildly informative. We expect that even stronger priors can be motivated theoretically. Baldauf et al. (2013) find that for their most massive halos shot-noise is reduced wrt. Poisson expectation by a factor of ≈ 2 , indicating that $\alpha_0 \gtrsim 0.5$, while for halo masses comparable to REDMAGiC halo masses (cf. Clampitt et al. (2017)) they find shot-noise to be close to Poissonian. Also, there is evidence that the fraction of red galaxies that are satellites (resp. the fraction of satellite galaxies that are red) increases with environment density (see e.g. Mandelbaum et al. (2006); Peng et al. (2012)). According to Baldauf et al. (2013) this will cause an increase of galaxy stochasticity with environment density, corresponding to $\alpha_1 > 0.0$. We intend to investigate implications of models for halo occupation distributions (HOD) on our shot-noise parameterizations in future studies (see e.g. the work by Cacciato et al. (2012); Dvornik et al. (2018) on connecting HOD models and parametric models of galaxy bias and stochasticity).

To compare this parametric shot-noise model to our simulations, we are nevertheless interested in the particular value of α_0 and α_1 that describe these simulations. From the tuples of $(N_T, \delta_{m,T})_i$ measured in our simulated maps, we can constrain bias and the $\alpha_{0/1}$ parameters with a likelihood L that is simply the product of the probabilities of the individual tuples from Equation 16.91,

$$\begin{aligned}\ln L = \sum_i & \left[N_i / \alpha(\delta_{m,T}^i) \right] \ln \left[(\bar{N}_i / \alpha(\delta_{m,T}^i))(1 + b \times \delta_{m,T}^i) \right] \\ & - \left[\bar{N}_i / \alpha(\delta_{m,T}^i) \right] (1 + b \times \delta_{m,T}^i) \\ & - \ln G \left[N_i / \alpha(\delta_{m,T}^i) + 1 \right] - \ln \alpha(\delta_{m,T}^i).\end{aligned}\quad (16.94)$$

Because the tuples have correlated counts and densities, this is not an exact expression for the likelihood of our measurements, but it should be sufficient to obtain reasonable best fit values for b , α_0 and α_1 . We determine the uncertainties of these best-fit values by finding the maximum of Equation 16.94 on jackknife re-samplings of the simulations. The resulting parameter values are shown in Table 16.3 and displayed in Figure 16.7 and Figure 16.8. We find that our simulated REDMAGiC galaxies are indeed well described as linearly biased tracers of the density field with a small, but significant, scale and density dependent super-Poissonian shot-noise.

16.4.4 Summary of fiducial model and approximations therein

For each ingredient (i) to (iii) of the framework described in section 16.2 we have introduced at least two different modeling ansatzes. We want to once more describe our baseline model built from these ansatzes (cf. also section 16.2.2). This is the model we consider in section 16.5 and which we use in the data analysis presented in Gruen et al. (2018).

- (i) $p(\delta_{m,T})$: We find that the log-normal model (section 16.4.1) and our model based in cylindrical collapse (section 16.4.1) describe the PDF of projected density contrast equally well. The computations based on the log-normal model are however significantly faster. Hence in our fiducial analysis we employ the log-normal model.
- (ii) $\langle \kappa_{<\theta} | \delta_{m,T} \rangle$: We also introduced a log-normal model (section 16.4.2) and a model based on cylindrical collapse (section 16.4.2) for the convergence profile around lines-of-sight with fixed density contrast $\delta_{m,T}$. Both models lead to almost identical predictions for the density split lensing signal. Hence we again choose the log-normal model for our fiducial analysis, because of the shorter computation time.
- (iii) $P(N_T | \delta_{m,T})$: We introduced two models for the distribution of tracer counts N_T in lines-of-sight of matter density $\delta_{m,T}$. The first was based on linear galaxy bias b and galaxy-matter-correlation coefficient r (section 16.4.3). The second was based on an alternative definition of galaxy bias and on two parameters α_0 and α_1 describing density dependent deviations from Poissonian shot-noise (section 16.4.3). Both models describe the behavior of our simulated tracer galaxies in Buzzard-v1.1 similarly well. But anticipating that real galaxies might behave in a more complicated way, we will consider both ansatzes in our fiducial analysis.

In the following list, we are summarizing the approximations that went into the derivation of our baseline model.

- 1.) We assumed, that for fixed value of $\delta_{m,T}$ the convergence within angular radius θ is not dependent on N_T (cf. equation 16.8).
- 2.) All second order moments in our formalism are computed with a *halofit* power spectrum (Takahashi et al., 2012) using an analytic approximation for the transfer function (Eisenstein & Hu, 1998).
- 3.) Equations 16.36, 16.37, 16.53 and 16.57 employ a small angle and a Limber-like approximation (following Bernardeau & Valageas, 2000).
- 4.) We compute the cumulant generating function of density contrast in long cylinders by means of the cylindrical collapse approximation (cf. section 16.4.1).
- 5.) We assume that the tree-level result of the cumulant generating function can be corrected for the full non-linear evolution of the density field by means of equations 16.47 and 16.48.

6.) We approximate the PDF of $\delta_{m,T}$ resulting from a cylindrical collapse approximation by a log-normal PDF (cf. section 16.4.1).

7.) We employed approximations similar to 4.), 5.) and 6.) for the joint distribution of $\delta_{m,T}$ and $\kappa_{<\theta}$ (cf. section 16.4.2, equation 16.70 and section 16.4.2).

8.) We assume that galaxies are linearly biased tracers of the density field. We consider two different models for shot-noise (resp. stochasticity), assuming that the full distribution $P(N_T|\delta_{m,T})$ is well described by either two parameters (b, r) or three parameters $(\tilde{b}, \alpha_0, \alpha_1)$.

Despite this long list of approximations, this baseline model describes our measurements in the Buzzard simulations well within DES Y1 errorbars (cf. figure 16.1). As shown in the next section, the model is also accurate enough to recover the true cosmology of our simulation within DES Y1 uncertainties in a simulated likelihood analysis. In Gruen et al. (2018) (and using an extended set of simulations) we furthermore show that the values of χ^2 found between our fiducial model and individual simulation measurements are consistent with the χ^2 -distribution expected from our number of data points, and that the coverage (i.e. the fraction of times the true simulation cosmology is within the confidence interval) matches expectations.

16.5 Recovering cosmology in N-body simulations

In this section we want to test whether the modeling that was described in sections 16.2 and 16.4 is sufficient to recover the cosmology underlying a density split data vector measured in N-body simulations. The simulations we use are described in section 16.3.1. They are the same simulations against which we tested the ingredients of our model in the previous section. A likelihood analysis based on a density split data vector measured in these simulations is presented in section 16.5.1. We only run a cosmological analysis on the mean data vector measured on 4 DES-Y1 realizations. The goal of this is to show that any possible systematic deviations between our modeling of density split statistics and the behavior of our N-body simulations is smaller than the statistical uncertainties of DES-Y1. A more extensive validation of our likelihood pipeline is presented in Gruen et al. (2018).

16.5.1 Simulated likelihood analysis

We now measure the data vector that was described in section 16.2.3 in 4 N-body realizations of DES-Y1. We always use the mean of these 4 data vectors. In order to further reduce the noise of this measurement, we turn off the shape noise in our simulated source catalogs, i.e. we measure our signal directly from the gravitational shear acting each galaxy. We then run Monte-Carlo Markov Chains of our model around this data vector. For this we assume a Gaussian likelihood function with the full covariance (i.e. including shape noise) that was estimated by Gruen et al. (2018) for a DES-Y1 data set. The goal of this analysis is to test whether the fiducial cosmology of the Buzzard simulations is well con-

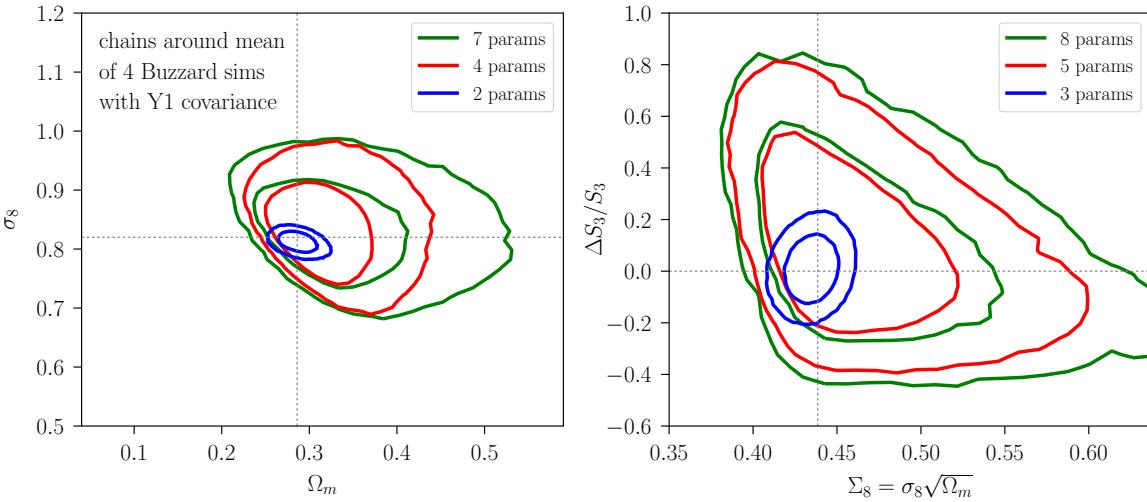


Figure 16.10: To test our model for possible systematic deviations from N-body simulations, we try to recover the Buzzard cosmology in a simulated likelihood analysis. Top panel: 1σ and 2σ contours in the Ω_m - σ_8 plane from a likelihood computed around the mean of 4 shape noise free realizations of DES Y1 (but assuming the full covariance matrix for a single DES Y1). The green contours are marginalized over Ω_b , n_s , h_{100} , REDMAGIC galaxy bias b and galaxy-matter correlation coefficient r . For the parameters Ω_b , n_s , h_{100} we have assumed the same flat priors as used in the DES Y1 combined probes analysis presented in DES Collaboration et al. (2017). The red contours are marginalized only over b and r and the blue contours only vary Ω_m and σ_8 . Even when going to this small parameter space, our model agrees with Buzzard within 1σ errors of DES Y1. Bottom panel: Same contours but in the Σ_8 - $\Delta S_3/S_3$ plane and varying one additional parameter, $\Delta S_3/S_3$.

tained within the 1σ constraints derived from this likelihood analysis. A more extensive validation of our likelihood pipeline is presented in Gruen et al. (2018).

In the top panel of Figure 16.10 we show the 1σ and 2σ constraints obtained from our simulated likelihood analysis in the Ω_m - σ_8 plane after marginalizing over different sets of parameters. First, we only vary Ω_m and σ_8 , setting other cosmological parameters to the inputs of the Buzzard simulations and the parameters connecting galaxy count and matter density to the values we found from the Buzzard galaxy and density maps. Note that those values would be inaccessible in a real measurement. The corresponding constraints are very tight, but the fiducial values of our parameters are still well contained in the 1σ contour. Then, we also marginalize over the galaxy bias b and the galaxy stochasticity r , demanding that $0 < r \leq 1$. The contours now widen, and the fiducial values of Ω_m and σ_8 are still located well within the corresponding 1σ contour. Finally, we also marginalize over Ω_b , n_s and h_{100} , using the same informative priors that have been used in the DES-Y1 combined probes analysis (DES Collaboration et al., 2017). The contours widen further, but our model and our simulations still agree well within 1σ uncertainty.

In the bottom panel of Figure 16.10 we repeat this analysis, but now also vary the parameter $\Delta S_3/S_3$ that was introduced in section 16.2.3. This parameter allows for deviations of the 3-point statistics of the density field from our fiducial model. Within our statistical uncertainties we find that the scaling between 3-point and 2-point statistics in our simulations is well described by our fiducial assumptions ($\Delta S_3/S_3 = 0$).

We repeat this analysis with our alternative shot-noise model in appendix 16.E.

16.6 Discussion & conclusions

In this work we introduced density split statistics, a technique to separately measure contributions to weak lensing and counts-in-cells from regions of different foreground galaxy density. Based on the pioneering work of Bernardeau (1994), Bernardeau & Valageas (2000) and Valageas (2002a) (see also references therein) on modeling the cosmic density PDF we were able to model the density split lensing signal as well as the counts-in-cells histogram from basic principles. With the help of this model, we then showed that density split statistics has two features that make it a potentially powerful cosmological probe:

- it is able to constrain the cosmological parameters Ω_m and σ_8 even if the relation of galaxy density and matter density is assumed to have 2 degrees of freedom: galaxy bias and galaxy-matter-correlation coefficient,
- it is able to constrain the amplitude of 3-point statistics of the density field with almost no degeneracy to constraints on the amplitude of 2-point statistics.

In our fiducial model we predict 3-point statistics from cosmological perturbation theory. Deviations from that fiducial prediction may hint to non-standard physics, that affect over-dense and underdense parts of the matter field differently, or to any non-linear dynamics or small scale physics that break the scaling relations of Λ CDM perturbation theory. We

showed that a DES-Y5 data set combined with data from the cosmic microwave background can measure the amplitude of 3-points statistics to a 1σ accuracy of $\lesssim 5\%$. This is a conservative estimate since our projections neglect the fact that DES-Y5 will be a deeper data set than DES-Y1. Also, we so far neglected the possibility of a combined analysis including density split statistics and measurements of 2-point correlation functions.

Using measurements in high-resolution N-body simulations we showed that our model of the density split lensing signal and the counts-in-cells histogram is accurate to well within the statistical uncertainties of the DES-Y1 data set. Especially, in a mock likelihood analysis we were able to recover the input cosmology of our simulations to well within DES-Y1 parameter errors. Cosmological constraints from DES-Y1 data based on density split statistics are presented in Gruen et al. (2018).

Acknowledgments

OF was supported by SFB-Transregio 33 ‘The Dark Universe’ by the Deutsche Forschungsgemeinschaft (DFG). Support for DG was provided by NASA through Einstein Postdoctoral Fellowship grant number PF5-160138 awarded by the Chandra X-ray Center, which is operated by the Smithsonian Astrophysical Observatory for NASA under contract NAS8-03060. OF and SH acknowledge support by the DFG cluster of excellence ‘Origin and Structure of the Universe’ (www.universe-cluster.de). Part of our computations have been carried out on the computing facilities of the Computational Center for Particle and Astrophysics (C2PAP).

This paper has gone through internal review by the DES collaboration. We want to thank all the members of the DES WL, LSS and Theory working groups that have contributed with helpful comments and discussions. We also want to thank the anonymous journal referee for very helpful comments.

Funding for the DES Projects has been provided by the U.S. Department of Energy, the U.S. National Science Foundation, the Ministry of Science and Education of Spain, the Science and Technology Facilities Council of the United Kingdom, the Higher Education Funding Council for England, the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign, the Kavli Institute of Cosmological Physics at the University of Chicago, the Center for Cosmology and Astro-Particle Physics at the Ohio State University, the Mitchell Institute for Fundamental Physics and Astronomy at Texas A&M University, Financiadora de Estudos e Projetos, Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro, Conselho Nacional de Desenvolvimento Científico e Tecnológico and the Ministério da Ciência, Tecnologia e Inovação, the Deutsche Forschungsgemeinschaft and the Collaborating Institutions in the Dark Energy Survey. The DES data management system is supported by the National Science Foundation under Grant Number AST-1138766.

The Collaborating Institutions are Argonne National Laboratory, the University of California at Santa Cruz, the University of Cambridge, Centro de Investigaciones Enérgéticas, Medioambientales y Tecnológicas-Madrid, the University of Chicago, University College

London, the DES-Brazil Consortium, the University of Edinburgh, the Eidgenössische Technische Hochschule (ETH) Zürich, Fermi National Accelerator Laboratory, the University of Illinois at Urbana-Champaign, the Institut de Ciències de l'Espai (IEEC/CSIC), the Institut de Física d'Altes Energies, Lawrence Berkeley National Laboratory, the Ludwig-Maximilians Universität München and the associated Excellence Cluster Universe, the University of Michigan, the National Optical Astronomy Observatory, the University of Nottingham, The Ohio State University, the University of Pennsylvania, the University of Portsmouth, SLAC National Accelerator Laboratory, Stanford University, the University of Sussex, and Texas A&M University.

The DES participants from Spanish institutions are partially supported by MINECO under grants AYA2012-39559, ESP2013-48274, FPA2013-47986, and Centro de Excelencia Severo Ochoa SEV-2012-0234. Research leading to these results has received funding from the European Research Council under the European Unions Seventh Framework Programme (FP7/2007-2013) including ERC grant agreements 240672, 291329, and 306478.

16.A Friedmann equations, linear growth, spherical collapse and cylindrical collapse

Throughout this section we set $G = 1 = c$ and we assume a flat ΛCDM universe. In proper co-moving time t the Friedmann equations take the form

$$H^2 = \frac{8\pi}{3} (\bar{\rho}_m + \bar{\rho}_\Lambda) \quad (16.95)$$

$$\frac{dH}{dt} + H^2 = -\frac{4\pi}{3} (\bar{\rho}_m - 2\bar{\rho}_\Lambda) , \quad (16.96)$$

where $H = \frac{d}{dt} \ln a$. In conformal time, defined by $dt = ad\tau$, this changes to

$$\mathcal{H}^2 = \frac{8\pi}{3} a^2 (\bar{\rho}_m + \bar{\rho}_\Lambda) \quad (16.97)$$

$$\frac{d\mathcal{H}}{d\tau} = -\frac{4\pi}{3} a^2 (\bar{\rho}_m - 2\bar{\rho}_\Lambda) , \quad (16.98)$$

where $\mathcal{H} = \frac{d}{d\tau} \ln a$. We will from now put $\frac{d}{d\tau} \equiv \dot{\cdot}$.

In the Newtonian approximation, i.e. on scales much smaller than the curvature horizon of the universe, the evolution of a spherical, cylindrical or planar perturbation δ is given by the equation

$$\frac{d^2\delta}{dt^2} + 2H\frac{d\delta}{dt} - \frac{N+1}{N} \frac{1}{1+\delta} \left(\frac{d\delta}{dt} \right)^2 = 4\pi\bar{\rho}_m\delta(1+\delta) , \quad (16.99)$$

where $N = 3$ for a spherical perturbation, $N = 2$ for a cylindrical perturbation and $N = 1$ for a planar perturbation (see Mukhanov (2005) where this is demonstrated for $N = 3$ and $N = 1$).

In conformal time this equation reads

$$\ddot{\delta} + \mathcal{H}\dot{\delta} - \frac{N+1}{N} \frac{\dot{\delta}^2}{1+\delta} = 4\pi\bar{\rho}_m a^2 \delta(1+\delta) . \quad (16.100)$$

To linear order in δ this becomes

$$\ddot{\delta} + \mathcal{H}\dot{\delta} = 4\pi\bar{\rho}_m a^2 \delta , \quad (16.101)$$

which is indeed independent of the particular shape of the perturbation.

16.B Λ CDM perturbation theory

Consider the matter density contrast δ and the divergence of the velocity field $\theta = \nabla \cdot \mathbf{v}$. In the Newtonian approximation the Fourier space equations of motion of δ and θ are (cf. Bernardeau et al., 2002)

$$\begin{aligned} \frac{\partial \tilde{\delta}(\mathbf{k}, \tau)}{\partial \tau} + \tilde{\theta}(\mathbf{k}, \tau) &= - \int d^3 k_1 d^3 k_2 \delta_D(\mathbf{k} - \mathbf{k}_{12}) \alpha(\mathbf{k}_1, \mathbf{k}_2) \tilde{\delta}(\mathbf{k}_1, \tau) \tilde{\theta}(\mathbf{k}_2, \tau) \\ \frac{\partial \tilde{\theta}(\mathbf{k}, \tau)}{\partial \tau} + \mathcal{H} \tilde{\theta}(\mathbf{k}, \tau) + \frac{3\Omega_m^0 H_0^2}{2a} \tilde{\delta}(\mathbf{k}, \tau) &= - \int d^3 k_1 d^3 k_2 \delta_D(\mathbf{k} - \mathbf{k}_{12}) \beta(\mathbf{k}_1, \mathbf{k}_2) \tilde{\theta}(\mathbf{k}_1, \tau) \tilde{\theta}(\mathbf{k}_2, \tau) \end{aligned}$$

where $\mathbf{k}_{12} = \mathbf{k}_1 + \mathbf{k}_2$ and α and β are given by

$$\begin{aligned} \alpha(\mathbf{k}_1, \mathbf{k}_2) &= 1 + \frac{1}{2} \frac{\mathbf{k}_1 \cdot \mathbf{k}_2}{k_1 k_2} \left(\frac{k_1}{k_2} + \frac{k_2}{k_1} \right) \\ \beta(\mathbf{k}_1, \mathbf{k}_2) &= \frac{1}{2} \frac{\mathbf{k}_1 \cdot \mathbf{k}_2}{k_1 k_2} \left(\frac{k_1}{k_2} + \frac{k_2}{k_1} \right) + \frac{(\mathbf{k}_1 \cdot \mathbf{k}_2)^2}{k_1^2 k_2^2} . \end{aligned} \quad (16.103)$$

In the following we will abbreviate the integrals involving α and β as $\alpha[\tilde{\delta}, \tilde{\theta}, \mathbf{k}]$ and $\beta[\tilde{\theta}, \tilde{\theta}, \mathbf{k}]$. In perturbation theory we write $\tilde{\delta}$ and $\tilde{\theta}$ as

$$\tilde{\delta}(\mathbf{k}, \tau) = \sum_{n=1}^{\infty} \delta_n(\mathbf{k}, \tau) \quad \text{and} \quad \tilde{\theta}(\mathbf{k}, \tau) = -\frac{\partial \ln D_+(\tau)}{\partial \tau} \sum_{n=1}^{\infty} \theta_n(\mathbf{k}, \tau) , \quad (16.104)$$

where δ_n and θ_n are of order n in the linearly approximated fields δ_1 and θ_1 and $D_+(\tau)$ is the linear growth factor. (We will ignore the decaying mode of linear growth here.) At linear order we have

$$\delta_1(\mathbf{k}, \tau) = \theta_1(\mathbf{k}, \tau) = \frac{D_+(\tau)}{D_+(\tau_0)} \delta_1(\mathbf{k}, \tau_0) \equiv D_+(\tau) \delta_{1,1}(\mathbf{k}) , \quad (16.105)$$

where we have assumed that $D_+(\tau_0) = 1$ at present time τ_0 and introduced the notation $\delta_{1,1}(\mathbf{k}) = \delta_1(\mathbf{k}, \tau_0)$ whose purpose will become clear at the end of this section. To get $\tilde{\delta}$ at second order we first note that

$$\frac{\partial \tilde{\theta}(\mathbf{k}, \tau)}{\partial \tau} + \mathcal{H} \tilde{\theta}(\mathbf{k}, \tau) = \frac{1}{a} \frac{\partial}{\partial \tau} \left(a \tilde{\theta}(\mathbf{k}, \tau) \right) . \quad (16.106)$$

Hence, multiplying the first of equations 16.102 with a and differentiating wrt. τ and then multiplying with $1/a$ we get

$$\frac{\partial^2 \tilde{\delta}(\mathbf{k}, \tau)}{\partial^2 \tau} + \mathcal{H} \frac{\partial \tilde{\delta}(\mathbf{k}, \tau)}{\partial \tau} + \frac{\partial \tilde{\theta}(\mathbf{k}, \tau)}{\partial \tau} + \mathcal{H} \tilde{\theta}(\mathbf{k}, \tau) = -\alpha \left[\frac{\partial \tilde{\delta}}{\partial \tau}, \tilde{\theta}, \mathbf{k} \right] - \alpha \left[\tilde{\delta}, \frac{\partial \tilde{\theta}}{\partial \tau}, \mathbf{k} \right] - \mathcal{H} \alpha \left[\tilde{\delta}, \tilde{\theta}, \mathbf{k} \right]. \quad (16.107)$$

Now the second of equations 16.102 can be used to eliminate $\tilde{\theta}$ from the right-hand-side, giving

$$\frac{\partial^2 \tilde{\delta}(\mathbf{k}, \tau)}{\partial^2 \tau} + \mathcal{H} \frac{\partial \tilde{\delta}(\mathbf{k}, \tau)}{\partial \tau} - \frac{3\Omega_m^0 H_0^2}{2a} \tilde{\delta}(\mathbf{k}, \tau) = \beta[\tilde{\theta}, \tilde{\theta}, \mathbf{k}] - \alpha \left[\frac{\partial \tilde{\delta}}{\partial \tau}, \tilde{\theta}, \mathbf{k} \right] - \alpha \left[\tilde{\delta}, \frac{\partial \tilde{\theta}}{\partial \tau}, \mathbf{k} \right] - \mathcal{H} \alpha \left[\tilde{\delta}, \tilde{\theta}, \mathbf{k} \right]. \quad (16.108)$$

At second order in perturbation theory this equation becomes

$$\begin{aligned} & \frac{\partial^2 \delta_2(\mathbf{k}, \tau)}{\partial^2 \tau} + \mathcal{H} \frac{\partial \delta_2(\mathbf{k}, \tau)}{\partial \tau} - \frac{3\Omega_m^0 H_0^2}{2a} \delta_2(\mathbf{k}, \tau) \\ &= \left(\frac{\partial D_+}{\partial \tau} \right)^2 \beta[\delta_{1,1}, \delta_{1,1}, \mathbf{k}] + \left(D \frac{\partial^2 D_+}{\partial \tau^2} + D \mathcal{H} \frac{\partial D_+}{\partial \tau} + \left(\frac{\partial D_+}{\partial \tau} \right)^2 \right) \alpha[\delta_{1,1}, \delta_{1,1}, \mathbf{k}] \\ &= \left(\frac{\partial D_+}{\partial \tau} \right)^2 \beta[\delta_{1,1}, \delta_{1,1}, \mathbf{k}] + \left(\frac{3\Omega_m^0 H_0^2}{2a} D^2 + \left(\frac{\partial D_+}{\partial \tau} \right)^2 \right) \alpha[\delta_{1,1}, \delta_{1,1}, \mathbf{k}] \\ &= \alpha[\delta_{1,1}, \delta_{1,1}, \mathbf{k}] \left(\frac{3\Omega_m^0 H_0^2}{2a} D^2 + 2 \left(\frac{\partial D_+}{\partial \tau} \right)^2 \right) \\ &+ (\beta[\delta_{1,1}, \delta_{1,1}, \mathbf{k}] - \alpha[\delta_{1,1}, \delta_{1,1}, \mathbf{k}]) \left(\frac{\partial D_+}{\partial \tau} \right)^2. \end{aligned} \quad (16.109)$$

This is solved by

$$\delta_2(\mathbf{k}, \tau) = D_{2,1}(\tau) \delta_{2,1}(\mathbf{k}) + D_{2,2}(\tau) \delta_{2,2}(\mathbf{k}) \quad (16.110)$$

where

$$D_{2,1}(\tau) \equiv D_+^2(\tau), \quad \delta_{2,1}(\mathbf{k}) = \alpha[\delta_{1,1}, \delta_{1,1}, \mathbf{k}], \quad \delta_{2,2}(\mathbf{k}) = \beta[\delta_{1,1}, \delta_{1,1}, \mathbf{k}] - \alpha[\delta_{1,1}, \delta_{1,1}, \mathbf{k}] \quad (16.111)$$

and $D_{2,2}$ is given by the differential equation

$$\frac{\partial^2 D_{2,2}(\tau)}{\partial^2 \tau} + \mathcal{H} \frac{\partial D_{2,2}(\tau)}{\partial \tau} - \frac{3\Omega_m^0 H_0^2}{2a} D_{2,2}(\tau) = \left(\frac{\partial D_+}{\partial \tau} \right)^2. \quad (16.112)$$

16.B.1 Second order of δ in Einstein-de Sitter universe

Let us define $1 - \mu \equiv D_{2,2}/D_+^2$. Then the general solution to 16.109 is given by

$$\delta_2(\mathbf{k}, \tau) = D_+^2 ([1 - \mu] \beta[\delta_{1,1}, \delta_{1,1}, \mathbf{k}] + \mu \alpha[\delta_{1,1}, \delta_{1,1}, \mathbf{k}]). \quad (16.113)$$

In an Einstein-de Sitter Universe where $\Omega_m^0 = 1$ and $D \equiv a$ we have

$$D_{2,2} = \frac{2}{7} D_+^2, \quad \mu = \frac{5}{7} \quad (16.114)$$

and δ_2 is hence given by

$$\begin{aligned} \delta_2(\mathbf{k}, \tau) &= D_+^2 \left(\frac{2}{7} \beta[\delta_{1,1}, \delta_{1,1}, \mathbf{k}] + \frac{5}{7} \alpha[\delta_{1,1}, \delta_{1,1}, \mathbf{k}] \right) \\ &= \int d^3k_1 d^3k_2 \delta_D(\mathbf{k} - \mathbf{k}_{12}) F_2(\mathbf{k}_1, \mathbf{k}_2) \delta_{1,1}(\mathbf{k}_1) \delta_{1,1}(\mathbf{k}_2) \end{aligned} \quad (16.115)$$

with

$$\begin{aligned} F_2(\mathbf{k}_1, \mathbf{k}_2) &= \frac{5}{7} \alpha(\mathbf{k}_1, \mathbf{k}_2) + \frac{2}{7} \beta(\mathbf{k}_1, \mathbf{k}_2) \\ &= \frac{5}{7} + \frac{1}{2} \frac{\mathbf{k}_1 \cdot \mathbf{k}_2}{k_1 k_2} \left(\frac{k_1}{k_2} + \frac{k_2}{k_1} \right) + \frac{2}{7} \frac{(\mathbf{k}_1 \cdot \mathbf{k}_2)^2}{k_1^2 k_2^2}. \end{aligned} \quad (16.116)$$

16.B.2 Second order of δ in Λ CDM universe

In a general Λ CDM universe the function F_2 becomes time dependent. It is given by

$$F_2(\mathbf{k}_1, \mathbf{k}_2, \tau) = \mu(\tau) + \frac{1}{2} \frac{\mathbf{k}_1 \cdot \mathbf{k}_2}{k_1 k_2} \left(\frac{k_1}{k_2} + \frac{k_2}{k_1} \right) + [1 - \mu(\tau)] \frac{(\mathbf{k}_1 \cdot \mathbf{k}_2)^2}{k_1^2 k_2^2}. \quad (16.117)$$

A useful property of this kernel is that

$$F_2(\mathbf{k}, -\mathbf{k}, \tau) = \mu(\tau) + \frac{1}{2} \frac{-k^2}{k^2} (1 + 1) + [1 - \mu(\tau)] \frac{k^4}{k^4} = \mu(\tau) - 1 + 1 - \mu(\tau) = 0. \quad (16.118)$$

Denoting the angle between \mathbf{k}_1 and \mathbf{k}_2 with ϕ one can also arrive at the following form of $F_2(\mathbf{k}_1, \mathbf{k}_2, \tau)$ which will be useful when computing the skewness of matter inside a long cylinder:

$$F_2(\mathbf{k}_1, \mathbf{k}_2, \tau) = \frac{1}{2} \left\{ \left(1 + \frac{k_1}{k_2} \cos \phi \right) + \left(1 + \frac{k_2}{k_1} \cos \phi \right) \right\} + [1 - \mu(\tau)] (\cos^2 \phi - 1). \quad (16.119)$$

16.B.3 Bispectrum and 3-point function at leading order

The bispectrum $B(k_1, k_2, k_3, \tau)$ of δ is defined by

$$\langle \tilde{\delta}(\mathbf{k}_1, \tau) \tilde{\delta}(\mathbf{k}_2, \tau) \tilde{\delta}(\mathbf{k}_3, \tau) \rangle = \delta_D(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) B(k_1, k_2, k_3, \tau). \quad (16.120)$$

At leading order in perturbation theory this can be calculated as

$$\begin{aligned} &\langle \tilde{\delta}(\mathbf{k}_1, \tau) \tilde{\delta}(\mathbf{k}_2, \tau) \tilde{\delta}(\mathbf{k}_3, \tau) \rangle_{\text{2nd}} \\ &= D_+^2 \int d^3q_1 d^3q_2 \delta_D(\mathbf{k}_3 - \mathbf{q}_{12}) F_2(\mathbf{q}_1, \mathbf{q}_2, \tau) \langle \delta_{1,1}(\mathbf{k}_1) \delta_{1,1}(\mathbf{k}_2) \delta_{1,1}(\mathbf{q}_1) \delta_{1,1}(\mathbf{q}_2) \rangle \\ &\quad + \text{cycl.}, \end{aligned} \quad (16.121)$$

where 'cycl.' indicates that the integral on the right-hand-side should appear for all possible permutations of \mathbf{k}_1 , \mathbf{k}_3 and \mathbf{k}_3 . Since we assume the linear density field to be a Gaussian random field, the expectation value on the left-hand-side factorizes as

$$\begin{aligned}
& \langle \delta_{1,1}(\mathbf{k}_1) \delta_{1,1}(\mathbf{k}_2) \delta_{1,1}(\mathbf{q}_1) \delta_{1,1}(\mathbf{q}_2) \rangle \\
= & \langle \delta_{1,1}(\mathbf{k}_1) \delta_{1,1}(\mathbf{k}_2) \rangle \langle \delta_{1,1}(\mathbf{q}_1) \delta_{1,1}(\mathbf{q}_2) \rangle + \langle \delta_{1,1}(\mathbf{k}_1) \delta_{1,1}(\mathbf{q}_1) \rangle \langle \delta_{1,1}(\mathbf{k}_2) \delta_{1,1}(\mathbf{q}_2) \rangle \\
& + \langle \delta_{1,1}(\mathbf{k}_1) \delta_{1,1}(\mathbf{q}_2) \rangle \langle \delta_{1,1}(\mathbf{k}_2) \delta_{1,1}(\mathbf{q}_1) \rangle \\
= & \delta_D(\mathbf{k}_1 + \mathbf{k}_2) \delta_D(\mathbf{q}_1 + \mathbf{q}_2) P_{\text{lin},0}(k_1) P_{\text{lin},0}(q_1) + \delta_D(\mathbf{k}_1 + \mathbf{q}_1) \delta_D(\mathbf{k}_2 + \mathbf{q}_2) P_{\text{lin},0}(k_1) P_{\text{lin},0}(q_2) \\
& + \delta_D(\mathbf{k}_1 + \mathbf{q}_2) \delta_D(\mathbf{k}_2 + \mathbf{q}_1) P_{\text{lin},0}(k_1) P_{\text{lin},0}(q_1) \tag{16.122}
\end{aligned}$$

Because of equation 16.118 the contribution of the first term to the bispectrum is zero. Using the symmetry $1 \leftrightarrow 2$ between the second and third term we hence get

$$\langle \tilde{\delta}(\mathbf{k}_1, \tau) \tilde{\delta}(\mathbf{k}_2, \tau) \tilde{\delta}(\mathbf{k}_3, \tau) \rangle_{\text{2nd.}} = 2D_+^2 \delta_D(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) F_2(\mathbf{k}_1, \mathbf{k}_2, \tau) P_{\text{lin},0}(k_1) P_{\text{lin},0}(k_2) + \text{cycl.} \tag{16.123}$$

16.B.4 Variance and skewness of the density contrast in long cylinders at leading order in perturbation theory

Consider a cylinder with radius R and length L . In Fourier space the top-hat filter for this cylinder is given by

$$W_{R,L}(\mathbf{k}) = \frac{1}{(2\pi)^3} W_L(k_{\parallel}) W_R(\mathbf{k}_{\perp}) \tag{16.124}$$

where we denote the component of \mathbf{k} parallel to the cylinder axis with k_{\parallel} and the components orthogonal to it are represented by the two-dimensional vector \mathbf{k}_{\perp} and W_L and W_R given by

$$W_L(k_{\parallel}) = \frac{\sin(Lk_{\parallel}/2)}{Lk_{\parallel}/2}, \quad W_R(\mathbf{k}_{\perp}) = \frac{2J_1(Rk_{\perp})}{Rk_{\perp}}. \tag{16.125}$$

Here $k_{\perp} = |\mathbf{k}_{\perp}|$ and J_{ν} are the cylindrical Bessel functions. At leading order or tree level in perturbation theory the variance of matter contrast within the cylinder is then given by

$$\begin{aligned}
& \langle \delta_{R,L}^2 \rangle_{\text{tree}}(\tau) \\
= & D_+^2(\tau) \int dk_{\parallel,1} dk_{\parallel,2} dk_{\perp,1} dk_{\perp,2} W_L(k_{\parallel,1}) W_L(k_{\parallel,2}) W_R(\mathbf{k}_{\perp,1}) W_R(\mathbf{k}_{\perp,2}) \langle \delta_{1,1}(\mathbf{k}_1) \delta_{1,1}(\mathbf{k}_2) \rangle \\
= & D_+^2(\tau) \int dk_{\parallel} dk_{\perp} W_L(k_{\parallel})^2 W_R(\mathbf{k}_{\perp})^2 P_{\text{lin},0}(\mathbf{k}) . \tag{16.126}
\end{aligned}$$

Here $P_{\text{lin},0}(\mathbf{k})$ is today's linear power spectrum. For $L \gg R$ we can actually approximate W_L by

$$W_L(k_{\parallel})^2 \approx \frac{2\pi}{L} \delta_D(k_{\parallel}) \tag{16.127}$$

such that in this limit we get

$$\langle \delta_{R,L}^2 \rangle_{\text{tree}}(\tau) \approx \frac{2\pi D_+^2}{L} \int dk \ k \ W_R(k)^2 \ P_{\text{lin},0}(k) . \quad (16.128)$$

The third moment at tree level is given by

$$\begin{aligned} & \langle \delta_{R,L}^3 \rangle_{\text{tree}}(\tau) \\ = & 3D_+^2 \int dk_{\parallel,1} dk_{\parallel,2} dq_{\parallel} d^2 k_{\perp,1} d^2 k_{\perp,2} d^2 q_{\perp} W_L(k_{\parallel,1}) W_L(k_{\parallel,2}) W_L(q_{\parallel}) W_R(\mathbf{k}_{\perp,1}) W_R(\mathbf{k}_{\perp,2}) W_R(\mathbf{q}_{\perp}) \\ & \langle \delta_{1,1}(\mathbf{k}_1) \delta_{1,1}(\mathbf{k}_2) \delta_2(\mathbf{q}, \tau) \rangle \\ = & 3D_+^4 \int dk_{\parallel,1} dk_{\parallel,2} dq_{\parallel} d^2 k_{\perp,1} d^2 k_{\perp,2} d^2 q_{\perp} d^3 q_1 d^3 q_2 W_L(k_{\parallel,1}) W_L(k_{\parallel,2}) W_L(q_{\parallel}) W_R(\mathbf{k}_{\perp,1}) W_R(\mathbf{k}_{\perp,2}) \\ & W_R(\mathbf{q}_{\perp}) \delta_D(\mathbf{q} - \mathbf{q}_1 - \mathbf{q}_2) F_2(\mathbf{q}_1, \mathbf{q}_2, \tau) \langle \delta_{1,1}(\mathbf{k}_1) \delta_{1,1}(\mathbf{k}_2) \delta_{1,1}(\mathbf{q}_1) \delta_{1,1}(\mathbf{q}_2) \rangle \\ = & 3D_+^4 \int dk_{\parallel,1} dk_{\parallel,2} d^2 k_{\perp,1} d^2 k_{\perp,2} d^3 q_1 d^3 q_2 W_L(k_{\parallel,1}) W_L(k_{\parallel,2}) W_R(\mathbf{k}_{\perp,1}) W_R(\mathbf{k}_{\perp,2}) \\ & W_R(\mathbf{q}_{\perp,1} + \mathbf{q}_{\perp,2}) W_L(q_{\parallel,1} + q_{\parallel,2}) F_2(\mathbf{q}_1, \mathbf{q}_2, \tau) \langle \delta_{1,1}(\mathbf{k}_1) \delta_{1,1}(\mathbf{k}_2) \delta_{1,1}(\mathbf{q}_1) \delta_{1,1}(\mathbf{q}_2) \rangle . \end{aligned} \quad (16.129)$$

Since we assume the linear density field to be a Gaussian random field, the expectation value on the left-hand-side factorizes as

$$\begin{aligned} & \langle \delta_{1,1}(\mathbf{k}_1) \delta_{1,1}(\mathbf{k}_2) \delta_{1,1}(\mathbf{q}_1) \delta_{1,1}(\mathbf{q}_2) \rangle \\ = & \langle \delta_{1,1}(\mathbf{k}_1) \delta_{1,1}(\mathbf{k}_2) \rangle \langle \delta_{1,1}(\mathbf{q}_1) \delta_{1,1}(\mathbf{q}_2) \rangle + \langle \delta_{1,1}(\mathbf{k}_1) \delta_{1,1}(\mathbf{q}_1) \rangle \langle \delta_{1,1}(\mathbf{k}_2) \delta_{1,1}(\mathbf{q}_2) \rangle \\ & + \langle \delta_{1,1}(\mathbf{k}_1) \delta_{1,1}(\mathbf{q}_2) \rangle \langle \delta_{1,1}(\mathbf{k}_2) \delta_{1,1}(\mathbf{q}_1) \rangle \\ = & \delta_D(\mathbf{k}_1 + \mathbf{k}_2) \delta_D(\mathbf{q}_1 + \mathbf{q}_2) P_{\text{lin},0}(k_1) P_{\text{lin},0}(q_1) + \delta_D(\mathbf{k}_1 + \mathbf{q}_1) \delta_D(\mathbf{k}_2 + \mathbf{q}_2) P_{\text{lin},0}(k_1) P_{\text{lin},0}(q_2) \\ & + \delta_D(\mathbf{k}_1 + \mathbf{q}_2) \delta_D(\mathbf{k}_2 + \mathbf{q}_1) P_{\text{lin},0}(k_1) P_{\text{lin},0}(q_1) \end{aligned} \quad (16.130)$$

Because of equation 16.118 the contribution of the first term to the skewness is zero. Using the symmetry $1 \leftrightarrow 2$ between the second and third term we hence get

$$\begin{aligned} \langle \delta_{R,L}^3 \rangle_{\text{tree}}(\tau) &= 6D_+^4 \int dq_{\parallel,1} dq_{\parallel,2} d^2 q_{\perp,1} d^2 q_{\perp,2} W_L(q_{\parallel,1}) W_L(q_{\parallel,2}) W_L(q_{\parallel,1} + q_{\parallel,2}) \\ & W_R(\mathbf{q}_1) W_R(\mathbf{q}_2) W_R(\mathbf{q}_1 + \mathbf{q}_2) P_{\text{lin},0}(q_1) P_{\text{lin},0}(q_2) F_2(\mathbf{q}_1, \mathbf{q}_2, \tau) . \end{aligned} \quad (16.131)$$

For $L \gg R$ we can use the approximation

$$W_L(q_{\parallel,1}) W_L(q_{\parallel,2}) W_L(q_{\parallel,1} + q_{\parallel,2}) \approx \frac{1}{L^2} \delta_D^2(q_{\parallel,1}, q_{\parallel,2}) . \quad (16.132)$$

This gives

$$\langle \delta_{R,L}^3 \rangle_{\text{tree}}(\tau) = \frac{6D_+^4}{L^2} \int d^2 q_1 d^2 q_2 W_R(\mathbf{q}_1) W_R(\mathbf{q}_2) W_R(\mathbf{q}_1 + \mathbf{q}_2) P_{\text{lin},0}(q_1) P_{\text{lin},0}(q_2) F_2(\mathbf{q}_1, \mathbf{q}_2, \tau) , \quad (16.133)$$

where we will consider all vectors to be 2-dimensional from now on. Using equation 16.119 to express F_2 in terms of q_1 , q_2 and ϕ we can simplify this to

$$\begin{aligned} \langle \delta_{R,L}^3 \rangle_{\text{tree}}(\tau) &= \frac{12\pi D_+^4}{L^2} \int dq_1 dq_2 q_1 W_R(q_1) q_2 W_R(q_2) P_{\text{lin},0}(q_1) P_{\text{lin},0}(q_2) * \\ &\quad * \int d\phi W_R \left[\sqrt{q_1^2 + q_2^2 + 2q_1 q_2 \cos \phi} \right] F_2(q_1, q_2, \phi, \tau) . \end{aligned} \quad (16.134)$$

Using relations given in Bernardeau (1995) or Buchalter et al. (2000) one can simplify the integral over ϕ as

$$\begin{aligned} &\int d\phi W_R \left[\sqrt{q_1^2 + q_2^2 + 2q_1 q_2 \cos \phi} \right] F_2(q_1, q_2, \phi, \tau) \\ &= \frac{1}{2} \int d\phi W_R \left[\sqrt{q_1^2 + q_2^2 + 2q_1 q_2 \cos \phi} \right] \left\{ \left(1 + \frac{k_1}{k_2} \cos \phi \right) + \left(1 + \frac{k_2}{k_1} \cos \phi \right) \right\} \\ &\quad + [1 - \mu(\tau)] \int d\phi W_R \left[\sqrt{q_1^2 + q_2^2 + 2q_1 q_2 \cos \phi} \right] (\cos^2 \phi - 1) \\ &= 2\pi W_R(q_1) W_R(q_2) + \frac{\pi}{2} \left\{ W_R(q_1) R q_2 \frac{\partial W_R(x)}{\partial x} \Big|_{x=Rq_2} + W_R(q_2) R q_1 \frac{\partial W_R(x)}{\partial x} \Big|_{x=Rq_1} \right\} \\ &\quad - \pi [1 - \mu(\tau)] W_R(q_1) W_R(q_2) \\ &= \pi [1 + \mu(\tau)] W_R(q_1) W_R(q_2) + \frac{\pi}{2} \frac{\partial}{\partial \ln R} \{W_R(q_1) W_R(q_2)\} . \end{aligned} \quad (16.135)$$

For the third moment of $\delta_{R,L}$ this gives

$$\begin{aligned} &\langle \delta_{R,L}^3 \rangle_{\text{tree}}(\tau) \\ &= [1 + \mu(\tau)] \frac{12\pi^2 D_+^4(\tau)}{L^2} \int dq_1 dq_2 q_1 q_2 W_R(q_1)^2 W_R(q_2)^2 P_{\text{lin},0}(q_1) P_{\text{lin},0}(q_2) \\ &\quad + \frac{6\pi^2 D_+^4(\tau)}{L^2} \int dq_1 dq_2 q_1 q_2 W_R(q_1) W_R(q_2) \frac{\partial}{\partial \ln R} \{W_R(q_1) W_R(q_2)\} P_{\text{lin},0}(q_1) P_{\text{lin},0}(q_2) \\ &= 3[1 + \mu(\tau)] \left(\frac{2\pi D_+^2(\tau)}{L} \int dq_1 q W_R(q)^2 P_{\text{lin},0}(q) \right)^2 \\ &\quad + \frac{3\pi^2 D_+^4(\tau)}{L^2} \frac{\partial}{\partial \ln R} \int dq_1 dq_2 q_1 q_2 W_R(q_1)^2 W_R(q_2)^2 P_{\text{lin},0}(q_1) P_{\text{lin},0}(q_2) \\ &= 3[1 + \mu(\tau)] \left(\frac{2\pi D_+^2(\tau)}{L} \int dq_1 q W_R(q)^2 P_{\text{lin},0}(q) \right)^2 \\ &\quad + \frac{3}{4} \frac{\partial}{\partial \ln R} \left(\frac{2\pi D_+^2(\tau)}{L^2} \int dq_1 q W_R(q)^2 P_{\text{lin},0}(q) \right)^2 \\ &= 3[1 + \mu(\tau)] (\langle \delta_{R,L}^2 \rangle_{\text{tree}}(\tau))^2 + \frac{3}{4} \frac{\partial}{\partial \ln R} (\langle \delta_{R,L}^2 \rangle_{\text{tree}}(\tau))^2 . \end{aligned} \quad (16.136)$$

Especially we have

$$S_3 \equiv \frac{\langle \delta_{R,L}^3 \rangle_{\text{tree}}(\tau)}{\langle \delta_{R,L}^2 \rangle_{\text{tree}}(\tau)^2} = 3[1 + \mu(\tau)] + \frac{3}{2} \frac{\partial \ln \langle \delta_{R,L}^2 \rangle_{\text{tree}}(\tau)}{\partial \ln R} . \quad (16.137)$$

For an Einstein-de Sitter universe and a power law power spectrum $P(k) \sim k^n$ this gives $S_3 = 36/7 - 3/2(n+2)$. The leading order prediction for S_3 is surprisingly good, even in the mildly non-linear regime (see Bernardeau et al., 2002, and references therein). Hence in order to predict the non-linear skewness, we simply employ the approximation

$$\langle \delta_{R,L}^3 \rangle_{\text{non-lin.}}(\tau) \approx S_3 \langle \delta_{R,L}^2 \rangle_{\text{non-lin.}}(\tau)^2 , \quad (16.138)$$

where we compute the non-linear variance with the use of halofit as detailed in Takahashi et al. (2014) which is a revised version of Smith et al. (2003).

16.B.5 The moment $\langle \delta_{R_A,L}^2 \delta_{R_B,L} \rangle_{\text{tree}}$

For predicting the density split lensing signal we are also interested in the moment $\langle \delta_{R_A,L}^2 \delta_{R_B,L} \rangle_{\text{tree}}$, where R_A and R_B are two different Radii. The above derivations can be generalized to give

$$\begin{aligned} & \langle \delta_{R_A,L}^2 \delta_{R_B,L} \rangle_{\text{tree}}(\tau) \\ = & \text{Var}(R_A) \text{Cov}(R_A, R_B) \left\{ 2[1 + \mu(\tau)] + \frac{1}{2} \frac{\partial \ln \text{Var}(R_A)}{\partial \ln R_A} + \frac{\partial \ln \text{Cov}(R_A, R_B)}{\partial \ln R_A} \right\} \\ & + \text{Cov}(R_A, R_B)^2 \left\{ [1 + \mu(\tau)] + \frac{\partial \ln \text{Cov}(R_A, R_B)}{\partial \ln R_B} \right\} \end{aligned} \quad (16.139)$$

where we defined

$$\text{Var}(R_A) = \langle \delta_{R_A,L}^2 \rangle_{\text{tree}} , \quad \text{Cov}(R_A, R_B) = \langle \delta_{R_A,L} \delta_{R_B,L} \rangle_{\text{tree}} . \quad (16.140)$$

To correct this expression for the non-linear evolution of the power spectrum, we compute $\text{Var}(R_A)$ and $\text{Cov}(R_A, R_B)$ with our halofit power spectrum whenever they appear outside of the logarithmic derivatives. This is a generalization of the rescaling of $\text{Var}(R_A)$ by means of S_3 .

For $R_B \gg R_A$ this rescaling is dominated by first term on the right hand-side of equation 16.139. For $R_B \approx R_A$ it reduces to equation 16.138. As a consequence, using the procedure described around 16.70 to correct for the non-linear power spectrum evolution yields a prediction for the density split lensing signal that is almost identical to the procedure described here. Also, it can be considered accurate to the extend that equation 16.138 is accurate. We nevertheless rescale the 3rd order moments in the more elaborate way described here.

16.B.6 The moment $\langle \delta_{R_A,L}^n \delta_{R_B,L} \rangle_{\text{tree}}$

Using a diagrammatic representation of perturbation theory (see, e.g. , Bernardeau et al. 2002) one can see that the tree-level result for the moment $\langle \delta_{R_A,L}^n \delta_{R_B,L} \rangle_c$ will consist of terms that scale as

$$\sim \text{Cov}(R_A, R_B)^k \text{Var}(R_A)^{n-k}, \quad 1 \leq k \leq n. \quad (16.141)$$

For $R_B \approx R_A$ each of these scalings reduces to $\sim \text{Var}(R_A)^n$ (cf. 16.138 and the definition of S_{n+1} in 16.47). On the other hand, for $R_B \gg R_A$ the terms scaling as $\sim \text{Cov}(R_A, R_B) \text{Var}(R_A)^{n-1}$ are the dominant contributions (cf. the last section for the case $n = 2$). This is why we use 16.70 when rescaling moments with $n > 2$ in $G_{\text{cyl},\theta_T w,\theta w,L}(q_l(w) Ly, w)$ (see also 16.57).

16.C Comparison with Millennium simulation

In Figure 16.11, we compare our model for the PDF of projected density contrast to another set of N-body simulations, the Millennium Run (MR Springel et al., 2005). The MR has a smaller simulation volume of only $(500h^{-1} \text{Mpc})^3$ co-moving, but a force resolution of $5h^{-1} \text{kpc}$ that is 4-10 times higher than that of the Buzzard simulations. The fiducial model and the log-normal model describe the distribution of $\delta_{m,T}$ measured from the MR well considering the large statistical uncertainty on $p(\delta_{m,T})$ due to the limited simulated sky area.

16.D Galaxy stochasticity

Consider the field of galaxy density contrast $\delta_{g,T}$ and the field of matter density contrast $\delta_{m,T}$, where both fields are assumed to be smoothed over a fixed circular aperture. The number of galaxies found inside such an aperture is assumed to be a Poissonian random variable with first and second moments for a given value of $\delta_{g,T}$ are given by

$$\langle \hat{N} | \delta_{g,T} \rangle = \bar{N}(1 + \delta_{g,T}) \quad (16.142)$$

and

$$\langle \hat{N}^2 | \delta_{g,T} \rangle = \bar{N}(1 + \delta_{g,T}) + \bar{N}^2(1 + \delta_{g,T})^2. \quad (16.143)$$

Let Var_m be the variance of $\delta_{m,T}$ and $\text{Var}_g = b^2 \text{Var}_m$ the variance of $\delta_{g,T}$, where b is the galaxy bias. Then the galaxy stochasticity r is defined by $\text{Cov}_{mg} = rb \text{Var}_m$, i.e. it is the correlation coefficient of $\delta_{g,T}$ and $\delta_{m,T}$.

We will now assume both $\delta_{g,T}$ and $\delta_{m,T}$ to be joint log-normal random variables, i.e.

$$\begin{aligned} \delta_{m,T} &= [e^{n_m} - 1] \delta_{m,0} \\ \delta_{g,T} &= [e^{n_g} - 1] \delta_{g,0}, \end{aligned} \quad (16.144)$$

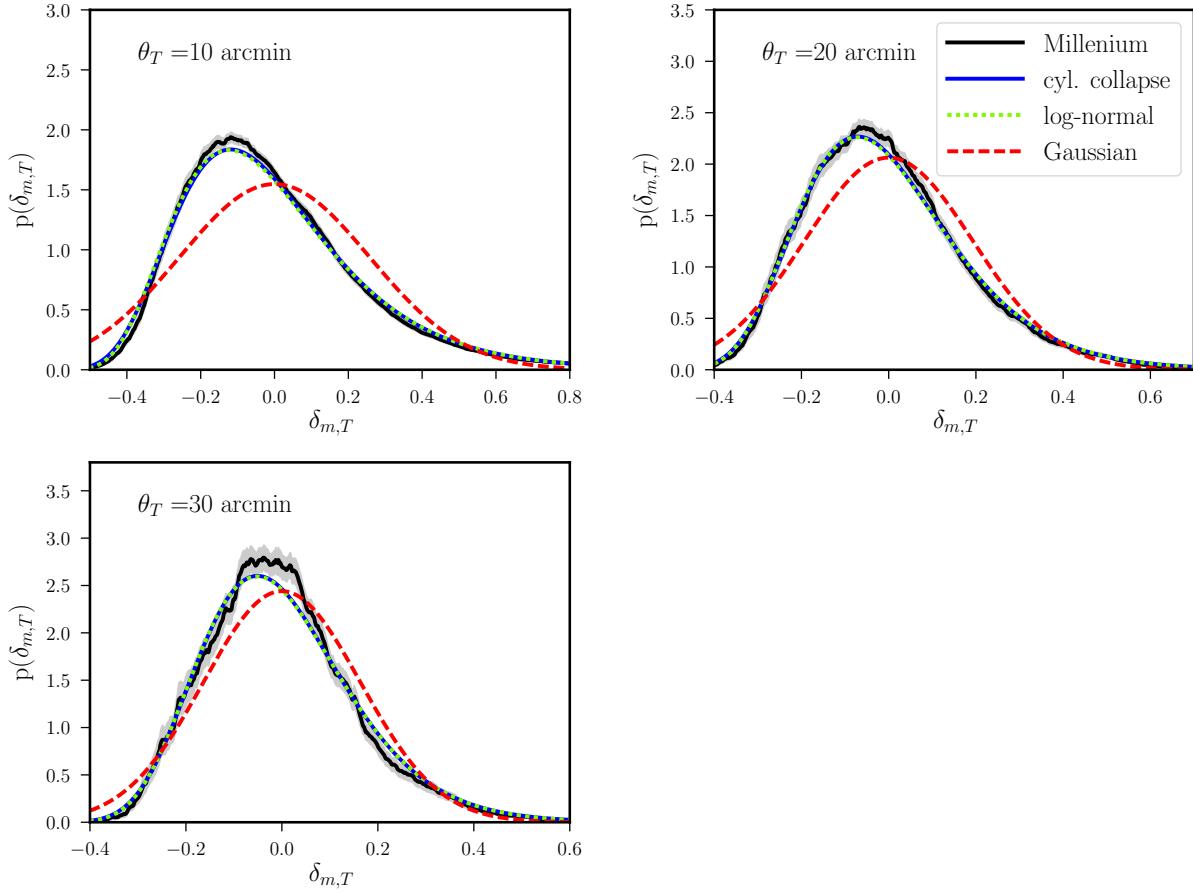


Figure 16.11: The PDF of projected density contrast $\delta_{m,T}$ in the Millennium Run (MR) compared to our model. In each plot, the black line shows a histogram of $\delta_{m,T}$ measured from 64 patches of $4 \times 4 \text{ deg}^2$ made from the MR by projecting the 3D density contrast with a constant selection function q_l between $0.19 \lesssim z \lesssim 0.43$, i.e. with a constant comoving density between those redshifts. The blue lines display the PDF predicted by our PT-motivated log-normal model, and the red lines show a Gaussian PDF with the same variance. The grey band is using the subsample covariance to estimate the error on the mean of all patches (Friedrich et al., 2016).

where n_m and n_g have a joint Gaussian distribution and $\delta_{g,0} = b\delta_{m,0}$. The variances of n_m and n_g are given by

$$\begin{aligned}\sigma_m^2 &= \ln \left\{ 1 + \frac{\text{Var}_m}{\delta_{m,0}^2} \right\} \\ \sigma_g^2 &= \ln \left\{ 1 + \frac{\text{Var}_g}{\delta_{g,0}^2} \right\} \\ &= \sigma_m^2\end{aligned}\tag{16.145}$$

and their covariance is given by

$$\begin{aligned}\xi_{mg} &= \ln \left\{ 1 + \frac{\text{Cov}_{mg}}{\delta_{m,0}\delta_{g,0}} \right\} \\ &= \ln \left\{ 1 + r \frac{\text{Var}_m}{\delta_{m,0}^2} \right\}.\end{aligned}\tag{16.146}$$

Let us denote the correlation coefficient of the Gaussian field by

$$\begin{aligned}\rho &= \frac{\xi_{mg}}{\sigma_m^2} \\ &= \frac{\ln \left\{ 1 + r \frac{\text{Var}_m}{\delta_{m,0}^2} \right\}}{\ln \left\{ 1 + \frac{\text{Var}_m}{\delta_{m,0}^2} \right\}}.\end{aligned}\tag{16.147}$$

Note that ρ will depend on scale even of b and r do not.

Now we want to compute the conditional moments $\langle \delta_{g,T} | \delta_{m,T} \rangle$ and $\langle \delta_{g,T}^2 | \delta_{m,T} \rangle$. First,

$$\begin{aligned}\langle e^{n_g} | n_m \rangle &= e^{\langle n_g | n_m \rangle + \sigma_g^2(1-\rho^2)/2} \\ &= e^{\rho(n_m + \sigma_m^2/2) - \sigma_g^2\rho^2/2} \\ &= e^{\sigma_g^2(\rho-\rho^2)/2} e^{\rho n_m}.\end{aligned}\tag{16.148}$$

Second,

$$\begin{aligned}\text{Var}(e^{n_g} | n_m) &= \left(e^{\sigma_g^2(1-\rho^2)} - 1 \right) e^{2\langle n_g | n_m \rangle + \sigma_g^2(1-\rho^2)} \\ &= \left(e^{\sigma_g^2(1-\rho^2)} - 1 \right) e^{\sigma_g^2(\rho-\rho^2)} e^{2\rho n_m}.\end{aligned}\tag{16.149}$$

Now what is $\text{Var}(\hat{N}|\delta_{m,T})$?

$$\begin{aligned}
 \langle \hat{N}^2 | \delta_{g,T} \rangle &= \langle \hat{N}^2 | n_m \rangle \\
 &= \int d\delta_{g,T} p(\delta_{g,T} | n_m) \times \\
 &\quad \times (\bar{N}[1 + \delta_{g,T}] + \bar{N}^2[1 + \delta_{g,T}]^2) \\
 &= \bar{N} + \bar{N}^2 + \int d\delta_{g,T} p(\delta_{g,T} | n_m) \times \\
 &\quad \times (\delta_{g,T}[\bar{N} + 2\bar{N}^2] + \bar{N}^2\delta_{g,T}^2) \\
 &= \bar{N} + \bar{N}^2 + [\bar{N} + 2\bar{N}^2]\langle \delta_{g,T} | n_m \rangle + \\
 &\quad + \bar{N}^2 (\text{Var}(\delta_{g,T} | n_m) + \langle \delta_{g,T} | n_m \rangle^2) \\
 \Rightarrow \text{Var}(\hat{N} | \delta_{m,T}) &= \bar{N}(1 + \langle \delta_{g,T} | n_m \rangle) + \bar{N}^2 \text{Var}(\delta_{g,T} | n_m) \\
 &= \bar{N}(1 + \langle \delta_{g,T} | n_m \rangle) + \bar{N}^2 \delta_{g,0}^2 \text{Var}(e^{n_g} | n_m)
 \end{aligned} \tag{16.150}$$

The probability $P(N_A | \delta_{m,T})$ can be computed in a similar way, by numerically evaluating

$$P(N_A | \delta_{m,T}) = \int d\delta_{g,T} p(\delta_{g,T} | \delta_{m,T}) P(N_A | \delta_{g,T}), \tag{16.151}$$

where $p(\delta_{g,T} | \delta_{m,T})$ can be computed from basic relations for joint log-normal random variables.

16.E Validation of alternative shot-noise model

In our data analysis (Gruen et al., 2018) we investigate both shot-noise parameterizations introduced in section 16.4.3. We hence check whether our alternative shot-noise parametrization, i.e. the one that uses three parameters to describe the relation between matter and galaxies (b , α_0 and α_1 , cf. section 16.4.3), recovers the true cosmology and shot-noise parameters of our mock data.

In figure 16.12 we show the posterior constraints derived for the two shot-noise parameters α_0 and α_1 , when marginalizing over different sets of model parameters. Our priors $0.1 < \alpha_0 < 3.0$ and $-1.0 < \alpha_1 < 4.0$ are mildly informative. We however expect that even stronger priors can be motivated (cf. our discussion in section 16.4.3) and will investigate this in future work. Figure 16.13 shows that our alternative shot-noise parametrization also recovers the correct Buzzard cosmology (cf. figure 16.10, which presents the same test for our baseline model).

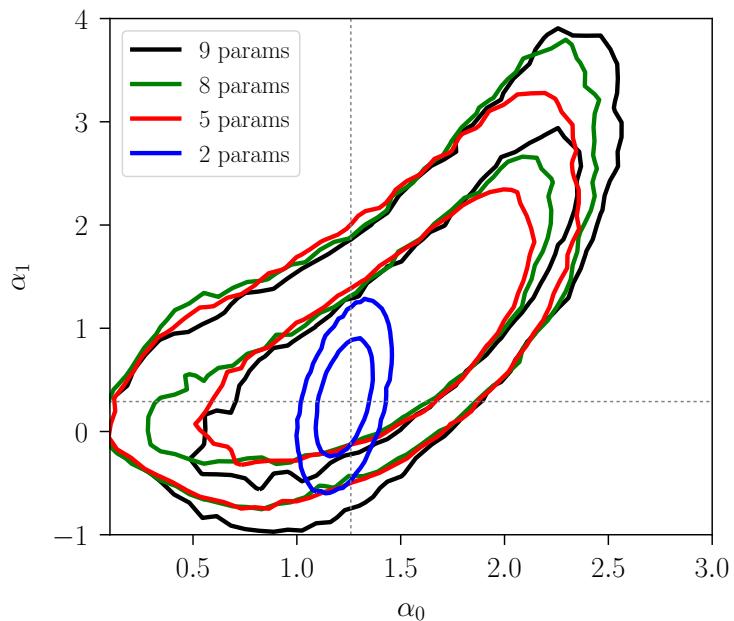


Figure 16.12: 1σ and 2σ contours in the α_0 - α_1 plane from a likelihood computed around the mean of 4 shape noise free realizations of DES Y1 (but assuming the full covariance matrix for a single DES Y1). The blue contour only varies α_0 and α_1 . The red contour marginalizes over Ω_m , σ_8 and galaxy bias b . The green contour additionally marginalizes over Ω_b , n_s , h_{100} , assuming the priors used by DES Collaboration et al. (2017). And the black contour also allows variation of the parameter $\Delta S_3/S_3$. Dotted lines show the values of α_0 and α_1 that were found to describe our mock data best in section 16.4.3.

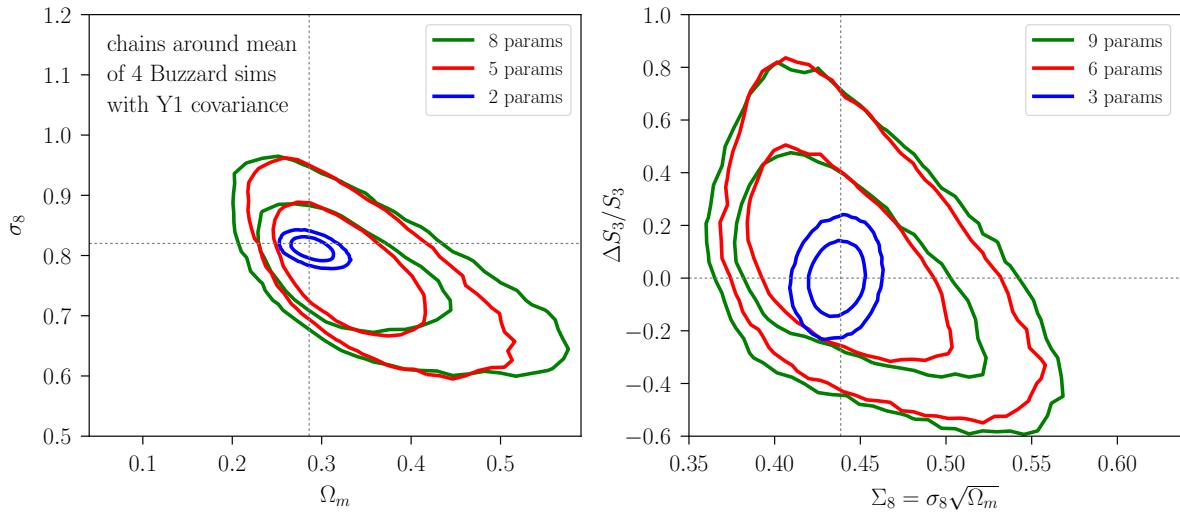


Figure 16.13: In analogy to figure 16.10, we test whether our alternative shot-noise parametrization can recover the Buzzard cosmology in a simulated likelihood analysis. Top panel: 1σ and 2σ contours in the Ω_m - σ_8 plane from a likelihood computed around the mean of 4 shape noise free realizations of DES Y1 (but assuming the full covariance matrix for a single DES Y1). The green contours are marginalized over Ω_b , n_s , h_{100} , REDMAGIC galaxy bias b as well as the shot-noise parameters α_0 and α_1 . For the parameters Ω_b , n_s , h_{100} we have assumed the same flat priors as used in the DES Y1 combined probes analysis presented in Abbott et al. (in prep.). The red contours are marginalized only over bias and shot-noise parameters and the blue contours only vary Ω_m and σ_8 . Even when going to this small parameter space, our model agrees with Buzzard within 1σ errors of DES Y1. Bottom panel: Same contours but in the Σ_8 - $\Delta S_3/S_3$ plane and varying one additional parameter, $\Delta S_3/S_3$. Dotted lines show the true Buzzard cosmology and our fiducial value of $\Delta S_3/S_3 = 0$.

Chapter 17

Density split statistics: cosmological constraints from counts and lensing in cells in DES Y1 and SDSS

Based on the modeling framework I developed in the previous chapter, we have analyzed year 1 data of the Dark Energy Survey (DES Y1) as well as data from the Sloan Digital Sky Survey (SDSS). This analysis has been published as Gruen, Friedrich, Krause, Cawthon, Davis, Elvin-Poole, Rykoff, Wechsler et al. (2018) in PRD (from here Gruen, Friedrich et al., 2018). In the following I will summarize the content of that paper and present its main results.

My own contributions to Gruen, Friedrich et al. (2018) are the following:

- I provided the code used to model the measured signals. In that context I also wrote parts of section III.
- I computed the input power spectra and the so-called log-normal shift parameters (Xavier et al., 2016) that were required to generate the log-normal simulations used in this work. In that context I also wrote parts of appendix B.
- I contributed to general discussion and proofreading.

17.1 Analysis overview

The main steps of the analysis presented in Gruen, Friedrich et al. (2018) are the following:

1. Preparing a sample of lens galaxies and a set of source galaxies from DES Y1 and SDSS data. In particular we used the REDMAGiC sample available for both surveys (Rozo et al., 2016) as lenses and shape catalogs obtained with METACALIBRATION (Sheldon & Huff, 2017) as sources.

2. Using the lens sample to create a map of galaxy density, smoothed by a circular top-hat radius. For this we used an angular radius of $\theta_T = 20'$ because this was the smallest radius at which we could successfully validate our model in Friedrich, Gruen et al., 2018.
3. Splitting the sky into density quantiles (based on the above map) and measuring the density split lensing signals. Also, measuring the Counts-in-Cells histogram (the CiC histogram, which is equivalent to the galaxy density PDF).
4. Generating log-normal realizations of the projected matter density field as well as the lensing convergence field (using the FLASK tool by Xavier et al., 2016) in order to
 - estimate the cosmic variance part of the covariance matrix of our signals (i.e. the part of the covariance that is shape noise free),
 - validate the likelihood pipeline with which we constrain the different parameters of our model.
5. Computing the shape noise part of our covariance matrix from random rotations of the ellipticities of our source galaxies. Also, estimating a covariance matrix with the jackknife method as a cross check.
6. Validating our likelihood pipeline wrt. the above mentioned log-normal simulations as well as wrt. an extended set of N-body realizations (that where not yet available when we prepared Friedrich, Gruen et al., 2018).
7. Testing whether changes in our analysis choices significantly impact our results. In particular the following changes to our analysis were investigated:
 - using a shape catalog from an alternative shape measurement algorithm,
 - excluding the lowest and highest redshift bin from the analysis,
 - using a different photometric redshift estimation procedure to estimate our redshift distributions,
 - changing our angular scale cuts,
 - neglecting a systematic effect that arises when lenses and sources overlap in redshift (see section III.D of the paper for details).
8. Deriving cosmological parameter constraints. In particular, we obtained information about the following aspects of the large scale structure of the universe:
 - the bias and stochasticity of redMaGiC galaxies wrt. the total matter density field,
 - the skewness of the matter density PDF,

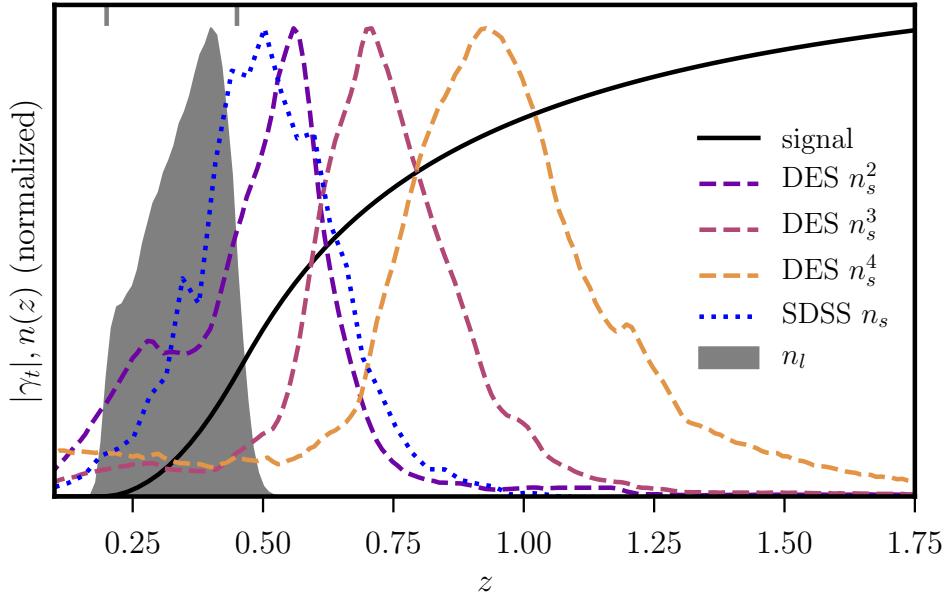


Figure 17.1: Taken from Gruen, Friedrich et al. (2018). The gray, shaded area shows the redshift distribution of our lens sample. This is almost identical for the DES and SDSS parts of the analysis. The dashed lines (purple, red and yellow) show the redshift distributions of the 3 source galaxy bins used in the DES analysis. The blue dotted line shows the redshift distribution of the single source galaxy bin used for the SDSS analysis. The black line indicates how the amplitude of our lensing signals increases with redshift. The fact that some of the source bins overlap in redshift with our lens sample has to be accounted for with a new nuisance parameter - see the discussion in section III.D of Gruen, Friedrich et al. (2018).

- the parameters Ω_m and σ_8 of the Λ CDM model that was underlying our analysis.

The results of our analysis are robust wrt. the tested changes in the analysis setup, indicating that they are also robust wrt. a wide range of systematic uncertainties. Also, when applying our likelihood pipeline to simulated data we recovered parameter constraints that are consistent wrt. the input models of that data.

Our fiducial analysis was carried out using one lens redshift bin and overall 3 source redshift bins of DES Y1. SDSS data was analyzed mainly as a consistency check for the DES analysis. Because SDSS has a lower depth than DES we only used one source redshift bin in the SDSS part of the analysis. Figure 17.1 shows the redshift distribution of the different samples and figure 17.2 shows the distribution of galaxy counts in apertures of $\theta_T = 20'$ radius that we found in DES and SDSS resp.

In figure 17.3 I show the density split lensing signals obtained from one of our source bins. The signals obtained for the other DES source bins as well as for SDSS can be found

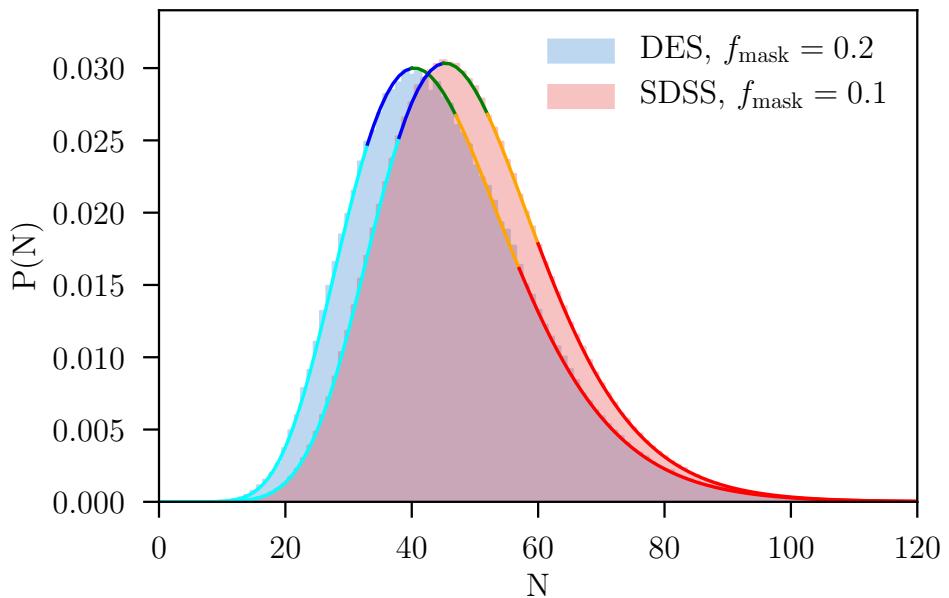


Figure 17.2: Taken from Gruen, Friedrich et al. (2018). The blue histogram shows the distribution of REDMAGiC galaxy counts found within apertures of $\theta_T = 20'$ radius in DES Y1 (the CiC histogram). The red histogram shows the corresponding distribution found in SDSS. The DES and SDSS REDMAGiC samples used in our analysis had almost identical redshift distributions. The reason why the SDSS sample nevertheless has a larger number density, is that DES Y1 has a larger masking fraction due to holes in the survey mask. See section II.A of Gruen, Friedrich et al. (2018) for how we treated the masking fraction in our analysis. The solid lines show our best fit model for the CiC histograms and the color coding thereof indicates the 5 density quantiles into which we split the survey areas (cf. figure 16.1).

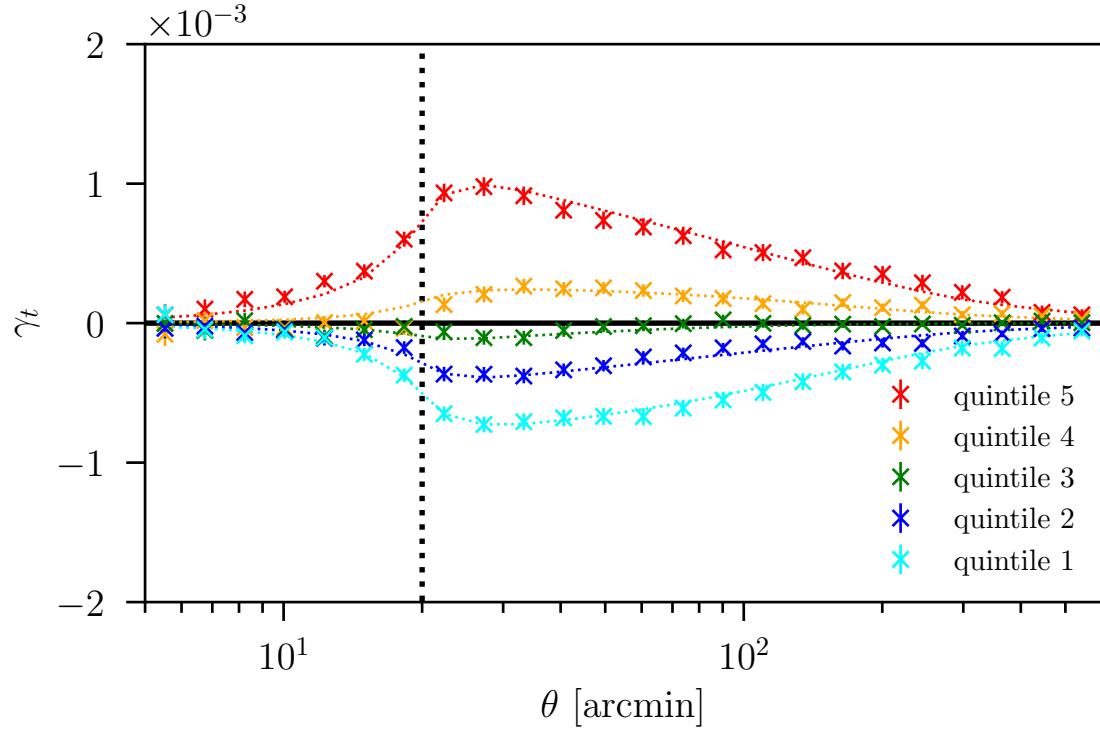


Figure 17.3: Taken from Gruen, Friedrich et al. (2018). The points show our measurements of the tangential shear profile $\gamma_t(\theta)$ around REDMAGIC density quantiles in DES Y1 as measured with our intermediate source redshift bin (from the most underdense quantile in cyan to the most overdense quantile in red). The dotted line indicate our best fit model.

in the paper. Both figure 17.2 and figure 17.3 demonstrate that our best fit model is able to describe the measured signals of our density split analysis. The best-fit values of our fiducial model give a reduced χ^2 of

$$\frac{\chi^2_{\text{best fit}}}{\text{degrees of freedom}} \approx 0.88 \quad (17.1)$$

with expected uncertainty

$$\sigma \left\{ \frac{\chi^2_{\text{best fit}}}{\text{degrees of freedom}} \right\} \approx 0.10 . \quad (17.2)$$

We will investigate in future work, whether our log-normal simulations slightly overestimate the covariance matrix.

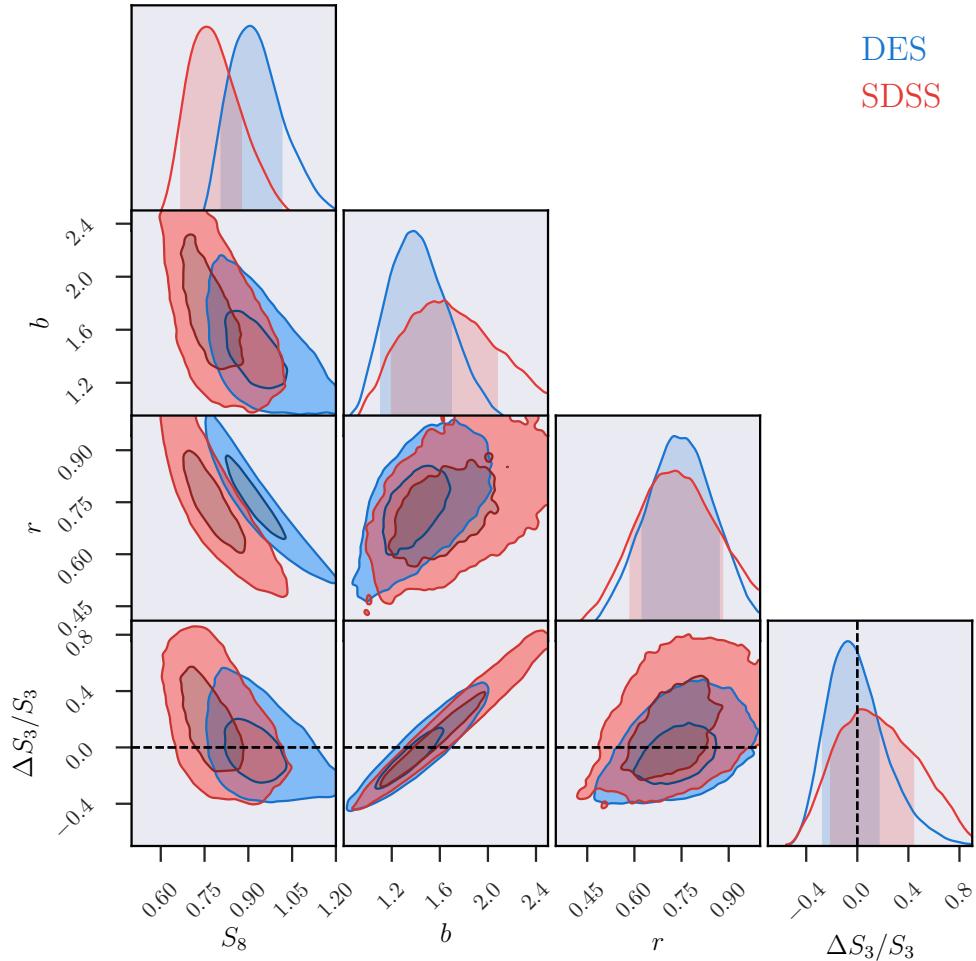


Figure 17.4: Taken from Gruen, Friedrich et al. (2018). Constraints on galaxy bias b , galaxy stochasticity r , the parameter S_3 that describes the scaling between 2nd and 3rd moments of the density field and the Λ CDM parameters Ω_m and σ_8 (i.e. the fraction of matter in the total energy content of the universe and the amplitude of density fluctuations) which are here compressed into the parameter $S_8 = \sigma_8 \sqrt{\Omega_m}$ that is best constrained by cosmic shear 2-point statistics. The ratio $\Delta S_3/S_3$ describes the relative deviation of S_3 wrt. its Λ CDM value.

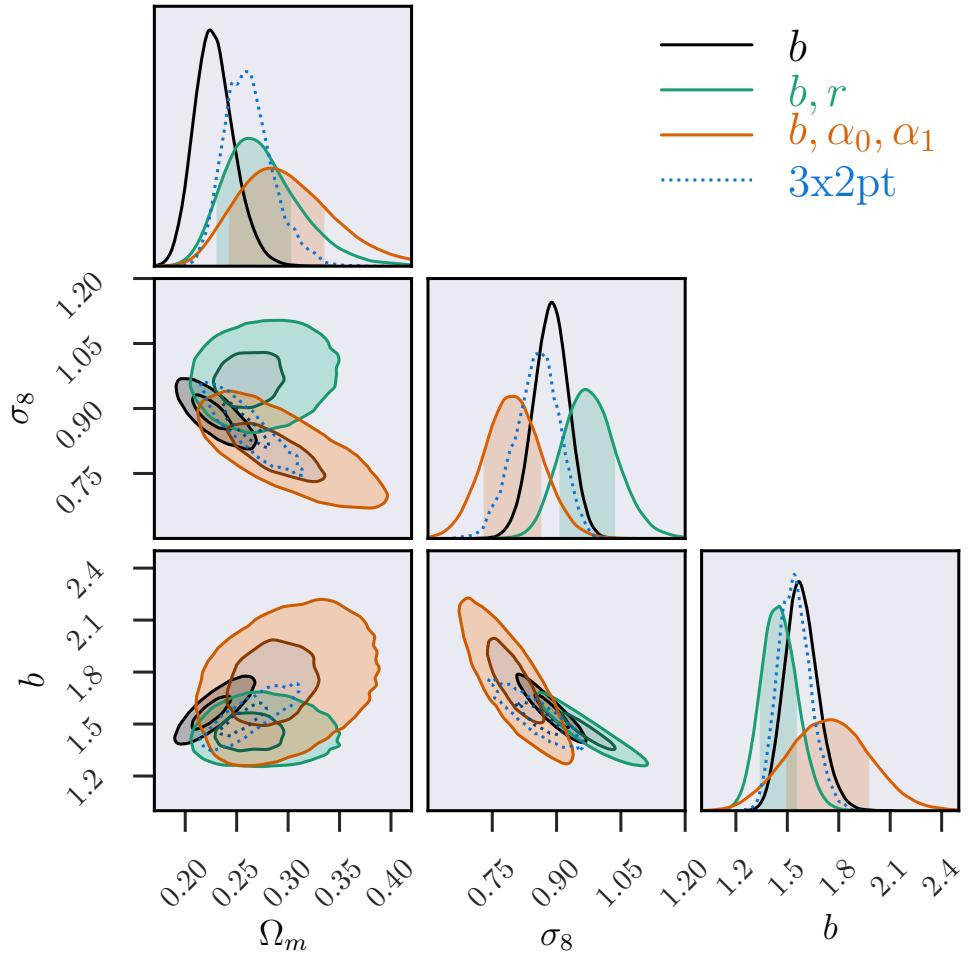


Figure 17.5: Taken from Gruen, Friedrich et al. (2018). Constraints on Ω_m and σ_8 and galaxy bias b when fixing the scaling parameter S_3 to its Λ CDM value. The black contours are obtained with a model that assumed linear galaxy bias and Poisson shot noise of galaxies wrt. matter (model 1 in the main text). The green contours are obtained from a model that allows for an additional stochasticity between galaxies and matter beyond Poisson noise (model 2 in the main text). The red contours are obtained with a model that allows for an arbitrary slope of the density dependence in the shot noise amplitude (model 3 in the main text). For comparison we show the contours of the DES Y1 2-point analysis (DES Collaboration et al., 2017) in blue.

17.2 Constraints on the skewness of the density PDF

During the validation of our analysis we have employed a blinding strategy to avoid confirmation bias. This strategy is described in section V.D of the paper. Once the validation was successful we unblinded and performed a cosmological analysis of the measured data vectors based on the model presented in the previous chapter. Our parameter constraints were obtained in the framework of Bayesian parameter inference (cf. section 12.3; also, see table I of the paper for our parameter priors). To compare different models we compute the so-called *Bayes factors* or *Bayesian evidence ratios*. Recall from equation 12.8 that the posterior distribution $p(\boldsymbol{\pi}|\hat{\mathbf{x}})$ of the model parameters $\boldsymbol{\pi}$ given the measured data vector $\hat{\mathbf{x}}$ is computed through

$$p(\boldsymbol{\pi}|\hat{\mathbf{x}}) = \frac{p(\hat{\mathbf{x}}|\boldsymbol{\pi}) \text{pr}(\boldsymbol{\pi})}{\mathcal{N}} , \quad (17.3)$$

where $p(\hat{\mathbf{x}}|\boldsymbol{\pi})$ is the PDF of $\hat{\mathbf{x}}$ given the parameters $\boldsymbol{\pi}$ and $\text{pr}(\boldsymbol{\pi})$ is the prior distribution assumed for the parameters $\boldsymbol{\pi}$. The factor \mathcal{N} normalizes $p(\boldsymbol{\pi}|\hat{\mathbf{x}})$ such that its integral over the parameters is 1. \mathcal{N} is also called the *evidence*. Two models 1 and 2 will have different evidences \mathcal{N}_1 and \mathcal{N}_2 and their ratio

$$E = \frac{\mathcal{N}_1}{\mathcal{N}_2} \quad (17.4)$$

is the Bayes factor. A Bayes factor close to 1 is supposed to indicate that both models describe the data equally well. A Bayes factor $\ll 1$ is supposed to indicate that model 2 should be favored over model 1. We investigate the behavior of E when comparing different versions of our model with the help of our log-normal simulations. This is described in appendix D of the paper and it confirms the above interpretation.

In figure 17.4 I show the constraints obtained on parameters of our fiducial model when allowing that the skewness of the density PDF deviated from its Λ CDM value. This figure demonstrates the claim made previously that density split statistics can simultaneously provide information about 3 different aspects of the large scale structure of the universe:

- the relation of galaxies and dark matter (in the form of galaxy bias b and galaxy stochasticity r),
- the skewness of the cosmic density PDF (in the form of the scaling parameter S_3),
- the parameters Ω_m and σ_8 of the Λ CDM model (i.e. the fraction of matter in the total energy content of the universe and the amplitude of density fluctuations). In figure 17.4 these parameters have been compressed into the combined parameter $S_8 = \sigma_8 \sqrt{\Omega_m/0.3}$.

Figure 17.4 shows that DES Y1 data is indeed consistent with the Λ CDM predictions of the skewness of the density field. Fixing the scaling parameter S_3 to its Λ CDM prediction will then improve the constraints on our other model parameters.

17.3 Comparing different models of the matter-galaxy relation

Figure 17.5 shows the parameter constraints we obtain from DES Y1 when fixing S_3 in our analysis. This figure also shows how our parameter constraints change if we use different models for the relation of galaxies and matter. These models are:

1. Linear bias + Poisson noise:

a model in which we assume a linear relation $\langle \delta_g \rangle = b \cdot \delta_m$ between galaxy density contrast and matter density contrast and a Poissonian shot noise (without any additional stochasticity).

2. Linear bias + stochasticity:

a model that allows for stochasticity beyond Poisson noise.

3. Linear bias + density dependent stochasticity:

a model that allows for an arbitrary slope of the density dependence in the shot noise amplitude.

These models have been discussed in detail in section 16.4.3. Figure 17.5 shows that the exact choice of our bias and stochasticity model indeed impacts our cosmological constraints. An analysis based on Bayesian evidence ratios implies that these differences are not significant (see section VI.A of the paper). Nevertheless, this results is surprising, considering the fact that our data from N-body simulations was equally well described by the model 2 and model 3 of the enumeration above.

The main difference of our findings in the data wrt. our N-body simulations is the following: the shot noise between matter and galaxies seem to be a steeper function of density in DES Y1 data than is in our simulated data. To understand this see figure 16.8, where both model 2 and model 3 describe the shot noise of galaxy counts as a function of matter density equally well. In our data however, model 3 prefers this function to be steeper than allowed by the parametrization of model 2. We are still investigating whether these findings are significant. Tables 17.1 and 17.2 summarizes the results of our analysis.

Table 17.1: Taken from Gruen, Friedrich et al. (2018). The table shows constraints on our model parameters when fixing the scaling coefficient S_3 to its Λ CDM prediction. See sections V.C and VI.A of the paper for a discussion of Bayes factors. The last row shows corresponding constraints obtained from the analysis of 2-point statistics in DES Y1 (DES Collaboration et al., 2017).

| Data | Model | Bayes factor | S_8 | Ω_m | σ_8 | b | r | α_0 | α_1 |
|-----------------------------------|-------------------------|--------------|------------------------|------------------------|------------------------|------------------------|---------------------------------|---------------------|---------------------|
| DES | b, r | $\equiv 1.0$ | $0.90^{+0.11}_{-0.08}$ | $0.26^{+0.04}_{-0.03}$ | $0.97^{+0.07}_{-0.06}$ | $1.45^{+0.10}_{-0.11}$ | $0.77^{+0.10}_{-0.13}$ | - | - |
| SDSS | b, r | $\equiv 1.0$ | $0.78^{+0.13}_{-0.08}$ | $0.25^{+0.05}_{-0.04}$ | $0.86^{+0.06}_{-0.05}$ | $1.48^{+0.09}_{-0.09}$ | $0.70^{+0.16}_{-0.14}$ | - | - |
| DES | b, α_0, α_1 | 0.7 | $0.78^{+0.05}_{-0.04}$ | $0.28^{+0.05}_{-0.04}$ | $0.80^{+0.06}_{-0.07}$ | $1.75^{+0.22}_{-0.26}$ | - | $1.5^{+0.4}_{-0.6}$ | $1.7^{+1.1}_{-0.9}$ |
| SDSS | b, α_0, α_1 | 1.6 | $0.76^{+0.08}_{-0.07}$ | $0.28^{+0.07}_{-0.05}$ | $0.80^{+0.08}_{-0.11}$ | $1.18^{+0.37}_{-0.23}$ | - | $2.3^{+0.3}_{-0.5}$ | $2.9^{+1.1}_{-1.0}$ |
| $3 \times 2\text{pt, fixed } \nu$ | | | $0.80^{+0.02}_{-0.02}$ | $0.26^{+0.02}_{-0.03}$ | $0.85^{+0.06}_{-0.05}$ | $1.54^{+0.09}_{-0.10}$ | DES Collaboration et al. (2017) | | |

Table 17.2: Taken from Gruen, Friedrich et al. (2018). The table shows constraints on our model parameters when varying the scaling coefficient S_3 ($\Delta S_3/S_3$ is the relative deviation from its Λ CDM prediction). See sections V.C and VI.A of the paper for a discussion of Bayes factors. The latter have been computed wrt. to the model where S_3 is fixed to its Λ CDM prediction.

| Data | Model | Bayes factor | S_8 | Ω_m | σ_8 | b | r | α_0 | α_1 | $\Delta S_3/S_3$ |
|------|-------------------------------------|--------------|------------------------|------------------------|------------------------|------------------------|------------------------|---------------------|---------------------|-------------------------|
| DES | $b, r, \Delta S_3$ | 0.3 | $0.91^{+0.10}_{-0.10}$ | $0.26^{+0.07}_{-0.05}$ | $0.96^{+0.17}_{-0.13}$ | $1.37^{+0.32}_{-0.27}$ | $0.72^{+0.14}_{-0.10}$ | - | - | $-0.08^{+0.25}_{-0.20}$ |
| SDSS | $b, r, \Delta S_3$ | 0.4 | $0.76^{+0.12}_{-0.09}$ | $0.28^{+0.09}_{-0.07}$ | $0.72^{+0.17}_{-0.13}$ | $1.64^{+0.44}_{-0.46}$ | $0.73^{+0.15}_{-0.15}$ | - | - | $0.06^{+0.40}_{-0.27}$ |
| DES | $b, \alpha_0, \alpha_1, \Delta S_3$ | 0.3 | $0.80^{+0.06}_{-0.05}$ | $0.26^{+0.07}_{-0.05}$ | $0.86^{+0.10}_{-0.13}$ | $1.48^{+0.41}_{-0.32}$ | - | $1.7^{+0.4}_{-0.6}$ | $2.0^{+1.1}_{-0.9}$ | $-0.18^{+0.25}_{-0.22}$ |
| SDSS | $b, \alpha_0, \alpha_1, \Delta S_3$ | 0.6 | $0.78^{+0.07}_{-0.08}$ | $0.27^{+0.09}_{-0.05}$ | $0.80^{+0.14}_{-0.12}$ | $0.98^{+0.49}_{-0.17}$ | - | $2.5^{+0.3}_{-0.5}$ | $4.0^{+0.0}_{-1.9}$ | $-0.14^{+0.44}_{-0.39}$ |

Part V

Summary and open questions

Chapter 18

Contributions of my thesis to the field of large scale structure cosmology

In cosmological analyses, higher order moments of the density field (i.e. moments of order > 2) play a 2-fold role: they are used to describe the covariance matrix of 2-point correlation functions and power spectra and they can serve as cosmological probes by themselves (cf. chapter 3). With this thesis I helped to improve our understanding of higher order moments (and higher order statistics in general, such as the cosmic density PDF) in both of these contexts. Since there have already been separate concluding sections in chapters 13, 14 and 16, I will not interpret the results of the corresponding papers again here. Instead, I will summarize the original contributions of this thesis in the following and outline a number of open questions pointing to future work in chapters 19, 20 and 21.

The original contributions of this work can be summarized as follows:

- In Friedrich et al. (2016) I have for the first time systematically analyzed the performance of internal covariance estimators for the covariance matrix of cosmic shear correlation functions. This includes a detailed investigation of different ways to treat galaxy pairs between the different subregions used for the internal re-sampling techniques. That is also relevant for other 2-point correlation functions of the cosmic density field and has been lacking in the most cited study of internal covariance estimators for 2-point correlation functions (Norberg et al., 2009). The main conclusion of Friedrich et al. (2016) is that parameter uncertainties in the $\Omega_m - \sigma_8$ plane are underestimated by 10 – 15% when using internal covariance estimates in a non-tomographic cosmic shear analysis (see section 13.6 for the detailed conclusions of that paper).
- In Friedrich & Eifler (2018) I have developed a completely new estimator for the inverse covariance matrix (the *precision matrix*) that allows the integration of a priori knowledge about the covariance matrix into matrix estimates from a set of simulated data. This method derives corrections to a modeled precision matrix in

256. Contributions of my thesis to the field of large scale structure cosmology

the form of a power series (see equation 14.10) and estimates the terms of that series from simulated data. In Friedrich & Eifler (2018) we show that this method can correct severe deficits in a covariance model. Also, we show that this method decreases the number of (e.g. N-body) simulations required for a good estimate of the precision matrix by a factor between 10 and 100. Here *good* means that the fraction of analyses in which the true cosmological parameters lie within a 68% confidence region derived from the estimated precision matrix should be between 67% and 69%. The detailed conclusions of that paper are given in section 14.5.

- My part in Gruen, Friedrich et al., 2016 was to develop a model for the newly discovered *trough lensing* effect based on the assumptions that
 - a) both the matter density contrast and the lensing convergence are Gaussian random fields
 - b) the relations between the matter density contrast and the density contrast of luminous red galaxies can be described with a linear bias model + Poissonian shot-noise.

Using this model we found hints in early DES data that there is either a skewness in the distribution of the matter density field or that there is a density dependence in the relation between galaxies and matter. (In Friedrich, Gruen et al., 2018; Gruen, Friedrich et al., 2018, we found that both hypotheses are correct.)

- In Friedrich, Gruen et al., 2018 I developed an improved model for *trough lensing* and its generalization *density split statistics*. This model uses an approach based on cosmological perturbation theory and spherical collapse to derive the PDF of the line-of-sight projected matter density contrast and to predict the relation between that density contrast and the lensing convergence field. I also improved our modeling of the relation between matter density and galaxy density by incorporating a stochastic noise between the two that can deviate from Poisson noise and whose amplitude can be density dependent. Using a set of high-resolution N-body simulations (see DeRose et al., 2018, for the details of those simulations) we showed that this model is accurate enough to perform a cosmological analysis on DES year 1 data.

The paper Gruen, Friedrich et al., 2018 presents such an analysis based on the above model. This is the first cosmological analysis of the low-redshift universe based on the PDF of the cosmic density field. In that study we were able to confirm Λ CDM predictions for the skewness of the matter density field and we also showed (with a mild, 2σ significance) that there is a stochastic bias between the matter density field and the density of luminous red galaxies on scales of $\lesssim 10$ Mpc.

Chapter 19

Open questions in covariance estimation

In this chapter I give a review of current open questions in the estimation and use of covariance matrices as well as an outlook on future work.

19.1 Non-Gaussian likelihoods

As discussed in chapter 12, cosmological analyses often assume that the statistical uncertainties of measured data vectors have a Gaussian distribution (Kilbinger et al., 2013; Heymans et al., 2013; Abbott et al., 2016; Hildebrandt et al., 2017; van Uitert et al., 2018; DES Collaboration et al., 2017). This mean that if $\hat{\mathbf{x}}$ is a measured data vector with expectation value $\boldsymbol{\mu}$ and covariance matrix \mathbf{C} it is often assumed that the PDF of $\hat{\mathbf{x}}$ is given by

$$p(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^D |\mathbf{C}|}} \exp \left\{ -\frac{1}{2} [\mathbf{x} - \boldsymbol{\mu}]^T \cdot \mathbf{C}^{-1} \cdot [\mathbf{x} - \boldsymbol{\mu}] \right\} . \quad (19.1)$$

It has recently been argued by Sellentin & Heavens (2018); Sellentin et al. (2018) that this assumption is not sufficiently accurate for the distribution of measurements of cosmic shear correlation functions. To explain their argument, let us consider any random field $f(\hat{\mathbf{n}})$, where $\hat{\mathbf{n}}$ is a unit vector pointing on a location of the sky (f might, e.g., be the lensing convergence field or the projected galaxy density contrast in the sky). A particular realization \hat{f} of f can be expanded into spherical harmonics $Y_{\ell m}(\hat{\mathbf{n}})$ as

$$\hat{f}(\hat{\mathbf{n}}) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} \hat{a}_{\ell m} Y_{\ell m}(\hat{\mathbf{n}}) . \quad (19.2)$$

If f is a Gaussian random field with vanishing expectation value, then the coefficients $a_{\ell m}$ are also Gaussian random variable with vanishing expectation value. Given a particular realization \hat{f} the angular power spectrum C_{ℓ}^f of f can be estimated from the coefficients

$\hat{a}_{\ell m}$ as (Peebles, 1993)

$$\hat{C}_{\ell}^f = \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} |\hat{a}_{\ell m}|^2. \quad (19.3)$$

Now, even if the $a_{\ell m}$ are Gaussian random variable, the above estimator of C_{ℓ}^f is clearly not since it is a sum of squared Gaussian random variables. Instead, \hat{C}_{ℓ}^f has a Gamma distribution (Sellentin et al., 2018). For large values of ℓ this is not a problem because the sum in the definition of \hat{C}_{ℓ}^f will be over many identically distributed random variables (that are also independent, if the field f is isotropic) and the central limit theorem will nevertheless lead to an approximately Gaussian distribution (Anderson, 2003). But for low values of ℓ estimators like the one in 19.3 will show significantly skewed distributions (see, e.g. , Planck Collaboration et al., 2014).

Sellentin et al. (2018) argue that any estimator of a particular 2-point correlation function $\xi^f(\theta)$ can be approximated as

$$\hat{\xi}(\theta) \approx \sum_{\ell} F_{\ell}(\theta) \hat{C}_{\ell}^f, \quad (19.4)$$

where the coefficients $F_{\ell}(\theta)$ depend on the particular type of correlation function (appendix B gives those coefficients for the correlations functions $\xi_{\pm}(\theta)$, $\gamma_t(\theta)$ and $w(\theta)$). The central limit theorem will again lead to an approximately Gaussian distribution for $\hat{\xi}$ when the sum in 19.4 gives similar weight to a lot of similarly distributed modes \hat{C}_{ℓ}^f . But if the coefficients $F_{\ell}(\theta)$ peak around a small number of modes, then one expects a significant non-Gaussianity. This is typically the case on scales comparable to the area on the sky that is observed by a particular survey and Sellentin & Heavens (2018); Sellentin et al. (2018) argue that this effect is indeed significant for current cosmic shear surveys since they are falsely approximating the skewed distribution of $\hat{\xi}$ with a symmetric, Gaussian one. However, the results of Sellentin et al. (2018) were derived for a survey comparable to the CFHTLenS survey (e.g. Kilbinger et al., 2013) whose area is ~ 10 times smaller than the area of DES year 1 and ~ 30 times smaller than that of the final DES survey. The DES collaboration is currently investigating whether the findings of Sellentin et al. (2018) affect such a large survey in a significant manner.

Even if non-Gaussian distribution should be found to have a significant impact on cosmological parameters constraints, any non-Gaussian ansatz for the data likelihood will most likely depend on the covariance matrix of the measured 2-point statistics (see, e.g. , Sellentin & Heavens 2016 and the appendix Hilbert et al. 2011 for how to construct log-normal PDFs from the corresponding Gaussian ones).

19.2 Covariance estimation for analyses that combine different cosmological probes

Current DES data has been analyzed with a number of different probes to constrain cosmological parameters. These are, e.g. , 2-point statistics of the shear and galaxy density

field (DES Collaboration et al., 2017), shear peak statistics (Kacprzak et al., 2016), density split statistics (Friedrich et al., 2018; Gruen et al., 2018) or the counts of galaxy clusters as a function of their mass (DES in prep.; see McClintock et al. 2018 the calibration of DES cluster masses using their lensing profile). To put cosmological theories to the most stringent test possible, it is desirable to combine several (or all) of these probes into a joint cosmological analysis. To do so, a joint covariance matrix for the data vectors obtained from all the combined probes is required.

It is unlikely, that the number of N-body simulations in which all of these probes can be obtained to a sufficient accuracy will be large enough to obtain a good covariance estimate for these combined probes (cf. MacCrann et al., 2018, where even for the DES 2-point analysis alone only 18 simulations could be considered). Unfortunately, also the modeling of such a combined probe covariance is very challenging (Krause & Eifler, 2016). But there are many subparts of such a covariance matrix that can be estimated from separate techniques:

- The small scale cosmic variance of 2-point functions can still be estimated from N-body simulations -, e.g. , through the so-called response approach (Barreira & Schmidt, 2017) or by cutting large N-body boxes into smaller sub-volumes that still capture the important modes.
- The part of the cluster lensing covariance that is due to intrinsic variations of the cluster mass profile can be estimated from the cluster-to-cluster variations in high resolution N-body simulations.
- Shape noise contributions to the covariance of shear peak statistics or cluster lensing profiles can be obtained by randomly rotating the orientation of galaxies (Kacprzak et al., 2016; McClintock et al., 2018).
- Shape noise and shot noise contributions to the covariance of 2-point functions can be estimated from the data itself (see appendix C for a demonstration of this approach). This could be especially useful in the light of Troxel et al. (2018), who found that these parts have been the most troublesome in covariance models used for recent 2-point analyses.
- Effects of masking and survey geometry on the large scale part of the 2-point function covariance can be estimated from (cheap) log-normal realizations of the density field (Xavier et al., 2016; Krause et al., 2017).

It is straight forward to generalize the PME method presented in chapter 14 (based on Friedrich & Eifler 2018) to the case where several different parts of the covariance matrix are estimated from different sources. To do so one would generalize the formalism derived in section 14.3 to the case where the covariance matrix \mathbf{C} is split into a number of contributions as

$$\mathbf{C} = \mathbf{A} + \mathbf{B}_1 + \dots + \mathbf{B}_N , \quad (19.5)$$

where \mathbf{A} is the part of the covariance model that is trusted and $\mathbf{B}_1, \dots, \mathbf{B}_N$ are parts of the covariance for which we have independent estimates $\hat{\mathbf{B}}_1, \dots, \hat{\mathbf{B}}_N$. An approach that integrates different covariance estimates, e.g. , in this way should be constructed in preparation for the upcoming multi-probe analyses.

Chapter 20

Open questions and tasks for density split statistics

The main benefit of density split statistics (DSS) is that it simultaneously probes the relation between the galaxy and matter density field as well as gravitational non-linearities (i.e. the evolution of higher order moments of the density field). There are several ways how to improve and extend the analysis of Friedrich, Gruen et al., 2018; Gruen, Friedrich et al., 2018 in order to make full use of these 2 pieces of information.

20.1 Improving the analysis of matter-galaxy relation

To gain a better understanding of the relation between matter and galaxies, the analysis should be extended in the following ways:

- We should study DSS at different smoothing scales θ_T in order to probe the scale dependence of galaxy bias and galaxy stochasticity.
- We should study DSS in several different lens redshift bins to study the redshift evolution of those parameters.
- We should perform the density splits in different lens galaxy samples to compare the relation of matter and galaxies for different galaxy types.

To interpret the information gained by such an extended study, a major improvement of our theoretical modeling must be achieved:

- We have to understand the relation between the mere nuisance parameters b (galaxy bias) and r (galaxy stochasticity) and the more physical parameters from the framework of halo occupation distributions (HOD; see, e.g. , Cacciato et al., 2012; Park et al., 2016; Dvornik et al., 2018).

Determining HOD parameters for different types of galaxies will enable us to infer the different environments in which these galaxies reside (resp. the different masses of typical

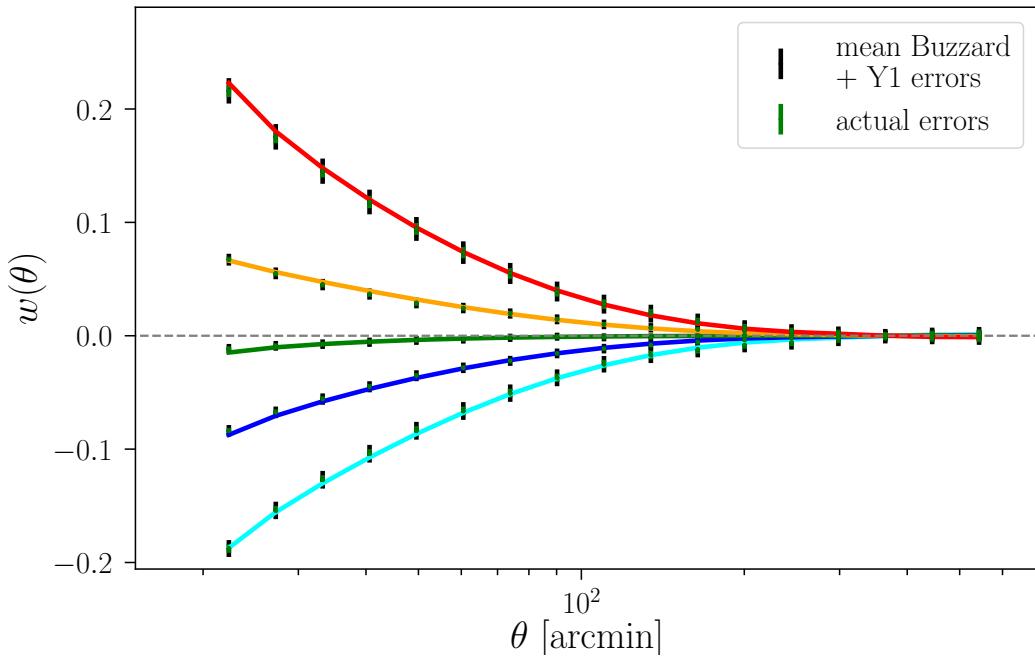


Figure 20.1: Instead of measuring the lensing profile around different quantiles of galaxy density we can also measure the galaxy clustering profile around those quantiles. As has been shown in the context of 2-point statistics (Park et al., 2016), this yields complementary information on the matter-galaxy relation and can improve our ability to disentangle non-linearities in that relation from gravitational non-linearities. The figure shows the average of measurements in 4 DES year 1 like patches from the Buzzard N-body simulations (DeRose et al., 2018) and can be directly compared to the corresponding lensing measurements in Friedrich et al. (2018) - cf. figure 16.1. The solid lines show theoretical predictions along the same lines as for the density split lensing signals presented in Friedrich et al. (2018).

host halos for these galaxies). This will allow us to study galaxy evolution in a cosmological context.

20.2 Improving the analysis of gravitational non-linearities

To gain a better understanding of gravitational non-linearities the analysis should be extended in the following ways:

- We should analyze density splits at larger scales and higher redshifts to study the onset of non-linear evolution.
- In addition we should analyze density splits on small, highly non-linear scales to study the end product of that evolution.

- We should increase the number of density quantiles used to perform the density splits and put focus not only on the moments of the density PDF but on the shape of its high and low density tails (see, e.g. , Brouwer et al., 2018, who extend our analysis to a finer binning in density).
- Instead of measuring only the lensing signals around different density quantiles we should also investigate the galaxy clustering profile around these quantiles. This would allow for an even better separation of gravitational non-linearities and non-linearities in the matter-galaxy relation (see, e.g. , Park et al., 2016, who do this in the context of 2-point statistics). I demonstrate in figure 20.1 that our model can be extended to these profiles.

To make full use of this extended information, several improvements of our theoretical modeling must be achieved:

- We have to improve our understanding of the moments of the density field. Right now, we approximate the cumulant generating function (CGF) of the density field with the rescaled tree level CGF that was described in chapters 7 and 16. This means we compute the tree level CGF as

$$\varphi_R^{\text{tree}}(y, \eta) = -S_y[\delta_L^*] \quad (20.1)$$

and then compute the non-linear CGF as a rescaling thereof (cf. section 7.3), i.e.

$$\varphi_R^{\text{exact}}(y, \eta) \approx \frac{\langle \delta_R(\eta)^2 \rangle_c^{\text{tree}}}{\langle \delta_R(\eta)^2 \rangle_c^{\text{exact}}} \quad \varphi_R^{\text{tree}} \left(y \frac{\langle \delta_R(\eta)^2 \rangle_c^{\text{exact}}}{\langle \delta_R(\eta)^2 \rangle_c^{\text{tree}}}, \eta \right). \quad (20.2)$$

In Friedrich et al. (2018) we have shown that this approximation breaks down on scales $\lesssim 5$ Mpc. A possible way to improve this approximation is to also include the 1-loop contribution of equation 7.15 , i.e. to use

$$\phi_R^{\text{1-loop}}(y, \eta) = -S_y[\delta_L^*] - \frac{1}{2} \text{Tr} \ln(\delta_D + A_y). \quad (20.3)$$

While the first term on the right hand-side of this equation can be computed purely through the equations of spherical collapse (cf. chapter 7 which summarized the results of Valageas 2002a) the second term has most likely contribution from non-spherical configurations. It should be possible though to compute these contributions, e.g. , along the lines of Zibin (2008); Clarkson et al. (2009).

- To analyze the high density and low density tails of the matter density PDF in more detail (following the suggestion of Brouwer et al., 2018) we have to improve our modeling of these tails. Ansatzes of how to do that have, e.g. , been laid out by Valageas (2002a,b). It would especially be interesting to work out, how primordial non-Gaussianities (i.e. non-Gaussianities in the initial density field) would impact these high and low density parts of the PDF.

- We have to improve our understanding of how baryonic effects impact the dynamics of dark matter on small scales. It has, e.g. , been shown that these effects can be incorporated within an effective field theory of the large scale structure (Senatore, 2015; Lewandowski et al., 2015). Usually, simulated data is required to calibrate such an approach (Lewandowski et al., 2015). However, a combined analysis of density splits of the lensing and clustering power spectrum could be used to disentangle non-linear effects in the dark matter dynamics from those baryonic effects (cf. figure 20.1).

An extended analysis based on such an improved modeling would allow for an unprecedented test of the non-linear dynamics of matter and especially of dark matter.

Chapter 21

Cosmology and thermodynamics of spacetime

Early analyses of cosmological gravitational lensing such as Heymans et al. (2013); Hildebrandt et al. (2017); Köhlinger et al. (2017) seemed to indicate a mild tension between the parameters of the Λ CDM model determined from the cosmic microwave background (CMB Planck Collaboration et al., 2016) and the parameters obtained from the large scale structure of the low-redshift universe. Especially, the amplitude of density fluctuations found in Heymans et al. (2013); Hildebrandt et al. (2017); Köhlinger et al. (2017) was slightly lower than expected from the CMB observations, indicating that structures possibly formed slower than predicted by the Λ CDM model. Analyses of the Dark Energy Survey such as Abbott et al. (2016); Kacprzak et al. (2016); DES Collaboration et al. (2017) or the density split analysis of Friedrich et al. (2018); Gruen et al. (2018) could so far not confirm this disagreement. And studies such as Efstathiou & Lemos (2018); Sellentin et al. (2018); Troxel et al. (2018) indicate that at least part of the initial disagreement between high-redshift and low-redshift observations was due to systematic effects.

However, even if the Λ CDM model continues to describe cosmological observations, this cannot not be considered a successful confirmation of cosmological theory because the three major ingredients of this model are as of yet not understood. It is still unclear, what mechanism has driven the epoch of cosmic inflation (though there is increasing evidence for a purely slow roll, single field scenario, cf. Planck Collaboration et al., 2018). The nature of dark matter - i.e. of the CDM component - is not yet understood (Plehn, 2017). And finally, the origin of dark energy - i.e. of the Λ component - is still unknown (Frieman et al., 2008).

If the Λ CDM model persists, then it will be a major task for physicists to explain these phenomenological model ingredients. I think that recent theoretical ideas about *emergent gravity* (EG) could provide such explanations. Therefore I want to conclude this work with a brief review of these ideas.

Proposals of EG try to describe gravity as an emergent phenomenon that arises from the thermodynamic behavior of microscopic degrees of freedom of spacetime (Padmanabhan, 2010b; Verlinde, 2011, 2016). In the following I give a (necessarily incomplete) list of

theoretical developments that led to these proposals.

- **Black holes and thermodynamics:**

Observing similarities between black hole physics and thermodynamic laws, Bekenstein (1973) argued that an entropy could be assigned to the black hole event horizon. Hawking (1974) showed that black holes radiate with the spectrum of thermodynamic black bodies and Hawking (1975) showed that these radiating black holes fulfill a relation reminiscent of the second law of thermodynamics if one assigns an entropy to the black hole horizon that is of the same order of magnitude as the expression proposed before by Bekenstein (1973). Hayward (1998) derived a unified version of black hole and ordinary thermodynamics.

- **Horizons and thermodynamics:**

It was found that horizons other than black hole event horizons - such as the Unruh horizon of an accelerated observer, the horizon of de Sitter space or the apparent horizon in general Friedmann universes - also radiate with black body spectra and that entropy (per unit area) can be assigned to these horizons as well (Unruh, 1976; Gibbons & Hawking, 1977; Padmanabhan, 2002; Cai et al., 2009, see Hayward 1998 for the definition of an apparent horizon).

- **General relativity as a thermodynamic equation:**

Cai & Kim (2005) have shown that the Friedmann equations can be put into the form

$$-\frac{dE}{dt} = T \frac{dS}{dt}, \quad (21.1)$$

where T and S are the temperature and entropy of the apparent Horizon (which in a flat universe is the Hubble horizon) and dE/dt is the energy that crosses this horizon per unit time. In the form 21.1 the Friedmann equations then appear as the first law of thermodynamics. Requiring a similar equation to hold for the apparent horizons of all accelerated observers (i.e. for their local Unruh horizon) the complete Einstein equations themselves can be formulated as a first law of thermodynamics (Jacobson, 1995). This caused Padmanabhan (2010b) to speculate that the theory of general relativity describes a Thermodynamic equilibrium between the degrees of freedom of ordinary matter and microscopic degrees of freedom of spacetime.

As proposed by Verlinde (2016); Padmanabhan & Padmanabhan (2017); Capozziello & Luongo (2018), a thermodynamic origin of gravity could also have implications for the dark matter and dark energy components of the universe. They point out that in a dark energy dominated universe the Hubble horizon approaches a constant maximal size. Since the horizon entropy is proportional to the horizon area (Bekenstein, 1973; Gibbons & Hawking, 1977), this can be interpreted as a complete thermalization of the horizon degrees of freedom in the near future. From this point of view, the accelerated expansion of the universe just signifies that the horizon approaches its maximum entropy bound. Verlinde

(2016) even argues, that dark matter is only an apparent phenomenon associated with this thermalization process (see also Cai et al., 2017, for similar arguments).

Different proposals of EG can be divided into roughly two points of view:

- **Verlinde EG:**

The version of EG proposed by Verlinde (2011, 2016) considers gravity as an entropic force, i.e. as an apparent force that originates from the 2nd law of spacetime thermodynamics. This approach has been criticized by Dai & Stojkovic (2017), who noted that this might contradict the fact that in the Newtonian limit gravity is in fact a conservative force.

- **Padmanabhan EG:**

The version of EG that is, e.g. , promoted in Padmanabhan (2010b); Padmanabhan & Padmanabhan (2017) states that the curvature of spacetime results from the 1st law of a combined thermodynamics of spacetime and ordinary degrees of freedom. The motion of particles in such a curved spacetime is then still governed by the usual geodesic equations, which means that this version of EG is not affected by the criticism of Dai & Stojkovic (2017).

I find the EG proposal intriguing since it promises a physical explanation of the accelerated expansion of the universe (and possibly of dark matter). Furthermore, it may allow to interpret different gravity theories in terms of different equations of state of the spacetime degrees of freedom (Eling et al., 2006; Akbar & Cai, 2006; Padmanabhan, 2010b). Steps that could be pursued to confront these ideas with cosmological observations are the following:

1. One should construct a concrete model for the spacetime degrees of freedom - see, e.g. , Cao et al. (2017) for an attempt based on quantum tensor networks or Padmanabhan (2010b) for an attempt based on a light-like vector field.
2. From these models one should derive predictions for the expansion history of the universe and for the evolution of density perturbations. For the latter, even linear perturbation theory might be sufficient to test the EG proposal on large cosmological scales.

Brouwer et al. (2017) have already presented a pioneering attempt to test the proposal of Verlinde (2016) with the help of galaxy-galaxy lensing in the KiDS and GAMMA surveys. As explained by Brouwer et al. themselves, their analysis has a number of shortcomings. Most notably, they applied formulae from Verlinde (2016) that are only valid in an exact de Sitter space. I want to contribute to these efforts in the future by working out the expansion history and gravitational dynamics of concrete EG models and confronting those predictions with observations of the (dark) matter density field such as density split statistics.

Appendix A

The Einstein tensor in conformal Newtonian gauge

A.1 Without Perturbations

In a flat universe the line element is

$$ds^2 = a(\eta)^2 \{ d\eta^2 - d\vec{x}^2 \}. \quad (\text{A.1})$$

First lets compute the Christoffel symbols:

$$\begin{aligned}
\Gamma_{00}^0 &= \frac{g^{00}}{2} \partial_0 g_{00} \\
&= \frac{1}{2a^2} \cdot 2aa' \\
&= \mathcal{H} \\
\Gamma_{0i}^0 &= \frac{1}{2a^2} (\partial_0 g_{0i} + \partial_i g_{00} - \partial_0 g_{0i}) \\
&= 0 \\
\Gamma_{ij}^0 &= \frac{1}{2a^2} (\partial_j g_{0i} + \partial_i g_{j0} - \partial_0 g_{ji}) \\
&= \frac{-\delta_{ij}}{2a^2} \partial_0 g_{ii} \\
&= \delta_{ij} \mathcal{H} \\
\Gamma_{00}^i &= \frac{-1}{2a^2} (2\partial_0 g_{0i} - \partial_i g_{00}) \\
&= 0 \\
\Gamma_{0j}^i &= \frac{-1}{2a^2} (\partial_0 g_{ji} + \partial_j g_{0i} - \partial_i g_{0j}) \\
&= \delta_{ij} \mathcal{H} \\
\Gamma_{ij}^k &= \frac{-1}{2a^2} (\partial_i g_{jk} + \partial_j g_{ki} - \partial_k g_{ij}) \\
&= 0.
\end{aligned} \quad (\text{A.2})$$

The Ricci tensor then is

$$\begin{aligned}
R_{00} &= \partial_\alpha \Gamma_{00}^\alpha - \partial_0 \Gamma_{0\alpha}^\alpha + \Gamma_{\alpha\kappa}^\alpha \Gamma_{00}^\kappa - \Gamma_{0\kappa}^\alpha \Gamma_{\alpha 0}^\kappa \\
&= -3\partial_0 \mathcal{H} + \Gamma_{\alpha 0}^\alpha \Gamma_{00}^0 - \Gamma_{0\kappa}^\alpha \Gamma_{\alpha 0}^\kappa \\
&= -3\partial_0 \mathcal{H} \\
R_{0j} &= \partial_\alpha \Gamma_{0j}^\alpha - \partial_j \Gamma_{0\alpha}^\alpha + \Gamma_{\alpha\kappa}^\alpha \Gamma_{0j}^\kappa - \Gamma_{j\kappa}^\alpha \Gamma_{\alpha 0}^\kappa \\
&= \Gamma_{\alpha k}^\alpha \Gamma_{0j}^k - \Gamma_{j\kappa}^0 \Gamma_{00}^\kappa - \Gamma_{j\kappa}^l \Gamma_{l0}^\kappa \\
&= 0 \\
R_{ij} &= \partial_\alpha \Gamma_{ij}^\alpha - \partial_j \Gamma_{i\alpha}^\alpha + \Gamma_{\alpha\kappa}^\alpha \Gamma_{ij}^\kappa - \Gamma_{j\kappa}^\alpha \Gamma_{\alpha i}^\kappa \\
&= \delta_{ij} \partial_0 \mathcal{H} + \Gamma_{\alpha 0}^\alpha \Gamma_{ij}^0 - \Gamma_{j\kappa}^\alpha \Gamma_{\alpha i}^\kappa \\
&= \delta_{ij} \partial_0 \mathcal{H} + 4\delta_{ij} \mathcal{H}^2 - \Gamma_{j0}^l \Gamma_{li}^0 - \Gamma_{jk}^0 \Gamma_{0i}^k \\
&= \delta_{ij} \partial_0 \mathcal{H} + 4\delta_{ij} \mathcal{H}^2 - \Gamma_{j0}^l \Gamma_{li}^0 - \Gamma_{jk}^0 \Gamma_{0i}^k \\
&= \delta_{ij} \partial_0 \mathcal{H} + 2\delta_{ij} \mathcal{H}^2.
\end{aligned} \tag{A.3}$$

The Ricci scalar is

$$R = \frac{-2}{a^2} (3\partial_0 \mathcal{H} + 3\mathcal{H}^2). \tag{A.4}$$

This can also be expressed in terms of t and H as

$$R = -6 \left(\frac{\partial H}{\partial t} + 2H^2 \right). \tag{A.5}$$

For the Einstein tensor $G_{\alpha\beta} = R_{\alpha\beta} - \frac{g_{\alpha\beta}}{2}R$ we get

$$\begin{aligned}
G_{00} &= -3\partial_0 \mathcal{H} + 3\partial_0 \mathcal{H} + 3\mathcal{H}^2 \\
&= 3\mathcal{H}^2 \\
G_{0i} &= 0 \\
G_{ij} &= \delta_{ij} \cdot (\partial_0 \mathcal{H} + 2\mathcal{H}^2 - 3\partial_0 \mathcal{H} - 3\mathcal{H}^2) \\
&= -\delta_{ij} \cdot (2\mathcal{H}' + \mathcal{H}^2).
\end{aligned} \tag{A.6}$$

A.2 With scalar Perturbations

With scalar perturbations the line element is

$$ds^2 = a(\eta)^2 \{ e^{2\phi} d\eta^2 - e^{-2\psi} d\vec{x}^2 \}. \tag{A.7}$$

Note that we slightly deviate here from the notation of equation 5.35 since this makes the following calculations simpler. But the two notations are identical at linear order in the perturbations ψ and ϕ .

The Christoffel symbols are:

$$\begin{aligned}
\Gamma_{00}^0 &= \frac{g^{00}}{2} \partial_0 g_{00} = \frac{e^{-2\phi}}{2a^2} \cdot (2aa' + 2a^2\phi') e^{2\phi} \\
&= \mathcal{H} + \phi' \\
\Gamma_{0i}^0 &= \frac{e^{-2\phi}}{2a^2} (\partial_0 g_{0i} + \partial_i g_{00} - \partial_0 g_{0i}) = \frac{e^{-2\phi}}{2a^2} \cdot 2a^2 \partial_i \phi e^{2\phi} \\
&= \partial_i \phi \\
\Gamma_{ij}^0 &= \frac{e^{-2\phi}}{2a^2} (\partial_j g_{0i} + \partial_i g_{j0} - \partial_0 g_{ji}) = -\delta_{ij} \frac{e^{-2\phi}}{2a^2} (-2aa' + 2a^2\psi') e^{-2\psi} \\
&= \delta_{ij} (\mathcal{H} - \psi') e^{-2(\phi+\psi)} \\
\Gamma_{00}^i &= \frac{-e^{2\psi}}{2a^2} (2\partial_0 g_{0i} - \partial_i g_{00}) \\
&= e^{2(\phi+\psi)} \partial_i \phi \\
\Gamma_{0j}^i &= \frac{-e^{2\psi}}{2a^2} (\partial_0 g_{ji} + \partial_j g_{0i} - \partial_i g_{0j}) = \delta_{ij} \frac{-e^{2\psi}}{2a^2} (-2aa' + 2a^2\psi') e^{-2\psi} \\
&= \delta_{ij} (\mathcal{H} - \psi') \\
\Gamma_{ij}^k &= \frac{-e^{2\psi}}{2a^2} (\partial_i g_{jk} + \partial_j g_{ki} - \partial_k g_{ij}) \\
&= \delta_{ij} \partial_k \psi - \delta_{ki} \partial_j \psi - \delta_{jk} \partial_i \psi. \tag{A.8}
\end{aligned}$$

Note especially the following equalities:

$$\begin{aligned}
\Gamma_{0\alpha}^\alpha &= 4\mathcal{H} + \phi' - 3\psi' \\
\Gamma_{ki}^k &= \delta_{ik} \partial_k \psi - \delta_{ki} \partial_k \psi - \delta_{kk} \partial_i \psi \\
&= -3\partial_i \psi \\
\Gamma_{\alpha i}^\alpha &= \partial_i \phi + \Gamma_{ki}^k \\
&= \partial_i \phi - 3\partial_i \psi \\
\partial_k \Gamma_{ij}^k &= \partial_k (\delta_{ij} \partial_k \phi - \delta_{ki} \partial_j \phi - \delta_{jk} \partial_i \phi) \\
&= \delta_{ij} \Delta \psi - 2\partial_i \partial_j \psi \\
\partial_k \psi \Gamma_{ij}^k &= \partial_k \psi (\delta_{ij} \partial_k \psi - \delta_{ki} \partial_j \psi - \delta_{jk} \partial_i \psi) \\
&= \delta_{ij} (\vec{\nabla} \psi)^2 - 2\partial_i \psi \partial_j \psi \\
\partial_k \phi \Gamma_{ij}^k &= \partial_k \phi (\delta_{ij} \partial_k \psi - \delta_{ki} \partial_j \psi - \delta_{jk} \partial_i \psi) \\
&= \delta_{ij} (\vec{\nabla} \phi \vec{\nabla} \psi) - 2\partial_i \phi \partial_j \psi. \tag{A.9}
\end{aligned}$$

The Ricci tensor then is

$$\begin{aligned}
R_{00} &= \partial_\alpha \Gamma_{00}^\alpha - \partial_0 \Gamma_{0\alpha}^\alpha + \Gamma_{\alpha\kappa}^\alpha \Gamma_{00}^\kappa - \Gamma_{0\kappa}^\alpha \Gamma_{\alpha 0}^\kappa \\
&= \frac{1}{2} [e^{2\psi} \Delta e^{2\phi} + (\nabla e^{2\psi}) \cdot (\nabla e^{2\phi})] - 3(\mathcal{H}' - \psi'') + \Gamma_{\alpha 0}^\alpha \Gamma_{00}^0 + \Gamma_{\alpha k}^\alpha \Gamma_{00}^k - \Gamma_{0\kappa}^\alpha \Gamma_{\alpha 0}^\kappa \\
&= \frac{1}{2} [e^{2\psi} \Delta e^{2\phi} + (\nabla e^{2\psi}) \cdot (\nabla e^{2\phi})] - 3(\mathcal{H}' - \psi'') + (4\mathcal{H} + \phi' - 3\psi')(\mathcal{H} + \phi') \\
&\quad + e^{2(\phi+\psi)} (\partial_k \phi - 3\partial_k \psi) \partial_k \phi - \Gamma_{0\kappa}^\alpha \Gamma_{\alpha 0}^\kappa \\
&= \frac{1}{2} [e^{2\psi} \Delta e^{2\phi} + (\nabla e^{2\psi}) \cdot (\nabla e^{2\phi})] - 3(\mathcal{H}' - \psi'') + (4\mathcal{H} + \phi' - 3\psi')(\mathcal{H} + \phi') \\
&\quad + e^{2(\phi+\psi)} (\nabla \phi - 3\nabla \psi) \cdot \nabla \phi - \Gamma_{00}^0 \Gamma_{00}^0 - 2\Gamma_{0k}^0 \Gamma_{00}^k - \Gamma_{0k}^l \Gamma_{l0}^k \\
&= \frac{1}{2} [e^{2\psi} \Delta e^{2\phi} + (\nabla e^{2\psi}) \cdot (\nabla e^{2\phi})] - 3(\mathcal{H}' - \psi'') + (4\mathcal{H} + \phi' - 3\psi')(\mathcal{H} + \phi') \\
&\quad - \frac{3}{4} (\nabla e^{2\psi}) \cdot (\nabla e^{2\phi}) + \frac{e^{2(\psi-\phi)}}{4} (\nabla e^{2\phi})^2 \\
&\quad - \Gamma_{00}^0 \Gamma_{00}^0 - 2\Gamma_{0k}^0 \Gamma_{00}^k - \Gamma_{0k}^l \Gamma_{l0}^k \\
&= \frac{1}{2} [e^{2\psi} \Delta e^{2\phi} + (\nabla e^{2\psi}) \cdot (\nabla e^{2\phi})] - 3(\mathcal{H}' - \psi'') + (4\mathcal{H} + \phi' - 3\psi')(\mathcal{H} + \phi') \\
&\quad - \frac{3}{4} (\nabla e^{2\psi}) \cdot (\nabla e^{2\phi}) + \frac{e^{2(\psi-\phi)}}{4} (\nabla e^{2\phi})^2 \\
&\quad - (\mathcal{H} + \phi')^2 - \frac{e^{2(\psi-\phi)}}{2} (\nabla e^{2\phi})^2 - 3(\mathcal{H} - \psi')^2 \\
&= \frac{1}{2} [e^{2\psi} \Delta e^{2\phi} + (\nabla e^{2\psi}) \cdot (\nabla e^{2\phi})] - 3(\mathcal{H}' - \psi'') + (4\mathcal{H} + \phi' - 3\psi')(\mathcal{H} + \phi') \\
&\quad - \frac{3}{4} (\nabla e^{2\psi}) \cdot (\nabla e^{2\phi}) - \frac{e^{2(\psi-\phi)}}{4} (\nabla e^{2\phi})^2 - (\mathcal{H} + \phi')^2 - 3(\mathcal{H} - \psi')^2 \\
&= \frac{1}{2} [e^{2\psi} \Delta e^{2\phi} + (\nabla e^{2\psi}) \cdot (\nabla e^{2\phi})] - 3(\mathcal{H}' - \psi'') + 3(\mathcal{H}(\phi' + \psi') - \psi'(\phi' + \psi')) \\
&\quad - \frac{3}{4} (\nabla e^{2\psi}) \cdot (\nabla e^{2\phi}) - \frac{e^{2(\psi-\phi)}}{4} (\nabla e^{2\phi})^2 \\
&= \frac{e^{2\psi}}{2} \Delta e^{2\phi} - 3(\mathcal{H}' - \psi'') + 3(\mathcal{H}(\phi' + \psi') - \psi'(\phi' + \psi')) \\
&\quad - \frac{1}{4} (\nabla e^{2\psi}) \cdot (\nabla e^{2\phi}) - \frac{e^{2(\psi-\phi)}}{4} (\nabla e^{2\phi})^2 \\
&= \Delta\phi - 3(\mathcal{H}' - \psi'') + 3\mathcal{H}(\phi' + \psi') + \mathcal{O}(\phi^2, \psi^2, \phi\psi). \tag{A.10}
\end{aligned}$$

$$\begin{aligned}
R_{0j} &= \partial_\alpha \Gamma_{0j}^\alpha - \partial_j \Gamma_{0\alpha}^\alpha + \Gamma_{\alpha\kappa}^\alpha \Gamma_{0j}^\kappa - \Gamma_{j\kappa}^\alpha \Gamma_{\alpha 0}^\kappa \\
&= \partial_j \phi' + \partial_j (\mathcal{H} - \psi') - \partial_j (4\mathcal{H} + \phi' - 3\psi') + \Gamma_{\alpha\kappa}^\alpha \Gamma_{0j}^\kappa - \Gamma_{j\kappa}^\alpha \Gamma_{\alpha 0}^\kappa \\
&= 2\partial_j \psi' + (4\mathcal{H} + \phi' - 3\psi') \partial_j \phi + \Gamma_{\alpha k}^\alpha \Gamma_{0j}^k - \Gamma_{j\kappa}^\alpha \Gamma_{\alpha 0}^\kappa \\
&= 2\partial_j \psi' + (4\mathcal{H} + \phi' - 3\psi') \partial_j \phi + (\mathcal{H} - \psi') (\partial_j \phi - 3\partial_j \psi) \\
&\quad - \Gamma_{j0}^0 \Gamma_{00}^0 - \Gamma_{j0}^l \Gamma_{l0}^0 - \Gamma_{jk}^0 \Gamma_{00}^k - \Gamma_{jk}^l \Gamma_{l0}^k \\
&= 2\partial_j \psi' + (4\mathcal{H} + \phi' - 3\psi') \partial_j \phi + (\mathcal{H} - \psi') (\partial_j \phi - 3\partial_j \psi) \\
&\quad - (\mathcal{H} + \phi') \partial_j \phi - (\mathcal{H} - \psi') \partial_j \phi - \Gamma_{jk}^0 \Gamma_{00}^k - \Gamma_{jk}^l \Gamma_{l0}^k \\
&= 2\partial_j \psi' + (3\mathcal{H} - 3\psi') \partial_j \phi - 3\partial_j \psi (\mathcal{H} - \psi') \\
&\quad - \Gamma_{jk}^0 \Gamma_{00}^k - \Gamma_{jk}^l \Gamma_{l0}^k \\
&= 2\partial_j \psi' + 2(\mathcal{H} - \psi') \partial_j \phi - 3\partial_j \psi (\mathcal{H} - \psi') - \Gamma_{jk}^l \Gamma_{l0}^k \\
&= 2\partial_j \psi' + (\mathcal{H} - \psi') (2\partial_j \phi - 3\partial_j \psi) + 3\partial_j \psi (\mathcal{H} - \psi') \\
&= 2\partial_j \psi' + 2(\mathcal{H} - \psi') \partial_j \phi \\
&= 2\partial_j \psi' + 2\mathcal{H} \partial_j \phi + \mathcal{O}(\phi\psi) . \tag{A.11}
\end{aligned}$$

$$\begin{aligned}
R_{ij} &= \partial_\alpha \Gamma_{ij}^\alpha - \partial_i \Gamma_{j\alpha}^\alpha + \Gamma_{\alpha\kappa}^\alpha \Gamma_{ij}^\kappa - \Gamma_{i\kappa}^\alpha \Gamma_{\alpha j}^\kappa \\
&= \partial_0 \Gamma_{ij}^0 + \delta_{ij} \Delta \psi - 2\partial_i \partial_j \psi - \partial_i \partial_j (\phi - 3\psi) + \Gamma_{\alpha\kappa}^\alpha \Gamma_{ij}^\kappa - \Gamma_{i\kappa}^\alpha \Gamma_{\alpha j}^\kappa \\
&= \delta_{ij} [\mathcal{H}' - \psi'' - 2(\phi' + \psi')(\mathcal{H} - \psi')] e^{-2(\phi+\psi)} + \delta_{ij} \Delta \psi - \partial_i \partial_j (\phi - \psi) + \Gamma_{\alpha\kappa}^\alpha \Gamma_{ij}^\kappa - \Gamma_{i\kappa}^\alpha \Gamma_{\alpha j}^\kappa \\
&= \delta_{ij} [\mathcal{H}' - \psi'' - 2(\phi' + \psi')(\mathcal{H} - \psi')] e^{-2(\phi+\psi)} + \delta_{ij} \Delta \psi - \partial_i \partial_j (\phi - \psi) \\
&\quad + \Gamma_{\alpha 0}^\alpha \Gamma_{ij}^0 + \Gamma_{\alpha k}^\alpha \Gamma_{ij}^k - \Gamma_{i\kappa}^\alpha \Gamma_{\alpha j}^\kappa \\
&= \delta_{ij} [\mathcal{H}' - \psi'' - 2(\phi' + \psi')(\mathcal{H} - \psi')] e^{-2(\phi+\psi)} + \delta_{ij} \Delta \psi - \partial_i \partial_j (\phi - \psi) \\
&\quad + \delta_{ij} (4\mathcal{H} + \phi' - 3\psi')(\mathcal{H} - \psi') e^{-2(\phi+\psi)} + \Gamma_{\alpha k}^\alpha \Gamma_{ij}^k - \Gamma_{i\kappa}^\alpha \Gamma_{\alpha j}^\kappa \\
&= \delta_{ij} [\mathcal{H}' - \psi'' + (4\mathcal{H} - \phi' - 5\psi')(\mathcal{H} - \psi')] e^{-2(\phi+\psi)} + \delta_{ij} \Delta \psi - \partial_i \partial_j (\phi - \psi) \\
&\quad + \delta_{ij} [(\nabla \phi \nabla \psi) - 3(\nabla \psi)^2] + 6\partial_i \psi \partial_j \psi - \partial_i \phi \partial_j \psi - \partial_i \psi \partial_j \phi \\
&\quad - \Gamma_{i0}^0 \Gamma_{0j}^0 - \Gamma_{i0}^k \Gamma_{kj}^0 - \Gamma_{ik}^0 \Gamma_{0j}^k - \Gamma_{ik}^l \Gamma_{lj}^k \\
&= \delta_{ij} [\mathcal{H}' - \psi'' + (4\mathcal{H} - \phi' - 5\psi')(\mathcal{H} - \psi')] e^{-2(\phi+\psi)} + \delta_{ij} \Delta \psi - \partial_i \partial_j (\phi - \psi) \\
&\quad + \delta_{ij} [(\nabla \phi \nabla \psi) - 3(\nabla \psi)^2] + 6\partial_i \psi \partial_j \psi - \partial_i \phi \partial_j \psi - \partial_i \psi \partial_j \phi \\
&\quad - \partial_i \phi \partial_j \phi - 2\delta_{ij} (\mathcal{H} - \psi')^2 e^{-2(\phi+\psi)} - \Gamma_{ik}^l \Gamma_{lj}^k \\
&= \delta_{ij} [\mathcal{H}' - \psi'' + (2\mathcal{H} - \phi' - 3\psi')(\mathcal{H} - \psi')] e^{-2(\phi+\psi)} + \delta_{ij} \Delta \psi - \partial_i \partial_j (\phi - \psi) \\
&\quad + \delta_{ij} [(\nabla \phi \nabla \psi) - 3(\nabla \psi)^2] + 6\partial_i \psi \partial_j \psi - \partial_i \phi \partial_j \psi - \partial_i \psi \partial_j \phi - \partial_i \phi \partial_j \phi \\
&\quad - (\delta_{ik} \partial_l \psi - \delta_{li} \partial_k \psi - \delta_{kl} \partial_i \psi) (\delta_{jl} \partial_k \psi - \delta_{kj} \partial_l \psi - \delta_{lk} \partial_j \psi) \\
&= \delta_{ij} [\mathcal{H}' - \psi'' + (2\mathcal{H} - \phi' - 3\psi')(\mathcal{H} - \psi')] e^{-2(\phi+\psi)} + \delta_{ij} \Delta \psi - \partial_i \partial_j (\phi - \psi) \\
&\quad + \delta_{ij} [(\nabla \phi \nabla \psi) - 3(\nabla \psi)^2] + 6\partial_i \psi \partial_j \psi - \partial_i \phi \partial_j \psi - \partial_i \psi \partial_j \phi - \partial_i \phi \partial_j \phi \\
&\quad - \partial_l \psi (\delta_{jl} \partial_i \psi - \delta_{ij} \partial_l \psi - \delta_{li} \partial_j \psi) \\
&\quad + \partial_k \psi (\delta_{ji} \partial_k \psi - \delta_{kj} \partial_i \psi - \delta_{ik} \partial_j \psi) \\
&\quad + \partial_i \psi (\delta_{jl} \partial_l \psi - \delta_{lj} \partial_l \psi - \delta_{ll} \partial_j \psi) \\
&= \delta_{ij} [\mathcal{H}' - \psi'' + (2\mathcal{H} - \phi' - 3\psi')(\mathcal{H} - \psi')] e^{-2(\phi+\psi)} \\
&\quad + \delta_{ij} [\Delta \psi - (\nabla \psi)^2 + (\nabla \phi \nabla \psi)] \\
&\quad + \partial_i \psi \partial_j \psi - \partial_i \phi \partial_j \psi - \partial_i \psi \partial_j \phi - \partial_i \phi \partial_j \phi - \partial_i \partial_j (\phi - \psi) \\
&= \delta_{ij} [2\mathcal{H}^2 + \mathcal{H}' - \psi'' + \Delta \psi - \mathcal{H}(\phi' + 5\psi') - 2(2\mathcal{H}^2 + \mathcal{H}')(\phi + \psi)] \\
&\quad - \partial_i \partial_j (\phi - \psi) + \mathcal{O}(\phi^2, \psi^2, \phi\psi) . \tag{A.12}
\end{aligned}$$

The Ricci scalar is given by

$$\begin{aligned}
a^2 R &= a^2 (g^{00} R_{00} + g^{ij} R_{ij}) \\
&= e^{-2\phi} \{ e^{2(\phi+\psi)} [\Delta\phi + (\nabla\phi)^2 - (\nabla\psi) \cdot (\nabla\phi)] \\
&\quad - 3(\mathcal{H}' - \psi'') + 3(\mathcal{H}(\phi' + \psi') - \psi'(\phi' + \psi')) \} \\
&\quad - e^{2\psi} \{ 3e^{-2(\phi+\psi)} [\mathcal{H}' - \psi'' + (2\mathcal{H} - \phi' - 3\psi')(\mathcal{H} - \psi')] \\
&\quad + 3\Delta\psi - 2(\nabla\psi)^2 - (\nabla\phi)^2 + (\nabla\phi\nabla\psi) - \Delta(\phi - \psi) \} \\
&= e^{2\psi} \{ \Delta\phi + (\nabla\phi)^2 - (\nabla\psi\nabla\phi) - 3\Delta\psi + 2(\nabla\psi)^2 + (\nabla\phi)^2 - (\nabla\phi\nabla\psi) + \Delta(\phi - \psi) \} \\
&\quad - 3e^{-2\phi} \{ \mathcal{H}' - \psi'' - (\mathcal{H}(\phi' + \psi') - \psi'(\phi' + \psi')) + [\mathcal{H}' - \psi'' + (2\mathcal{H} - \phi' - 3\psi')(\mathcal{H} - \psi')] \} \\
&= 2e^{2\psi} \{ \Delta\phi - 2\Delta\psi + (\nabla\phi)^2 - (\nabla\psi\nabla\phi) + (\nabla\psi)^2 \} \\
&\quad - 3e^{-2\phi} \{ 2\mathcal{H}^2 + 2\mathcal{H}' - 2\psi'' - 2\mathcal{H}(\phi' + 3\psi') + 2\phi'\psi' + 4\psi'^2 \} \\
&= -6\mathcal{H}^2 - 6\mathcal{H}' + 2\Delta\phi - 4\Delta\psi + 3\phi'' + 3\psi'' + 6\mathcal{H}(\phi' + 3\psi') + 12(\mathcal{H}^2 + \mathcal{H}')\phi + \mathcal{O}(\phi^2, \psi^2, \phi\psi)
\end{aligned} \tag{A.13}$$

For the Einstein tensor $G_{\alpha\beta} = R_{\alpha\beta} - \frac{g_{\alpha\beta}}{2}R$ this means

$$\begin{aligned}
G_{00} &= -3(\mathcal{H}' - \psi'') + 3(\mathcal{H}(\phi' + \psi') - \psi'(\phi' + \psi')) \\
&\quad + e^{2(\phi+\psi)} [\Delta\phi + (\nabla\phi)^2 - (\nabla\psi\nabla\phi)] \\
&\quad - e^{2(\phi+\psi)} [\Delta\phi - 2\Delta\psi + (\nabla\phi)^2 - (\nabla\psi\nabla\phi) + (\nabla\psi)^2] \\
&\quad + \frac{3}{2} (2\mathcal{H}^2 + 2\mathcal{H}' - 2\psi'' - 2\mathcal{H}(\phi' + 3\psi') + 2\phi'\psi' + 4\psi'^2) \\
&= 3\mathcal{H}^2 - 6\mathcal{H}\psi' + 3\psi'^2 + e^{2(\phi+\psi)} [2\Delta\psi - (\nabla\psi)^2] \\
G_{0i} &= 2\partial_j\psi' + 2(\mathcal{H} - \psi')\partial_j\phi \\
G_{ij} &= \delta_{ij} [\mathcal{H}' - \psi'' + (2\mathcal{H} - \phi' - 3\psi')(\mathcal{H} - \psi')] e^{-2(\phi+\psi)} \\
&\quad + \delta_{ij} [\Delta\phi - (\nabla\psi)^2 + (\nabla\phi\nabla\psi)] \\
&\quad + \partial_i\psi\partial_j\psi - \partial_i\phi\partial_j\psi - \partial_i\psi\partial_j\phi - \partial_i\phi\partial_j\phi - \partial_i\partial_j(\phi - \psi) \\
&\quad + \delta_{ij} [\Delta\phi - 2\Delta\psi + (\nabla\phi)^2 + (\nabla\psi)^2 - (\nabla\psi\nabla\phi)] \\
&\quad - 3\delta_{ij} [\mathcal{H}^2 + \mathcal{H}' - \psi'' - \mathcal{H}(\phi' + 3\psi') + \phi'\psi' + 2\psi'^2] e^{-2(\phi+\psi)} \\
&= \delta_{ij} [-\mathcal{H}^2 - 2\mathcal{H}' + 2\psi'' + 2\mathcal{H}(\phi' + 2\psi') - 2\phi'\psi' - 3\psi'^2] e^{-2(\phi+\psi)} \\
&\quad + \delta_{ij} [\Delta(\phi - \psi) + (\nabla\phi)^2] \\
&\quad + \partial_i\psi\partial_j\psi - \partial_i\phi\partial_j\psi - \partial_i\psi\partial_j\phi - \partial_i\phi\partial_j\phi - \partial_i\partial_j(\phi - \psi) .
\end{aligned} \tag{A.14}$$

To first order in the perturbation this gives

$$\begin{aligned}
G_{00} &= 3\mathcal{H}^2 - 6\mathcal{H}\psi' + 2\Delta\psi \\
G_{0i} &= 2\partial_j(\psi' + \mathcal{H}\phi) \\
G_{ij} &= \delta_{ij} [-\mathcal{H}^2 - 2\mathcal{H}' + 2(\mathcal{H}^2 + 2\mathcal{H}')(\phi + \psi) + 2\psi'' + 2\mathcal{H}(\phi' + 2\psi') + \Delta(\phi - \psi)] \\
&\quad + \partial_i\partial_j(\psi - \phi)
\end{aligned} \tag{A.15}$$

or

$$\begin{aligned}
 a^2 G_0^0 &= 3\mathcal{H}^2 - 6\mathcal{H}\psi' - 6\mathcal{H}^2\phi + 2\Delta\psi \\
 a^2 G_0^i &= -2\partial_j(\psi' + \mathcal{H}\phi) \\
 a^2 G_i^j &= -\delta_{ij} \left[-\mathcal{H}^2 - 2\mathcal{H}' + 2(\mathcal{H}^2 + 2\mathcal{H}')\phi + 2\psi'' + 2\mathcal{H}(\phi' + 2\psi') + \Delta(\phi - \psi) \right] \\
 &\quad - \partial_i\partial_j(\psi - \phi)
 \end{aligned} \tag{A.16}$$

or

$$\begin{aligned}
 a^4 G^{00} &= 3\mathcal{H}^2 - 6\mathcal{H}\psi' - 12\mathcal{H}^2\phi + 2\Delta\psi \\
 a^4 G^{0i} &= -2\partial_j(\psi' + \mathcal{H}\phi) \\
 a^4 G^{ij} &= \delta_{ij} \left[-\mathcal{H}^2 - 2\mathcal{H}' + 2(\mathcal{H}^2 + 2\mathcal{H}')(\phi - \psi) + 2\psi'' + 2\mathcal{H}(\phi' + 2\psi') + \Delta(\phi - \psi) \right] \\
 &\quad + \partial_i\partial_j(\psi - \phi).
 \end{aligned} \tag{A.17}$$

A.3 Perturbations of $T_{\mu\nu}$

The perturbed energy momentum tensor is

$$\begin{aligned}
 T_{00} &= (\rho^0 + \delta\rho + p^0 + \delta p)a^2(1 + 2\phi) - (1 + 2\phi)a^2(p^0 + \delta p) \\
 &= (\rho^0 + \delta\rho)a^2(1 + 2\phi) \\
 &= {}^0T_{00} + a^2(\delta\rho + 2\phi\rho^0) \\
 T_{ii} &= (1 - 2\phi)a^2(p^0 + \delta p) \\
 &= {}^0T_{ii} + a^2(\delta p - 2\phi p^0) \\
 T_{0i} &= (\rho^0 + \delta\rho + p^0 + \delta p)a\sqrt{1 + 2\phi}\delta u_i \\
 &= (\rho^0 + p^0)a\delta u_i.
 \end{aligned} \tag{A.18}$$

Note that in order to have pure scalar perturbations the velocity field δu_i needs to be the divergence of some scalar field.

Appendix B

2-point correlation function on curved sky

The angular clustering correlation function of galaxies $w(\theta)$ is given in terms of the galaxy clustering power spectrum C_ℓ^{gg} as

$$w(\theta) = \sum_\ell \frac{2\ell+1}{4\pi} P_\ell(\cos \theta) \left(C_\ell^{gg} + \frac{1}{n} \right) , \quad (\text{B.1})$$

where n is the galaxy density per steradian. The term proportional to $\frac{1}{n}$ is usually omitted (cf. Ross et al. 2011) since it sums up to¹

$$\begin{aligned} \frac{1}{2\pi n} \sum_\ell \frac{2\ell+1}{2} P_\ell(\cos \theta) &= \frac{1}{2\pi n} \delta_D(\cos \theta - 1) \\ &= \frac{\delta_D(\theta)}{2\pi n \sin \theta} , \end{aligned} \quad (\text{B.2})$$

which has to be interpreted as a 2-dimensional Dirac delta function on the sphere.

According to de Putter & Takada (2010) (see also Stebbins (1996)) the galaxy-galaxy lensing correlation function $\gamma_t(\theta)$ is given in terms of the galaxy-convergence cross-power spectrum $C_\ell^{g\kappa}$ as

$$\gamma_t(\theta) = \sum_\ell \frac{2\ell+1}{4\pi\ell(\ell+1)} P_\ell^2(\cos \theta) C_\ell^{g\kappa} , \quad (\text{B.3})$$

where P_ℓ^m are the associated Legendre Polynomials.

Finally, the cosmic shear correlation functions $\xi_\pm(\theta)$ are given by

$$\begin{aligned} \xi_\pm(\theta) &= \frac{1}{2\pi} \sum_{\ell \geq 2} \frac{2\ell+1}{\ell^2(\ell+1)^2} * \\ &* \left[C_\ell^E (G_{\ell,2}^+(x) \pm G_{\ell,2}^-(x)) + C_\ell^B (G_{\ell,2}^-(x) \pm G_{\ell,2}^+(x)) \right] , \end{aligned} \quad (\text{B.4})$$

¹from $\sum_\ell \frac{2\ell+1}{2} P_\ell(x) P_\ell(y) = \delta_D(x - y)$ - see N. Bronstein & A. Semendjajew (1979) for this and other properties of Legendre polynomials.

where C_ℓ^E and C_ℓ^B are the E-mode and B-mode power spectra of shear, $x = \cos \theta$ and the functions $G_{\ell,2}^\pm(x)$ are defined in eq. 4.18² of Stebbins (1996). Eq. B.4 can be expressed in terms of associated Legendre polynomials by using eq. 4.19 of Stebbins (1996), which gives

$$\begin{aligned} G_{\ell,2}^+(x) \pm G_{\ell,2}^-(x) &= P_\ell^2(x) \left\{ \frac{4 - \ell \pm 2x(\ell - 1)}{1 - x^2} - \frac{\ell(\ell - 1)}{2} \right\} \\ &\quad + P_{\ell-1}^2(x) \frac{(\ell + 2)(x \mp 2)}{1 - x^2}. \end{aligned} \quad (\text{B.5})$$

²Note that a factor of $1/i \sin(\theta)$ is missing in the second line of this equation.

Appendix C

Use case for combining internal covariance estimates and PME

At first glance it seems like the 4-point function of the density field that appears in the cosmic variance part of any 2-point function covariance is the most difficult part when modelling the covariance matrix of 2-point function measurements (Schneider, 2005; Joachimi et al., 2008; Sato et al., 2009; Hilbert et al., 2011; Eifler et al., 2014; Krause & Eifler, 2016). However, a recent study by Troxel et al. (2018) found that the covariance model used by Hildebrandt et al. (2017); van Uitert et al. (2018) (and to a smaller degree also the model used by DES Collaboration et al. (2017)) suffered most from inaccuracies in how they accounted for masking effects and survey geometry in the noise parts of their covariance (i.e. in the shape-noise and shot-noise parts).

Apart from masking and survey geometry, there are other effects that would mainly impact the noise-terms of the covariance. There is e.g. doubt on whether the shot-noise of galaxies is indeed Poissonian (Hamaus et al., 2010; Cacciato et al., 2012; Baldauf et al., 2013; Dvornik et al., 2018; Friedrich et al., 2018; Gruen et al., 2018). Also, the quality of galaxy shape measurements could be a function of the intrinsic shapes of the individual galaxies (Miller et al., 2013; Huff & Mandelbaum, 2017; Sheldon & Huff, 2017) and it has not been investigated yet, whether this affects the shape-noise part in the covariance matrix. As we have shown in Friedrich et al. (2016), internal covariance estimators can in principle give unbiased estimates of the noise parts of the covariance matrix. In combination with the PME hybrid estimator developed in Friedrich & Eifler (2018) this could be used to derive corrections to modelled noise terms in the covariance from the data itself.

To demonstrate the feasibility of this approach, I measured the 2-point correlation functions $\xi_+(\theta)$, $\xi_-(\theta)$, $\gamma_t(\theta)$ and $w(\theta)$ in one source galaxy bin and one lens bin of the log-normal simulations that we also used in Gruen, Friedrich et al., 2018 (cf. chapter 17). I measured those correlation functions in $N_s \approx 880$ of these simulations using angular scales comparable to the scales used in the 2-point function analysis of DES Collaboration et al. (2017). By construction, the shot-noise in these log-normal simulations should be Poissonian and the intrinsic ellipticities of the source galaxies were drawn randomly (i.e. no actual shape measurement took place). Hence, it is possible to construct a model for

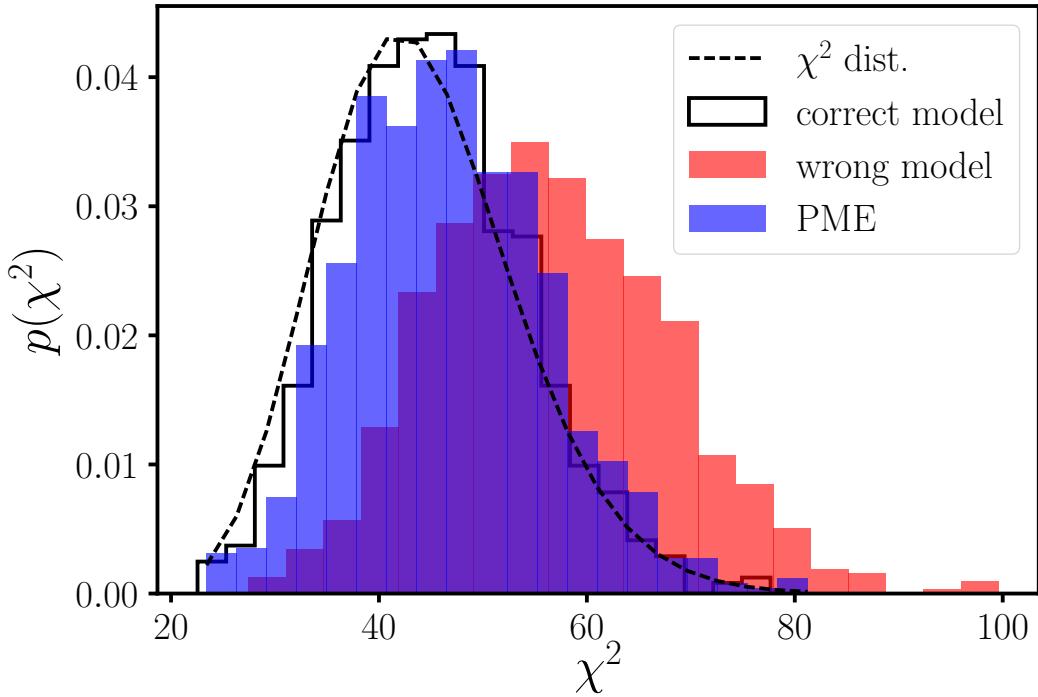


Figure C.1: Histogram of χ^2 values obtained for the residual of simulated data defined in equation C.2 using different approximations to the precision matrix: a model covariance that correctly describes the shape-noise and shot-noise in the simulated data (black solid line), a model covariance that underestimates this noise by 30% (red bars) and correction of this wrong covariance using data-internal estimates of the noise terms in the PME technique (blue bars). The black dashed line shows the distribution of χ^2 that would be expected if the model with the correct noise term would be the exact covariance matrix of the simulated data (and if that data had a perfectly Gaussian distribution). However, our model covariance is based on the Gaussian covariance model (Schneider et al., 2002; Joachimi et al., 2008) while the simulations are based on log-normal realisations of the density field (Hilbert et al., 2011; Xavier et al., 2016).

the shot-noise and shape-noise components of the covariance that exactly describes the simulations. In addition to this model I also constructed an estimate of these noise terms from each simulation using an internal resampling technique similar to the ones described in Friedrich et al. (2016). However, to make sure that this estimate only measures the noise parts of the covariance, I did the resampling without any regard for the location of galaxies, i.e. without cutting the simulations into spatial subregions. I then consider the following situation that mimics the use case of the *precision matrix expansion* we introduced in Friedrich & Eifler (2018):

- A covariance model that includes the correct noise terms acts as the (unknown) true covariance matrix. This would be the matrix \mathbf{C} in equation 14.7.
- A covariance model in which the noise terms are multiplied by a factor of 0.7 acts as a (wrong) model covariance matrix. This would be the matrix \mathbf{M} in equation 14.7, while the wrong noise part itself would correspond to the matrix \mathbf{B}_m in that equation.
- The internal estimate of the noise terms in the covariance then serve as a estimate $\hat{\mathbf{B}}$ of the matrix \mathbf{B} in equation 14.7.

Inserting the matrices \mathbf{M} , \mathbf{B}_m and $\hat{\mathbf{B}}$ equation 14.12 we can then correct the mistake made by the inverse model covariance \mathbf{M}^{-1} .

To judge the performance of this method I want to look at a χ^2 -like quantity. Let \mathbf{x}_i be the combined data vector of all measured correlation functions in simulation i . Then we can define the sample average

$$\bar{\mathbf{x}} = \frac{1}{N_s} \sum_i \mathbf{x}_i \quad (\text{C.1})$$

as well as the deviation vector

$$\Delta \mathbf{x}_i = \mathbf{x}_i - \bar{\mathbf{x}} . \quad (\text{C.2})$$

If \mathbf{C} is the covariance matrix of the \mathbf{x}_i and \mathbf{C}_Δ is the covariance matrix of the $\Delta \mathbf{x}_i$ the we have (Anderson, 2003)

$$\mathbf{C}_\Delta = \frac{N_s - 1}{N_s} \mathbf{C} . \quad (\text{C.3})$$

The quantities

$$\chi_i^2 \equiv \frac{N_s}{N_s - 1} \Delta \mathbf{x}_i^T \mathbf{C}^{-1} \Delta \mathbf{x}_i \quad (\text{C.4})$$

should then have a χ^2 -distribution where the number of degrees-of-freedom equals the dimension of the \mathbf{x}_i (at least if the \mathbf{x}_i have a Gaussian distribution, cf. the previous section). We can then compare the values of χ_i^2 obtained in each simulation with the alternative

quantities

$$\begin{aligned}\chi_{i,M}^2 &\equiv \frac{N_s}{N_s - 1} \Delta \mathbf{x}_i^T \mathbf{M}^{-1} \Delta \mathbf{x}_i \\ \chi_{i,\text{PME}}^2 &\equiv \frac{N_s}{N_s - 1} \Delta \mathbf{x}_i^T \hat{\boldsymbol{\Psi}}_{\text{2nd}} \Delta \mathbf{x}_i ,\end{aligned}\quad (\text{C.5})$$

where $\hat{\boldsymbol{\Psi}}_{\text{2nd}}$ is the correction of the modelled precision matrix derived with equation 14.12. The total number of data points in each data vector \mathbf{x}_i of this experiment was 44. This would be the expectation value of χ^2 if the \mathbf{x}_i have a perfectly Gaussian distribution. Using the 3 different precision matrices \mathbf{C}^{-1} , \mathbf{M}^{-1} and $\hat{\boldsymbol{\Psi}}_{\text{2nd}}$ we get the average values

$$\begin{aligned}\frac{1}{N_s} \sum_i \chi_i^2 &\approx 45.1 \\ \frac{1}{N_s} \sum_i \chi_{i,M}^2 &\approx 56.6 \\ \frac{1}{N_s} \sum_i \chi_{i,\text{PME}}^2 &\approx 46.8 .\end{aligned}\quad (\text{C.6})$$

This shows that the PME method significantly improves the overestimated χ^2 values that are obtained from the wrong model covariance \mathbf{M} . This can also be seen from figure C.1 where I compare the histograms of χ^2 values obtained from the 3 different methods.

Appendix D

Full author list of Friedrich et al. (2018)

O. Friedrich

Universitäts-Sternwarte, Fakultät für Physik, Ludwig-Maximilians Universität München, Scheinerstr. 1, 81679 München, Germany

Max Planck Institute for Extraterrestrial Physics, Giessenbachstrasse, 85748 Garching, Germany

D. Gruen

Kavli Institute for Particle Astrophysics & Cosmology, P. O. Box 2450, Stanford University, Stanford, CA 94305, USA

SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA

J. DeRose

Department of Physics, Stanford University, 382 Via Pueblo Mall, Stanford, CA 94305, USA

Kavli Institute for Particle Astrophysics & Cosmology, P. O. Box 2450, Stanford University, Stanford, CA 94305, USA

D. Kirk

Department of Physics & Astronomy, University College London, Gower Street, London, WC1E 6BT, UK

A. E. Krause

Kavli Institute for Particle Astrophysics & Cosmology, P. O. Box 2450, Stanford University, Stanford, CA 94305, USA

T. McClintock

Department of Physics, University of Arizona, Tucson, AZ 85721, USA

E. S. Rykoff

Kavli Institute for Particle Astrophysics & Cosmology, P. O. Box 2450, Stanford University, Stanford, CA 94305, USA

SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA

S. Seitz

Universitäts-Sternwarte, Fakultät für Physik, Ludwig-Maximilians Universität München, Scheinerstr. 1, 81679 München, Germany

Max Planck Institute for Extraterrestrial Physics, Giessenbachstrasse, 85748 Garching, Germany

R. H. Wechsler

Department of Physics, Stanford University, 382 Via Pueblo Mall, Stanford, CA 94305, USA

Kavli Institute for Particle Astrophysics & Cosmology, P. O. Box 2450, Stanford University, Stanford, CA 94305, USA

SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA

G. M. Bernstein

Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104, USA

J. Blazek

Center for Cosmology and Astro-Particle Physics, The Ohio State University, Columbus, OH 43210, USA

Institute of Physics, Laboratory of Astrophysics, École Polytechnique Fédérale de Lausanne (EPFL), Observatoire de Sauverny, 1290 Versoix, Switzerland

C. Chang

Kavli Institute for Cosmological Physics, University of Chicago, Chicago, IL 60637, USA

S. Hilbert

Faculty of Physics, Ludwig-Maximilians-Universität, Scheinerstr. 1, 81679 Munich, Germany

Excellence Cluster Universe, Boltzmannstr. 2, 85748 Garching, Germany

B. Jain

Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104, USA

A. Kovacs

Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Campus UAB, 08193 Bellaterra (Barcelona) Spain

O. Lahav

Department of Physics & Astronomy, University College London, Gower Street, London, WC1E 6BT, UK

F. B. Abdalla

Department of Physics & Astronomy, University College London, Gower Street, London, WC1E 6BT, UK

Department of Physics and Electronics, Rhodes University, PO Box 94, Grahamstown, 6140, South Africa

S. Allam

Fermi National Accelerator Laboratory, P. O. Box 500, Batavia, IL 60510, USA

J. Annis

Fermi National Accelerator Laboratory, P. O. Box 500, Batavia, IL 60510, USA

K. Bechtol

LSST, 933 North Cherry Avenue, Tucson, AZ 85721, USA

A. Benoit-Lévy

CNRS, UMR 7095, Institut d'Astrophysique de Paris, F-75014, Paris, France

Department of Physics & Astronomy, University College London, Gower Street, London, WC1E 6BT, UK

Sorbonne Universités, UPMC Univ Paris 06, UMR 7095, Institut d'Astrophysique de Paris, F-75014, Paris, France

E. Bertin

CNRS, UMR 7095, Institut d'Astrophysique de Paris, F-75014, Paris, France

Sorbonne Universités, UPMC Univ Paris 06, UMR 7095, Institut d'Astrophysique de Paris, F-75014, Paris, France

D. Brooks

Department of Physics & Astronomy, University College London, Gower Street, London, WC1E 6BT, UK

A. Carnero Rosell

Laboratório Interinstitucional de e-Astronomia - LIneA, Rua Gal. José Cristino 77, Rio de Janeiro, RJ - 20921-400, Brazil

Observatório Nacional, Rua Gal. José Cristino 77, Rio de Janeiro, RJ - 20921-400, Brazil

M. Carrasco Kind

Department of Astronomy, University of Illinois, 1002 W. Green Street, Urbana, IL 61801, USA

National Center for Supercomputing Applications, 1205 West Clark St., Urbana, IL 61801, USA

J. Carretero

Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Campus UAB, 08193 Bellaterra (Barcelona) Spain

C. E. Cunha

Kavli Institute for Particle Astrophysics & Cosmology, P. O. Box 2450, Stanford University, Stanford, CA 94305, USA

C. B. D'Andrea

Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104, USA

L. N. da Costa

Laboratório Interinstitucional de e-Astronomia - LIneA, Rua Gal. José Cristino 77, Rio de Janeiro, RJ - 20921-400, Brazil

Observatório Nacional, Rua Gal. José Cristino 77, Rio de Janeiro, RJ - 20921-400, Brazil

C. Davis

Kavli Institute for Particle Astrophysics & Cosmology, P. O. Box 2450, Stanford University, Stanford, CA 94305, USA

S. Desai

Department of Physics, IIT Hyderabad, Kandi, Telangana 502285, India

H. T. Diehl

Fermi National Accelerator Laboratory, P. O. Box 500, Batavia, IL 60510, USA

J. P. Dietrich

Faculty of Physics, Ludwig-Maximilians-Universität, Scheinerstr. 1, 81679 Munich, Germany

Excellence Cluster Universe, Boltzmannstr. 2, 85748 Garching, Germany

A. Drlica-Wagner

Fermi National Accelerator Laboratory, P. O. Box 500, Batavia, IL 60510, USA

T. F. Eifler

Department of Physics, California Institute of Technology, Pasadena, CA 91125, USA

Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr., Pasadena, CA 91109, USA

P. Fosalba

Institute of Space Sciences, IEEC-CSIC, Campus UAB, Carrer de Can Magrans, s/n, 08193 Barcelona, Spain

J. Frieman

Fermi National Accelerator Laboratory, P. O. Box 500, Batavia, IL 60510, USA

Kavli Institute for Cosmological Physics, University of Chicago, Chicago, IL 60637, USA

J. García-Bellido

Instituto de Fisica Teorica UAM/CSIC, Universidad Autonoma de Madrid, 28049 Madrid, Spain

E. Gaztanaga

Institute of Space Sciences, IEEC-CSIC, Campus UAB, Carrer de Can Magrans, s/n, 08193 Barcelona, Spain

D. W. Gerdes

Department of Astronomy, University of Michigan, Ann Arbor, MI 48109, USA

Department of Physics, University of Michigan, Ann Arbor, MI 48109, USA

T. Giannantonio

Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK

Kavli Institute for Cosmology, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK

Universitäts-Sternwarte, Fakultät für Physik, Ludwig-Maximilians Universität München, Scheinerstr. 1, 81679 München, Germany

R. A. Gruendl

Department of Astronomy, University of Illinois, 1002 W. Green Street, Urbana, IL 61801, USA

National Center for Supercomputing Applications, 1205 West Clark St., Urbana, IL 61801, USA

J. Gschwend

Laboratório Interinstitucional de e-Astronomia - LIneA, Rua Gal. José Cristino 77, Rio de Janeiro, RJ - 20921-400, Brazil

Observatório Nacional, Rua Gal. José Cristino 77, Rio de Janeiro, RJ - 20921-400, Brazil

G. Gutierrez

Fermi National Accelerator Laboratory, P. O. Box 500, Batavia, IL 60510, USA

K. Honscheid

Center for Cosmology and Astro-Particle Physics, The Ohio State University, Columbus, OH 43210, USA

Department of Physics, The Ohio State University, Columbus, OH 43210, USA

D. J. James

Astronomy Department, University of Washington, Box 351580, Seattle, WA 98195, USA

M. Jarvis

Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104, USA

K. Kuehn

Australian Astronomical Observatory, North Ryde, NSW 2113, Australia

N. Kuropatkin

Fermi National Accelerator Laboratory, P. O. Box 500, Batavia, IL 60510, USA

M. Lima

Departamento de Física Matemática, Instituto de Física, Universidade de São Paulo, CP 66318, São Paulo, SP, 05314-970, Brazil

Laboratório Interinstitucional de e-Astronomia - LIneA, Rua Gal. José Cristino 77, Rio de Janeiro, RJ - 20921-400, Brazil

M. March

Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104, USA

J. L. Marshall

George P. and Cynthia Woods Mitchell Institute for Fundamental Physics and Astronomy, and Department of Physics and Astronomy, Texas A&M University, College Station, TX 77843, USA

P. Melchior

Department of Astrophysical Sciences, Princeton University, Peyton Hall, Princeton, NJ 08544, USA

F. Menanteau

Department of Astronomy, University of Illinois, 1002 W. Green Street, Urbana, IL 61801, USA

National Center for Supercomputing Applications, 1205 West Clark St., Urbana, IL 61801, USA

R. Miquel

Institució Catalana de Recerca i Estudis Avançats, E-08010 Barcelona, Spain

Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Campus UAB, 08193 Bellaterra (Barcelona) Spain

J. J. Mohr

Excellence Cluster Universe, Boltzmannstr. 2, 85748 Garching, Germany

Faculty of Physics, Ludwig-Maximilians-Universität, Scheinerstr. 1, 81679 Munich, Germany

Max Planck Institute for Extraterrestrial Physics, Giessenbachstrasse, 85748 Garching, Germany

B. Nord

Fermi National Accelerator Laboratory, P. O. Box 500, Batavia, IL 60510, USA

A. A. Plazas

Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr., Pasadena, CA 91109, USA

E. Sanchez

Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), Madrid, Spain

V. Scarpine

Fermi National Accelerator Laboratory, P. O. Box 500, Batavia, IL 60510, USA

R. Schindler

SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA

M. Schubnell

Department of Physics, University of Michigan, Ann Arbor, MI 48109, USA

I. Sevilla-Noarbe
Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), Madrid, Spain

E. Sheldon
Brookhaven National Laboratory, Bldg 510, Upton, NY 11973, USA

M. Smith
School of Physics and Astronomy, University of Southampton, Southampton, SO17 1BJ, UK

M. Soares-Santos
Fermi National Accelerator Laboratory, P. O. Box 500, Batavia, IL 60510, USA

F. Sobreira
Instituto de Física Gleb Wataghin, Universidade Estadual de Campinas, 13083-859, Campinas, SP, Brazil
Laboratório Interinstitucional de e-Astronomia - LIneA, Rua Gal. José Cristino 77, Rio de Janeiro, RJ - 20921-400, Brazil

E. Suchyta
Computer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831

M. E. C. Swanson
National Center for Supercomputing Applications, 1205 West Clark St., Urbana, IL 61801, USA

G. Tarle
Department of Physics, University of Michigan, Ann Arbor, MI 48109, USA

D. Thomas
Institute of Cosmology & Gravitation, University of Portsmouth, Portsmouth, PO1 3FX, UK

M. A. Troxel
Center for Cosmology and Astro-Particle Physics, The Ohio State University, Columbus, OH 43210, USA
Department of Physics, The Ohio State University, Columbus, OH 43210, USA

V. Vikram
Argonne National Laboratory, 9700 South Cass Avenue, Lemont, IL 60439, USA

J. Weller
Excellence Cluster Universe, Boltzmannstr. 2, 85748 Garching, Germany
Max Planck Institute for Extraterrestrial Physics, Giessenbachstrasse, 85748 Garching, Germany
Universitäts-Sternwarte, Fakultät für Physik, Ludwig-Maximilians Universität München, Scheinerstr. 1, 81679 München, Germany

Bibliography

Abbott T., et al., 2016, Phys. Rev. D, 94, 022001

Abroe M. E., et al., 2002, MNRAS, 334, 11

Adler R. J., Casey B., Jacob O. C., 1995, American Journal of Physics, 63, 620

Akbar M., Cai R.-G., 2006, Physics Letters B, 635, 7

Alam S., et al., 2017, MNRAS, 470, 2617

Anderson T., 2003, An Introduction to Multivariate Statistical Analysis. Wiley Series in Probability and Statistics, Wiley, <https://books.google.de/books?id=Cmm9QgAACAAJ>

Anselmi S., Starkman G. D., Sheth R. K., 2016, MNRAS, 455, 2474

Baldauf T., Seljak U., Smith R. E., Hamaus N., Desjacques V., 2013, Phys. Rev. D, 88, 083507

Barreira A., Schmidt F., 2017, JCAP, 11, 051

Barreira A., Bose S., Li B., Llinares C., 2017, JCAP, 2, 031

Bartelmann M., Schneider P., 2001, Physics Reports, 340, 291

Becker M. R., 2013, MNRAS, 435, 115

Becker M. R., et al., 2016, Phys. Rev. D, 94, 022002

Bekenstein J. D., 1973, Phys. Rev. D, 7, 2333

Bender R., et al., 2001, in Cristiani S., Renzini A., Williams R. E., eds, Deep Fields. p. 96, doi:10.1007/10854354'18

Benítez N., 2000, ApJ, 536, 571

Benjamin J., et al., 2007, MNRAS, 381, 702

Bernardeau F., 1994, A&A, 291, 697

Bernardeau F., 1995, *A&A*, 301, 309

Bernardeau F., Valageas P., 2000, *A&A*, 364, 1

Bernardeau F., van Waerbeke L., Mellier Y., 1997, *A&A*, 322, 1

Bernardeau F., Colombi S., Gaztañaga E., Scoccimarro R., 2002, *Physics Reports*, 367, 1

Bernardeau F., Pichon C., Codis S., 2014, *Phys. Rev. D*, 90, 103519

Bernardeau F., Codis S., Pichon C., 2015, *MNRAS*, 449, L105

Bethe H. A., 1947, *Physical Review*, 72, 339

Birkhoff G. D., Langer R. E., 1923, *Relativity and modern physics*. Harvard University Press

Blazek J., MacCrann N., Troxel M. A., Fang X., 2017, preprint, ([arXiv:1708.09247](https://arxiv.org/abs/1708.09247))

Bolton A. S., et al., 2012, *AJ*, 144, 144

Borisov A., Jain B., 2009, *Phys. Rev. D*, 79, 103506

Brainerd T. G., Blandford R. D., Smail I., 1996, *ApJ*, 466, 623

Brimioulle F., Seitz S., Lerchster M., Bender R., Snigula J., 2013, *MNRAS*, 432, 1046

Brouwer M. M., et al., 2017, *MNRAS*, 466, 2547

Brouwer M. M., et al., 2018, preprint, ([arXiv:1805.00562](https://arxiv.org/abs/1805.00562))

Buchalter A., Kamionkowski M., Jaffe A. H., 2000, *ApJ*, 530, 36

Cabré A., Fosalba P., Gaztañaga E., Manera M., 2007, *Monthly Notices of the Royal Astronomical Society*, 381, 1347

Cacciato M., Lahav O., van den Bosch F. C., Hoekstra H., Dekel A., 2012, *MNRAS*, 426, 566

Cai R.-G., Kim S. P., 2005, *Journal of High Energy Physics*, 2, 050

Cai R.-G., Cao L.-M., Hu Y.-P., 2009, *Classical and Quantum Gravity*, 26, 155018

Cai R.-G., Sun S., Zhang Y.-L., 2017, preprint, ([arXiv:1712.09326](https://arxiv.org/abs/1712.09326))

Cao C., Carroll S. M., Michalakis S., 2017, *Phys. Rev. D*, 95, 024031

Capozziello S., Luongo O., 2018, *International Journal of Modern Physics D*, 27, 1850029

Chiang C.-T., Wagner C., Sánchez A. G., Schmidt F., Komatsu E., 2015, *Journal of Cosmology and Astroparticle Physics*, 9, 028

Clampitt J., et al., 2017, *MNRAS*, 465, 4204

Clarkson C., Clifton T., February S., 2009, *JCAP*, 6, 025

Clerkin L., Kirk D., Lahav O., Abdalla F. B., Gaztañaga E., 2015, *MNRAS*, 448, 1389

Clerkin L., et al., 2017, *MNRAS*, 466, 1444

Codis S., Pichon C., Bernardeau F., Uhlemann C., Prunet S., 2016a, *MNRAS*, 460, 1549

Codis S., Bernardeau F., Pichon C., 2016b, *MNRAS*, 460, 1598

Collister A. A., Lahav O., 2004, *PASP*, 116, 345

Conroy C., Wechsler R. H., Kravtsov A. V., 2006, *ApJ*, 647, 201

Cooper M. C., et al., 2011, *ApJS*, 193, 14

Cooray A., Hu W., 2001, *The Astrophysical Journal*, 554, 56

Crocce M., Pueblas S., Scoccimarro R., 2006, *MNRAS*, 373, 369

Crocce M., Cabré A., Gaztañaga E., 2011, *Monthly Notices of the Royal Astronomical Society*, 414, 329

Crocce M., et al., 2016, *Monthly Notices of the Royal Astronomical Society*, 455, 4301

DES Collaboration et al., 2017, preprint, ([arXiv:1708.01530](https://arxiv.org/abs/1708.01530))

Dai D.-C., Stojkovic D., 2017, *Journal of High Energy Physics*, 11, 7

Dark Energy Survey Collaboration et al., 2016, *MNRAS*, 460, 1270

Dark Energy Survey Collaboration et al., 2018, preprint, ([arXiv:1801.03181](https://arxiv.org/abs/1801.03181))

DeRose J., Wechsler R., Rykoff E., et al., 2018, in prep.

Dekel A., Lahav O., 1999, *ApJ*, 520, 24

Dewitt B. S., 1967, *Physical Review*, 160, 1113

Dodelson S., Schneider M. D., 2013, *Physical Review D*, 88, 063537

Dvornik A., et al., 2018, *MNRAS*, 479, 1240

Efron B., 1982, *The Jackknife, the Bootstrap and other resampling plans*. Society for Industrial and Applied Mathematics

Efstathiou G., Lemos P., 2018, MNRAS, 476, 151

Eifler T., Schneider P., Hartlap J., 2009, Astronomy & Astrophysics, 502, 721

Eifler T., Krause E., Schneider P., Honscheid K., 2014, Monthly Notices of the Royal Astronomical Society, 440, 1379

Eisenstein D. J., Hu W., 1998, ApJ, 496, 605

Eling C., Guedens R., Jacobson T., 2006, Physical Review Letters, 96, 121301

Elvin-Poole J., et al., 2017, preprint, ([arXiv:1708.01536](https://arxiv.org/abs/1708.01536))

Flaugh B., 2005, International Journal of Modern Physics A, 20, 3121

Flaugh B., et al., 2015, AJ, 150, 150

Fließbach T., 1990, Allgemeine Relativitätstheorie. BI Wissenschaftsverlag

Fosalba P., Gaztanaga E., 1998, MNRAS, 301, 503

Friedrich O., Eifler T., 2018, MNRAS, 473, 4150

Friedrich O., Seitz S., Eifler T. F., Gruen D., 2016, MNRAS, 456, 2662

Friedrich O., et al., 2018, Phys. Rev. D, 98, 023508

Frieman J. A., Turner M. S., Huterer D., 2008, ARA&A, 46, 385

Fry J. N., 1984, ApJ, 279, 499

Fry J. N., 1985, ApJ, 289, 10

Fu L., et al., 2008, A&A, 479, 9

Gibbons G. W., Hawking S. W., 1977, Phys. Rev. D, 15, 2738

Górski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M., Bartelmann M., 2005, ApJ, 622, 759

Greisel N., Seitz S., Drory N., Bender R., Saglia R. P., Snigula J., 2013, ApJ, 768, 117

Gruen D., et al., 2013, MNRAS, 432, 1455

Gruen D., et al., 2014, MNRAS, 442, 1507

Gruen D., Friedrich O., The DES Collaboration 2016, Monthly Notices of the Royal Astronomical Society, 455, 3367

Gruen D., et al., 2018, Phys. Rev. D, 98, 023507

Hamaus N., Seljak U., Desjacques V., Smith R. E., Baldauf T., 2010, Phys. Rev. D, 82, 043515

Harnois-Déraps J., van Waerbeke L., 2015, MNRAS, 450, 2857

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

Hartle J. B., Hawking S. W., 1983, Phys. Rev. D, 28, 2960

Hawking S. W., 1974, Nature, 248, 30

Hawking S. W., 1975, Communications in Mathematical Physics, 43, 199

Hayward S. A., 1998, Classical and Quantum Gravity, 15, 3147

Heymans C., et al., 2013, MNRAS, 432, 2433

Hilbert S., Hartlap J., Schneider P., 2011, Astronomy & Astrophysics, 536, A85

Hilbert S., Xu D., Schneider P., Springel V., Vogelsberger M., Hernquist L., 2017, MNRAS, 468, 790

Hildebrandt H., et al., 2017, MNRAS, 465, 1454

Hirata C. M., et al., 2004, MNRAS, 353, 529

Hobson M. P., Jaffe A. H., Liddle A. R., Mukherjee P., Parkinson D., 2014, Bayesian Methods in Cosmology. Cambridge University Press

Hogg D. W., 1999, ArXiv Astrophysics e-prints,

Hoyle B., Rau M. M., Bonnett C., Seitz S., Weller J., 2015, MNRAS, 450, 305

Hoyle B., et al., 2018, MNRAS, 478, 592

Hubble E., 1929, Proceedings of the National Academy of Science, 15, 168

Hudson M. J., Gwyn S. D. J., Dahle H., Kaiser N., 1998, ApJ, 503, 531

Huff E., Mandelbaum R., 2017, preprint, ([arXiv:1702.02600](https://arxiv.org/abs/1702.02600))

Ivezic Z., et al., 2008, preprint, ([arXiv:0805.2366](https://arxiv.org/abs/0805.2366))

Jacobson T., 1995, Physical Review Letters, 75, 1260

Jain B., Zhang P., 2008, Phys. Rev. D, 78, 063503

Jee M. J., Tyson J. A., Hilbert S., Schneider M. D., Schmidt S., Wittman D., 2016, ApJ, 824, 77

Jeong D., Komatsu E., 2009, *ApJ*, 703, 1230

Joachimi B., 2017, *MNRAS*, 466, L83

Joachimi B., Schneider P., Eifler T., 2008, *Astronomy & Astrophysics*, 477, 43

Kacprzak T., et al., 2016, *MNRAS*, 463, 3653

Kaiser N., 1984, *ApJ*, 284, L9

Kaiser N., Squires G., 1993, *ApJ*, 404, 441

Kaufman G. M., 1967, Report No. 6710, Center for Operations Research and Econometrics, Catholic University of Louvain, Heverlee, Belgium

Kilbinger M., Schneider P., 2004, *Astronomy & Astrophysics*, 413, 465

Kilbinger M., et al., 2009, *A&A*, 497, 677

Kilbinger M., et al., 2013, *MNRAS*, 430, 2200

Kitching T. D., et al., 2012, *MNRAS*, 423, 3163

Kitching T. D., Alsing J., Heavens A. F., Jimenez R., McEwen J. D., Verde L., 2017, *MNRAS*, 469, 2737

Köhlinger F., et al., 2017, *MNRAS*, 471, 4412

Krause E., Eifler T., 2016, preprint, ([arXiv:1601.05779](https://arxiv.org/abs/1601.05779))

Krause E., Eifler T., Blazek J., 2016, *MNRAS*, 456, 207

Krause E., et al., 2017, preprint, ([arXiv:1706.09359](https://arxiv.org/abs/1706.09359))

Kwan J., et al., 2017, *MNRAS*, 464, 4045

Laureijs R., et al., 2011, preprint, ([arXiv:1110.3193](https://arxiv.org/abs/1110.3193))

Lemos P., Challinor A., Efstathiou G., 2017, *JCAP*, 5, 014

Letac G., Massam H., 2004, *Scandinavian Journal of Statistics*, 31, 295

Lewandowski M., Perko A., Senatore L., 2015, *JCAP*, 5, 019

Lewis A., Challinor A., Lasenby A., 2000, *ApJ*, 538, 473

Limber D. N., 1953, *ApJ*, 117, 134

Lin C.-A., Kilbinger M., 2015, *A&A*, 576, A24

Liu X., et al., 2015, *MNRAS*, 450, 2888

Loh J. M., 2008, *ApJ*, 681, 726

Lue A., Scoccimarro R., Starkman G., 2004, *Phys. Rev. D*, 69, 044005

MacCrann N., et al., 2018, *MNRAS*,

Mandelbaum R., Seljak U., Kauffmann G., Hirata C. M., Brinkmann J., 2006, *MNRAS*, 368, 715

Mandelbaum R., Slosar A., Baldauf T., Seljak U., Hirata C. M., Nakajima R., Reyes R., Smith R. E., 2013, *MNRAS*, 432, 1544

Mantz A. B., et al., 2016, *MNRAS*, 463, 3582

Martin S., Schneider P., Simon P., 2012, *Astronomy & Astrophysics*, 540, A9

Martinet N., et al., 2018, *MNRAS*, 474, 712

McClintock T., et al., 2018, preprint, ([arXiv:1805.00039](https://arxiv.org/abs/1805.00039))

Milgrom M., 2001, *Acta Physica Polonica B*, 32, 3613

Miller L., et al., 2013, *MNRAS*, 429, 2858

Miyazaki S., et al., 2002, *PASJ*, 54, 833

Mukhanov V., 2005, *Physical Foundations of Cosmology*. Cambridge University Press, doi:10.2277/0521563984

Mukhanov V., 2013, *European Physical Journal C*, 73, 2486

Multamäki T., Gaztañaga E., Manera M., 2003, *MNRAS*, 344, 761

N. Bronstein I., A. Semendjajew K., 1979, *Taschenbuch der Mathematik*, 19 edn. BSB B. G. Teubner Verlagsgesellschaft, Nauka-Verlag, Leipzig, Moskau

Neyman J., 1937, *Philosophical Transactions of the Royal Society of London Series A*, 236, 333

Norberg P., Baugh C. M., Gaztañaga E., Croton D. J., 2009, *Monthly Notices of the Royal Astronomical Society*, 396, 19

Nordman D. J., Lahiri S. N., 2007, *The Indian Journal of Statistics*, 69, 468

Odom B., Hanneke D., D'Urso B., Gabrielse G., 2006, *Phys. Rev. Lett.*, 97, 030801

Padmanabhan T., 2002, *Modern Physics Letters A*, 17, 923

Padmanabhan T., 2010a, *Gravitation: Foundations and Frontiers*. Cambridge University Press

Padmanabhan T., 2010b, Reports on Progress in Physics, 73, 046901

Padmanabhan T., Padmanabhan H., 2017, Physics Letters B, 773, 81

Padmanabhan N., White M., Zhou H. H., O'Connell R., 2016, MNRAS, 460, 1567

Park Y., et al., 2016, Phys. Rev. D, 94, 063533

Parsa M., Eckart A., Shahzamanian B., Karas V., Zajaček M., Zensus J. A., Straubmeier C., 2017, ApJ, 845, 22

Paz D. J., Sánchez A. G., 2015, MNRAS, 454, 4326

Peebles P. J. E., 1993, Principles of Physical Cosmology. Princeton University Press

Peng Y.-j., Lilly S. J., Renzini A., Carollo M., 2012, ApJ, 757, 4

Pires S., Leonard A., Starck J.-L., 2012, MNRAS, 423, 983

Planck Collaboration et al., 2014, A&A, 571, A15

Planck Collaboration et al., 2016, A&A, 594, A13

Planck Collaboration et al., 2018, preprint, ([arXiv:1807.06211](https://arxiv.org/abs/1807.06211))

Plehn T., 2017, preprint, ([arXiv:1705.01987](https://arxiv.org/abs/1705.01987))

Pope A. C., Szapudi I., 2008, MNRAS, 389, 766

Prat J., et al., 2017, preprint, ([arXiv:1708.01537](https://arxiv.org/abs/1708.01537))

Rau M. M., Seitz S., Brimioule F., Frank E., Friedrich O., Gruen D., Hoyle B., 2015, MNRAS, 452, 3710

Reddick R. M., Wechsler R. H., Tinker J. L., Behroozi P. S., 2013, ApJ, 771, 30

Refregier A., Gamper L., Amara A., Heisenberg L., 2017, preprint, ([arXiv:1708.05177](https://arxiv.org/abs/1708.05177))

Rehmann R. L., et al., 2018, preprint, ([arXiv:1806.10614](https://arxiv.org/abs/1806.10614))

Riess A. G., et al., 1998, The Astronomical Journal, 116, 1009

Ross A. J., Percival W. J., Crocce M., Cabré A., Gaztañaga E., 2011, MNRAS, 415, 2193

Rovelli C., 2000, ArXiv General Relativity and Quantum Cosmology e-prints,

Rozo E., et al., 2016, MNRAS, 461, 1431

Sato M., Hamana T., Takahashi R., Takada M., Yoshida N., Matsubara T., Sugiyama N., 2009, ApJ, 701, 945

Sato M., Takada M., Hamana T., Matsubara T., 2011, *The Astrophysical Journal*, 734, 76

Schneider P., 1985, *Astronomy and Astrophysics*, 143, 413

Schneider P., 2005, ArXiv Astrophysics e-prints,

Schneider P., Ehlers J., Falco E. E., 1992, *Gravitational Lenses*. Springer-Verlag Berlin Heidelberg New York, doi:10.1007/978-3-662-03758-4

Schneider P., van Waerbeke L., Kilbinger M., Mellier Y., 2002, *Astronomy & Astrophysics*, 396, 1

Schrabback T., et al., 2010, *A&A*, 516, A63

Schwarzschild K., 1916, in *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin, Phys.-Math. Klasse*, 424-434 (1916).

Seitz C., Schneider P., 1997, *A&A*, 318, 687

Seitz S., Schneider P., Ehlers J., 1994, *Classical and Quantum Gravity*, 11, 2345

Seljak U., Warren M. S., 2004, *Monthly Notices of the Royal Astronomical Society*, 355, 129

Sellentin E., Heavens A. F., 2016, *MNRAS*, 456, L132

Sellentin E., Heavens A. F., 2017, *MNRAS*, 464, 4658

Sellentin E., Heavens A. F., 2018, *MNRAS*, 473, 2355

Sellentin E., Heymans C., Harnois-Déraps J., 2018, *MNRAS*, 477, 4879

Semboloni E., van Waerbeke L., Heymans C., Hamana T., Colombi S., White M., Mellier Y., 2007, *Monthly Notices of the Royal Astronomical Society*, 375, L6

Semboloni E., Schrabback T., van Waerbeke L., Vafaei S., Hartlap J., Hilbert S., 2011, *MNRAS*, 410, 143

Senatore L., 2015, *JCAP*, 11, 007

Shan H., et al., 2018, *MNRAS*, 474, 1116

Sheldon E. S., Huff E. M., 2017, *ApJ*, 841, 24

Simon P., King L. J., Schneider P., 2004, *Astronomy & Astrophysics*, 417, 873

Simpson F., James J. B., Heavens A. F., Heymans C., 2011, *Physical Review Letters*, 107, 271301

Simpson F., Heavens A. F., Heymans C., 2013, *Phys. Rev. D*, 88, 083510

Simpson F., Harnois-Déraps J., Heymans C., Jimenez R., Joachimi B., Verde L., 2016, Monthly Notices of the Royal Astronomical Society, 456, 278

Smith R. E., et al., 2003, Monthly Notices of the Royal Astronomical Society, 341, 1311

Springel V., 2005, MNRAS, 364, 1105

Springel V., et al., 2005, Nature, 435, 629

Stebbins A., 1996, ArXiv Astrophysics e-prints

Takada M., Jain B., 2002, MNRAS, 337, 875

Takada M., Jain B., 2009, Monthly Notices of the Royal Astronomical Society, 395, 2065

Takahashi R., et al., 2009, ApJ, 700, 479

Takahashi R., Sato M., Nishimichi T., Taruya A., Oguri M., 2012, ApJ, 761, 152

Takahashi R., Soma S., Takada M., Kayo I., 2014, Monthly Notices of the Royal Astronomical Society, 444, 3473

Taruya A., Takada M., Hamana T., Kayo I., Futamase T., 2002, The Astrophysical Journal, 571, 638

Taylor A., Joachimi B., 2014, Monthly Notices of the Royal Astronomical Society, 442, 2728

Taylor A., Joachimi B., Kitching T., 2013, Monthly Notices of the Royal Astronomical Society, 432, 1928

Tegmark M., Rees M. J., 1998, ApJ, 499, 526

Tegmark M., SDSS Collaboration 2002, in American Astronomical Society Meeting Abstracts #200. p. 777

The Dark Energy Survey Collaboration 2005, ArXiv Astrophysics e-prints,

Thomas S., Abdalla F., Lahav O., 2011, Monthly Notices of the Royal Astronomical Society, 412, 1669

Trotta R., 2017, preprint, ([arXiv:1701.01467](https://arxiv.org/abs/1701.01467))

Troxel M. A., Ishak M., 2015, Phys. Rep., 558, 1

Troxel M. A., et al., 2017, preprint, ([arXiv:1708.01538](https://arxiv.org/abs/1708.01538))

Troxel M. A., et al., 2018, MNRAS, 479, 4998

Uhlemann C., et al., 2018a, MNRAS, 473, 5098

Uhlemann C., Pichon C., Codis S., L'Huillier B., Kim J., Bernardeau F., Park C., Prunet S., 2018b, MNRAS, 477, 2772

Umetsu K., 2010, preprint, ([arXiv:1002.3952](https://arxiv.org/abs/1002.3952))

Unruh W. G., 1976, Phys. Rev. D, 14, 870

Valageas P., 2002a, A&A, 382, 412

Valageas P., 2002b, A&A, 382, 477

Vale C., White M., 2003, The Astrophysical Journal, 592, 699

Van Waerbeke L., et al., 2000, A&A, 358, 30

Verlinde E., 2011, Journal of High Energy Physics, 4, 29

Verlinde E. P., 2016, preprint, ([arXiv:1611.02269](https://arxiv.org/abs/1611.02269))

Wang Y., Brunner R. J., Dolence J. C., 2013, Monthly Notices of the Royal Astronomical Society, 432, 1961

Wang Y., et al., 2017, MNRAS, 469, 3762

Wechsler R., DeRose J., Busha M., et al., 2018, in prep.

Wilson G., Kaiser N., Luppino G. A., Cowie L. L., 2001, ApJ, 555, 572

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

Xavier H. S., Abdalla F. B., Joachimi B., 2016, MNRAS, 459, 3693

Yèche C., Ealet A., Réfrégier A., Tao C., Tilquin A., Virey J.-M., Yvon D., 2006, A&A, 448, 831

Zibin J. P., 2008, Phys. Rev. D, 78, 043504

Zuntz J., et al., 2017a, preprint, ([arXiv:1708.01533](https://arxiv.org/abs/1708.01533))

Zuntz J., et al., 2017b, preprint, ([arXiv:1708.01533](https://arxiv.org/abs/1708.01533))

de Putter R., Takada M., 2010, Phys. Rev. D, 82, 103522

de Simoni F., et al., 2013, Monthly Notices of the Royal Astronomical Society, 435, 3017

van Uitert E., Hoekstra H., Velander M., Gilbank D. G., Gladders M. D., Yee H. K. C., 2011, A&A, 534, A14

van Uitert E., et al., 2018, MNRAS, 476, 4662

Curriculum Vitæ

Oliver Friedrich

born on 12/02/1987 in Potsdam (Germany).

Nationality: German

Current address: Konigsteinstr. 12, 80807 Munich Germany

Education:

- 2000 to 2007:

student at Max-Reinhardt-Gymnasium, Berlin (equivalent to high school)

06/22/2007

completing the allgemeine Hochschulreife (high school degree)

- 2008 to 2012:

student at Ruprecht-Karls-University, Heidelberg

09/26/2011

completing the Bachelors degree in physics

08/31/2012

completing the Bachelors degree in mathematics

- 2012 to 2014:

student at Ludwig-Maximilians-University, Munich

09/19/2014

completing the Masters degree in physics

- Since October 2014:

PhD student of physics at the Ludwig-Maximilians-University, Munich

(supervised by Dr. Stella Seitz)

04/04/2016 to 07/27/2016

working with the group of Prof. Risa Wechsler at Stanford University, supported by a DAAD short term scholarship

Tutoring:

- Winter semester 2011/2012:
tutoring students in "higher Analysis (Analysis III)"
- Summer semester 2013:
tutoring students in "Mathematics for Physicists II"
- Winter semester 2013/2014:
tutoring students in "Mathematics for Physicists III"
- Since October 2014:
assisting in the supervision of Bachelor and Master students
- Summer semester 2017:
tutoring students in "Formation and Evolution of Cosmic Structures"

Grants / awards:

- Dark Energy Survey (DES) early career scientist award (1000\$) to support participation in the Summer 2015 DES collaboration meeting
- DAAD short term scholarship to cover living expenses for a 14 weeks visit to Stanford University

Employments:

- 2007/2008:
voluntary ecological year at the Stiftung Naturschutz Berlin
- April 2009 to September 2012:
working as student employee at the Cafe Botanic (Heidelberg)
since January 2011
promotion to shift supervisor at the Cafe Botanic

Publication List

Major / Leading Contributions:

- D. Gruen; **O. Friedrich**; et al. (2018)
Density split statistics: Cosmological constraints from counts and lensing in cells in DES Y1 and SDSS
Physical Review D, Volume 98, Issue 2, id.023507
arxiv:1710.05045
- **O. Friedrich**; D. Gruen; et al. (2018)
Density split statistics: joint model of counts and lensing in cells
Physical Review D, Volume 98, Issue 2, id.023508
arxiv:1710.05162
- **O. Friedrich**; T. F. Eifler (2017)
Precision matrix expansion - efficient use of numerical simulations in estimating errors on cosmological parameters
Monthly Notices of the Royal Astronomical Society, Volume 473, Issue 3, p.4150-4163
arxiv:1703.07786
- D. Gruen; **O. Friedrich**; et al. (2016)
Weak lensing by galaxy troughs in DES Science Verification data
Monthly Notices of the Royal Astronomical Society, Volume 455, Issue 3, p.3367-3380
arxiv:1507.05090
- **O. Friedrich**, S. Seitz, T. F. Eifler, D. Gruen (2016)
Performance of internal Covariance Estimators for Cosmic Shear Correlation Functions
Monthly Notices of the Royal Astronomical Society, Volume 456, Issue 3, p.2662-2680
arxiv:1508.00895

Substantial Contributions:

- Dark Energy Survey Collaboration; T. Abbott; ... **O. Friedrich**; et al. (2017)
Dark Energy Survey Year 1 Results: Cosmological Constraints from Galaxy Clustering and Weak Lensing
submitted to Phys. Rev. D
arxiv:1708.01530
- E. Krause; T. F. Eifler; J. Zuntz; **O. Friedrich**; M. A. Troxel; S. Dodelson; J. Blazek; L. F. Secco; N. MacCrann; E. Baxter, C. Chang; N. Chen; M. Crocce; J. DeRose; A. Ferte; N. Kokron; F. Lacasa; V. Miranda; Y. Omori; A. Porredon; R. Rosenfeld; S. Samuroff; M. Wang; R. H. Wechsler; et al (2017)
Dark Energy Survey Year 1 Results: Multi-Probe Methodology and Simulated Likelihood Analyses
submitted to Phys. Rev. D
arxiv:1706.09359
- T. Kacprzak; D. Kirk; **O. Friedrich**; A. Amara; A. Refregier; L. Marian; J. P. Dietrich; E. Suchyta; et al. (2016)
Cosmology constraints from shear peak statistics in Dark Energy Survey Science Verification data
Monthly Notices of the Royal Astronomical Society, Volume 463, Issue 4, p.3653-3673
arxiv:1603.05040
- M. R. Becker; M. A. Troxel; N. MacCrann; E. Krause; T. F. Eifler; **O. Friedrich**; A. Nicola; et al. (2016)
Cosmic Shear Measurements with DES Science Verification Data
Physical Review D, Volume 94, Issue 2, id.022002
arxiv:1507.05598
- M. M. Rau, S. Seitz, F. Brimioule, E. Frank, **O. Friedrich**, D. Gruen, B. Hoyle (2015)
Accurate photometric redshift probability density estimation - method comparison and application
Monthly Notices of the Royal Astronomical Society, Volume 452, Issue 4, p.3710-3725
arxiv:1503.08215

- D. Gruen, S. Seitz, M. R. Becker, **O. Friedrich**, A. Mana (2015)
 Cosmic variance of the galaxy cluster weak lensing signal
 Monthly Notices of the Royal Astronomical Society, Volume 449, Issue 4, p.4264-4276
 arxiv:1501.01632

Minor Contributions:

- M. A. Troxel, E. Krause, C. Chang, T. F. Eifler, **O. Friedrich**, D. Gruen, N. MacCrann et al. (2017)
 Survey geometry and the internal consistency of recent cosmic shear measurements
 Monthly Notices of the Royal Astronomical Society, Volume 479, Issue 4, p.4998-5004
 arxiv:1804.10663
- M. A. Troxel; N. MacCrann; J. Zuntz; T. F. Eifler; E. Krause; S. Dodelson; D. Gruen; J. Blazek; **O. Friedrich**; S. Samuroff; J. Prat; L. F. Secco; C. Davis; A. Ferte; J. DeRose; et al. (2017)
 Dark Energy Survey Year 1 Results: Cosmological Constraints from Cosmic Shear
 Physical Review D, Volume 98, Issue 4, id.043528
 arxiv:1708.01538
- Prat, J.; Sanchez, C.; ... ; **O. Friedrich**; et al. (2017)
 Dark Energy Survey Year 1 Results: Galaxy-Galaxy Lensing
 Physical Review D, Volume 98, Issue 4, id.042005
 arxiv:1708.01537
- Elvin-Poole, J.; ... ; **O. Friedrich**; et al. (2017)
 Dark Energy Survey Year 1 Results: Galaxy clustering for combined probes
 Physical Review D, Volume 98, Issue 4, id.042006
 arxiv:1708.01536
- Dark Energy Survey Collaboration; T. Abbott; ... ; **O. Friedrich**; et al. (2016)
 Cosmology from Cosmic Shear with DES Science Verification Data
 Physical Review D, Volume 94, Issue 2, id.022001
 arxiv:1507.05552

Note of Thanks

I want to thank my parents and grandparents for their support during my studies. Also I am grateful to my entire Family - Heike & Frank, Helga & Günter, Isabell & René, Franziska & Enrico, Olaf & Antje - for making my visits to Berlin and Potsdam always enjoyable.

I am happy to call Laura Adler my wife. She provides me with orientation when I am overwhelmed with academic stress and she is my most powerful motivation. Also, she is constantly correcting my perspective on the importance or unimportance of things.

Great thanks also goes to Laura's parents Petra and Wolfgang and to her sister Antonia. The time I spent with them was always an enjoyable break from everyday life.

I am grateful for the supervision by Dr. Stella Seitz. Her confidence in me has helped me grow as a scientist and her support of my work within the DES Collaboration has provided me with great research opportunities. Our discussions have prevented my thinking from becoming too one-sided and she made me appreciate the topic of galaxy evolution in the context of LSS cosmology. I also want to thank Prof. Dr. Ralf Bender for his continued support.

I am happy to have found an incredible colleague and friend in Dr. Daniel Gruen.

I am grateful for the company of Romy Rehmann and Tamas Varga. I always enjoy our conversations and find our exchange to be most fruitful. I hope that we will continue this when I am overseas.

I want to thank Tim Eifler and Elisabeth Krause for their continued support. I have learned a lot from them.

I want to thank Risa Wechsler for supporting and welcoming me during my time in Stanford. Also I am grateful to Anna Ogorzałek, Ondrej Orban and Joe DeRose as well as Renata and Daniel Gruen for making my time there so enjoyable!

I am thankful to Ulrich Hopp for letting me run the Wendelstein Telecope every once in a while.

Finally, I want to thank Ben Hoyle, Steffen Hagstotz, Giorgia Pollina, Nico Hamaus, Tea Tutis, Matteo Costanzi, Kerstin Paech and Jochen Weller for interesting discussions, light lunch breaks and for great times in Ringberg!

This thesis was supported by SFB-Transregio 33 'The Dark Universe' by the Deutsche Forschungsgemeinschaft (DFG), the DFG cluster of excellence 'Origin and Structure of the Universe' and the Dark Energy Survey Collaboration. My stay in Stanford was supported by a short term scholarship by the Deutsche Akademische Auslandsdienst.