The non-Gaussian matter power spectrum covariance in the halo model approach

Dissertation

zur

Erlangung des Doktorgrades (Dr. rer. nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Jasmin Pielorz

aus

Leverkusen

Bonn 2008

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

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

Tag der Promotion: 11.7.2008

Contents

Introduction

Concordance model of cosmology

Recent years have brought rapid progress in observational cosmology, establishing it as a precision science where cosmological parameters can be determined to a high accuracy. The reason for this dramatic improvement is twofold: on the one hand, advances in observational techniques have produced a wealth of data and, on the other hand, cosmologists have developed a consistent theoretical framework to interpret them over the last decades. As a result, the combination of measurements from type Ia supernovae (SN), cosmic microwave background (CMB) anisotropies and baryon acoustic oscillations (BAO) leads to the current paradigm of an accelerating flat ΛCDM model in which the Universe is composed of 5% baryons, 23% dark matter and 72% dark energy [47]. In spite of these successes, the current cosmological model raises a number of issues: we have only limited knowledge of the physical nature of the two dark components, which together account for 95% of the energy content of the Universe. Whereas particle physicists came up with several possible dark matter candidates, we literally grope in the dark with an explanation for dark energy. Is it a cosmological constant, is it a new kind of field that evolves dynamically with the expansion of the Universe or is a new law of gravity needed? Apart from these fundamental physical questions, we need to refine our current theory for the origin and evolution of cosmic structure. Are the initial perturbations indeed randomly distributed as suggested by standard inflationary models, how can we explain the diversity of galaxies and the complex processes involved in their formation and how does gravitational clustering work on small scales?

Large-scale structure

One way to tackle these problems is to investigate the large scale distribution of (dark) matter in the Universe in more detail. It contains valuable information about fundamental cosmological parameters, the properties of dark matter and the formation processes of structure. If we study the distribution of matter at different redshifts, we can learn more about the nature of dark energy e.g. through the way it affects the growth of structure. The most promising way to retrieve this information from large-scale structure is to use a statistical approach and consider quantities such as the matter density and velocity as random fields. The present Universe can then be interpreted as one realization of this matter random field whose properties are

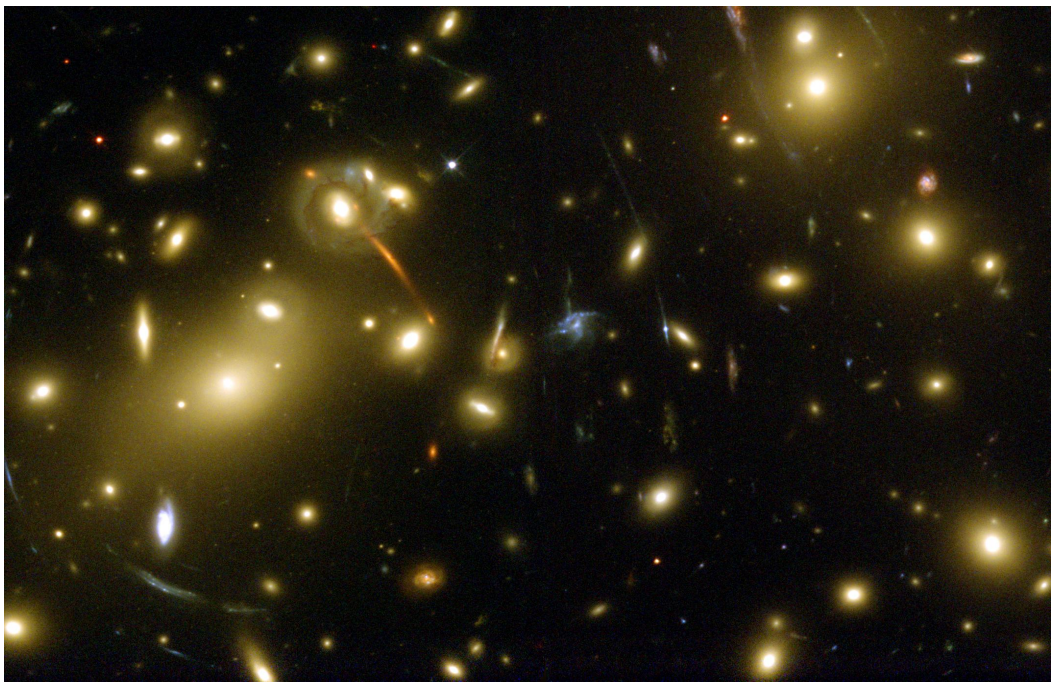

Figure 0.1: *Strong lensing effects in the galaxy cluster Abell 2218 as seen by the Hubble Space Telescope (HST). (Source: NASA, A. Fruchter and the ERO team)*

characterized by its moments. The most important moment in cosmology is the second order moment or - transformed to Fourier space - the power spectrum, as it contains all the information for Gaussian fields and matter is assumed to be initially Gaussian distributed. Since the processes that lead to the formation of structure are non-linear, they inevitably produce non-Gaussianities in the matter density field. In order to use the encoded information in the non-Gaussianities to constrain cosmological parameters, we need to measure higher-order moments and have a good theoretical understanding of the underlying mechanisms that lead to them.

Cosmic shear

An important tool to probe the large-scale structure of the Universe and to estimate cosmological parameters is provided by weak gravitational lensing. It describes the coherent distortion of light coming from distant galaxies caused by matter inhomogeneities on very large cosmological scales. Since gravitational light deflection occurs independent of the specific type of matter, weak lensing provides a unique tool to study dark matter and the distribution of large-scale structure in the Universe. For cosmological purposes, the most important effect of gravitational light deflection is that it causes distortions in the original shape of distant galaxies. These are typically much smaller than the intrinsic ellipticity of a galaxy and cannot be determined from a single galaxy image. Visible effects such as arcs (see Fig. 0.1) occur only in the

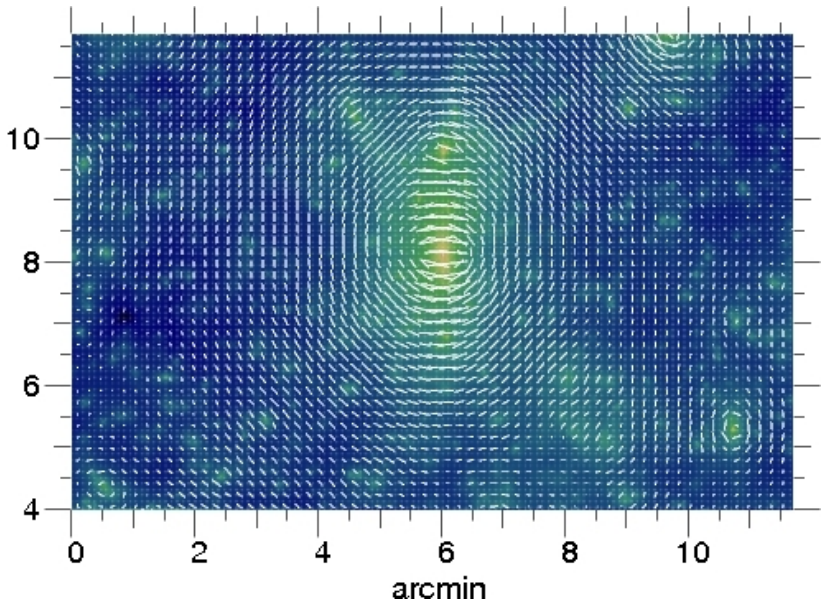

Figure 0.2: *Example of a projected mass distribution and the cosmic shear field induced as obtained from ray-tracing through the Millennium Run simulation (Source: S. Hilbert, J. Hartlap). The white sticks are a measure of the magnitude and direction of the local shear. Note the alignment of the shear to the foreground mass overdensities (green areas).*

strong lensing regime, when very massive objects as e.g. galaxy clusters are involved. To quantify the weak lensing effect nevertheless, cosmologists use high-quality images of a large number of distant galaxies and average over their shapes. The observational signal of interest has become known as *cosmic shear* and was first detected in 2000 by four independent groups (Bacon et al. [3], Kaiser et al. [42], Van Waerbeke et al. [85], Wittman et al. [89]). Since the intrinsic galaxy ellipticities are expected to vanish on average, the result is a direct measure of the projected mass density distribution in the observed patch of sky. This allows us to find constraints on cosmological parameters that are independent of and complementary to those found by other measurements such as CMB anisotropies, SN type Ia or galaxy surveys. Due to its sensitivity on the matter distribution on large scales it provides valuable constraints on the amount of matter in the Universe and the dark matter power spectrum normalization. If one determines the cosmic shear for source galaxies at different redshifts, it can be used to probe dark energy as well. This is due to two factors: on the one hand dark energy affects the cosmic shear geometrically in the sense that the amount of dark energy determines the distance to the observed galaxies, on the other hand it influences the growth of structure which can be seen in the evolution of the matter power spectrum.

Higher-order moments

So far, cosmic shear has been used to measure the convergence power spectrum and parameter constraints have been based on the Gaussian approximation of the power

spectrum covariance. The next generation of surveys will allow much more precise measurements of weak lensing effects, which will enable cosmologists to determine higher-order moments of the matter density field. For this reason, it is important to have a good understanding of the underlying physics and the expected errors. In this thesis, we focus on the fourth-order correlation function and its Fourier counterpart, the trispectrum, since it allows us to study the non-Gaussianities of the matter density field. Furthermore, it enables us to study the expected error in the power spectrum, since we can use the trispectrum to calculate the full non-Gaussian covariance of the power spectrum. In order to find an analytical expression for the trispectrum, we apply the semi-analytic *halo model* [17, 74]. It assumes that all dark matter in the Universe is bound in spherical halos and makes use of results from numerical N-body simulations to characterize halo properties as their profile, abundance and clustering behavior. The results we find from investigating the full non-Gaussian covariance of the projected power spectrum within the halo model approach are finally compared to the results found with numerical N-body simulations.

Overview

The calculation of the non-Gaussian covariance requires detailed background knowledge of structure formation in a ΛCDM Universe. The topics which are most important for this thesis are reviewed in the first four chapters. New results can be found in Chapters 5 and 6. The thesis is organized as follows:

- Chapter 1 gives an overview of the standard Hot Big-Bang model and its shortcomings and summarizes the most important equations that are necessary for describing structure formation in a ΛCDM Universe.
- Chapter 2 outlines how the equations of motion in structure formation can be solved analytically using linear and weakly non-linear perturbation theory.
- Chapter 3 introduces the formalism of random fields and considers the properties of homogeneous, isotropic and Gaussian random fields. Additionally, we consider how one obtains an estimate of a statistical quantity from a sample of measurements and the accuracy one can expect.
- Chapter 4 provides a detailed overview of the halo model description of matter in the Universe which allows one to calculate moments of arbitrary order.
- In Chapter 5, we calculate the full non-Gaussian power spectrum covariance in the three-dimensional and projected cases using the halo model approach. Subsequently, we analyze different approximations for the covariances in order to minimize the computational effort and investigate the impact of scatter in the halo concentration-mass relation.
- Chapter 6 compares the results found in the previous chapter with different numerical simulations. At first for the convergence power spectrum, then for the

corresponding covariance and finally for the non-Gaussian-to-Gaussian ratio of the power spectrum covariance.

The thesis concludes with a short summary and an outlook.

Chapter 1

Standard cosmology

From observations we know that our Universe is homogeneous and isotropic on scales larger than $\simeq 200\,h^{-1}\,\mathrm{Mpc}$, i.e. matter and radiation are uniformly distributed and without any privileged direction. This is often referred to in the literature as the *Cosmological Principle*. Together with Einstein's discovery of general relativity this allows us to come up with a consistent, testable theory of our Universe. Combining our theoretical knowledge with astronomical observations, the picture of an expanding Universe emerges, which was once much denser and hotter. This idea of a Universe which evolved from an initial singularity is called the *Hot Big-Bang model* and relies on three fundamental observations: the *recession of galaxies* increasing with distance from which we can infer the expansion of the Universe, the *light element abundance* indicating that this has been the case since an early era in which the Universe was hot and dense and the *cosmic microwave background* which is observed at present times and interpreted as the relic radiation originating from the decoupling of the photon-electron plasma at early times.

Recent observations confirm the Hot Big-Bang model, but hint that a more detailed description of the Universe beyond the standard model is needed. Measurements of rotation curves from galaxies which probe their gravitational field show a mismatch between observed and predicted mass. The only explanation seems to be that a large fraction of matter in the Universe is *dark*, i.e. nonbaryonic and only weakly interacting. Furthermore, investigations of cluster abundances indicate that roughly one third of the Universe consists of matter at all. In addition, we learn from cosmic microwave background observations that our Universe is close to a flat geometry. In order to explain both results we need an additional contribution which comes up for the missing $\sim 2/3$ of the energy content of the Universe and does not cluster. Even more surprising were the results from observations of distant supernovae. Against the prevailing opinion at that time the data clearly favor a Universe which undergoes a phase of *recent acceleration*. Combining the results with the data obtained from cosmic microwave background and cluster measurements, one sees the necessity of a form of *dark energy* which repulses the gravitational attraction.

Sect. 1.1 gives an overview of the standard hot Big-Bang model and the equations necessary for studying structure formation in an accelerating ΛCDM Universe, whereas

Sect. 1.2 deals with the most important issues beyond the standard model. For a comprehensive overview of these topics see e.g. the book by Kolb and Turner [46], Peacock [61] or Dodelson [18].

1.1 The homogeneous and isotropic Universe

1.1.1 Einstein's field equations

Of the four known fundamental forces in nature – gravity, electromagnetic force, strong and weak interaction – only the former two are long range forces and can act on cosmic scales. As we assume our Universe to be almost charge-neutral, gravity is the dominant force which governs its dynamics and evolution. The fundamental theory describing gravity is the *General Theory of Relativity* as formulated by Einstein in 1916. It models gravity as a property of space-time which can be described as a four-dimensional Riemannian manifold. As a consequence, the corresponding metric $g_{\mu\nu}$ includes time dependencies as well and the line-element of four space-time dimensions has the form

$$\mathrm{d}s^2 = g_{\mu\nu}\,\mathrm{d}x^\mu \mathrm{d}x^\nu\,, \qquad \mu,\nu = 0,\ldots,3 \tag{1.1}$$

where – following Einstein's sum convention – we have to sum over multiple indices. Note that the 0 index is always reserved for the time-like coordinate, while the other three are applied for the spatial coordinates. The advantage of a metric including gravity is that particles in a gravitational field can be considered as moving freely on the geodesics of a curved space-time, whereas in Newtonian physics one has to include gravity as an external force which then alters the particles trajectory. Before considering the metric of an expanding Universe in more detail in the next section, we first study the actual relation between space-time geometry and matter. It is described by the famous *Einstein's field equations*

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}\mathcal{R} = \frac{8\pi G_\mathrm{N}}{c^4}\,T_{\mu\nu} - \Lambda g_{\mu\nu}\,, \tag{1.2}$$

where the *Ricci tensor* $R_{\mu\nu} = R^\alpha_{\mu\nu\alpha}$ and the Ricci scalar $\mathcal{R} = R^\alpha_\alpha$ are contractions of the Riemann tensor $R^\mu_{\nu\lambda\rho}$ which describes the curvature of the manifold, G_N is Newton's constant and $T_{\mu\nu}$ denotes the *energy-momentum tensor* which describes the matter content of the Universe.

In order to allow for static solutions of the field equations, Einstein introduced a *cosmological constant* term $\propto \Lambda g_{\mu\nu}$ in Eq. (1.2) to counterbalance the gravitational attraction of matter. The physical interpretation of this term is still unclear, but a

non-zero cosmological constant in Einstein's field equations provides up to now the simplest explanation for a recently accelerating Universe (see Sect. 1.2.6).

1.1.2 The Robertson-Walker metric

Einstein's field equations can only be solved by making further assumptions on the energy content of the Universe and its metric. In case of a homogeneous and isotropic Universe one has a situation that explicitly allows to determine the time-dependence of Eq. (1.2). The most general metric for an expanding, homogeneous and isotropic Universe is the *Robertson-Walker* (RW) metric with the line element

$$\mathrm{d}s^2 = g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu = c^2\mathrm{d}t^2 - a^2(t)\,\left[\mathrm{d}w^2 + f_K^2(w)\left(\mathrm{d}\vartheta^2 + \sin^2\vartheta\,\mathrm{d}\varphi^2\right)\right] \tag{1.3}$$

where (w, ϑ, φ) are denoted as *comoving* coordinates, $a(t)$ as the *scale factor* (normalized to $a(t_0) = 1$ today) and $f_K(w)$ is the *comoving angular diameter distance*. The scale factor is a relative length which varies according to the expansion or contraction of the Universe, whereas the comoving angular diameter distance can take – depending on the underlying geometry of the Universe – the following forms:

$$f_K(w) = \begin{cases} K^{-1/2}\sin(K^{1/2}w) & \text{for} \quad K > 0\,, \\ w & \text{for} \quad K = 0\,, \\ (-K)^{-1/2}\sinh((-K)^{1/2}w) & \text{for} \quad K < 0\,. \end{cases} \tag{1.4}$$

The parameter K determines the *curvature* of the three-dimensional surface defined by the spatial part of the Robertson-Walker metric: for $K > 0$ it corresponds to a 3-sphere, for $K = 0$ one has a flat Euclidian space and $K < 0$ yields a hyperbolic, open space.

1.1.3 Light rays and cosmic redshift

An important consequence of an expanding Universe is the shift in wavelength experienced by propagating photons. To quantify this effect, we consider a photon of wavelength λ_1 that is emitted from a source at a time t_1 and arrives at a telescope at a time t_0 with a wavelength λ_0. Without loss of generality, we assume the photon to travel along a radial trajectory which satisfies $\mathrm{d}\vartheta = \mathrm{d}\varphi = 0$. Since the photon is massless, it propagates along a null geodesics, which fulfills $\mathrm{d}s = 0$. From the RW metric (see Eq. 1.3) it follows that in this case

$$c\,\mathrm{d}t = -a(t)\mathrm{d}w\,, \tag{1.5}$$

where the minus sign appears since we consider the backwards light cone of the observer. From this we can calculate the radial distance at which the photon is observed today by integrating Eq. (1.5) and obtain

$$w = \int_{t_1}^{t_0} \frac{c\,\mathrm{d}t}{a(t)} = \text{const.} \tag{1.6}$$

Differentiating with respect to t_1 yields

$$\frac{\mathrm{d}t_0}{\mathrm{d}t_1} = \frac{a(t_0)}{a(t_1)}\,. \tag{1.7}$$

This time-dilation is responsible for a shift in wavelength and frequency of photons and defines a *cosmological redshift*

$$\frac{a(t_0)}{a(t_1)} = \frac{\nu_1}{\nu_0} = \frac{\lambda_0}{\lambda_1} \equiv 1 + z\,. \tag{1.8}$$

Since $a(t_0) = 1$ by definition and $0 < a(t_1) < 1$ in an expanding Universe, this shift corresponds to an increase of the photon wavelength. In the visible spectrum this means a shift towards red wavelength which gives the *redshift* its name.

The recession of galaxies was already discovered by Slipher in 1912 and interpreted as cosmological Doppler effect. A systematic analysis of galaxy velocities in the 1920s by Hubble revealed that the recession of galaxies is proportional to their distance. This observation was later explained with the expansion of the Universe.

1.1.4 The Friedmann equations

Inserting the RW metric (1.3) into Einstein's field equations (1.2), constrains the matter content to the form of a perfect fluid, i.e. with no viscosity or heat flow. In this case the energy-momentum tensor reduces to

$$T^{\mu}{}_{\nu} = \mathrm{diag}(\rho c^2, -p, -p, -p)\,. \tag{1.9}$$

where $p = p(t)$ and $\rho = \rho(t)$ are the time-dependent, homogeneous pressure and energy density. With this choice for the energy-momentum tensor and the RW metric, one recovers from Einstein's field equations two independent equations which have become

known as *Friedmann equations* and describe the evolution of an expanding Universe:

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G_\mathrm{N}}{3}\rho - \frac{Kc^2}{a^2} + \frac{\Lambda}{3}\,, \tag{1.10}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G_\mathrm{N}}{3}\left(\rho + \frac{3p}{c^2}\right) + \frac{\Lambda}{3} \tag{1.11}$$

where the dots denote derivatives with respect to t. These equations indicate that the precise evolution of the scale factor is determined by the content of the Universe, i.e. ρ and p. In general, the energy density ρ is a sum of contributions from different species, e.g. from matter, radiation or exotic particles but does not include the cosmological constant. Nevertheless, one can use the Friedmann equations without the explicit occurrence of Λ and define a contribution from a cosmological constant which enters then in the energy density as

$$\rho_\Lambda \equiv \frac{\Lambda}{8\pi G_\mathrm{N}}\,. \tag{1.12}$$

It is useful to characterize the expansion rate of the Universe by introducing the *Hubble parameter* $H \equiv \dot{a}/a$. Its present value $H_0 = H(t_0)$ is denoted as *Hubble constant* and is often parametrized as

$$H_0 = 100\,h\,\mathrm{km\ s^{-1}\,Mpc^{-1}} \tag{1.13}$$

with h taking into account the observational uncertainty. The combination of distance measurements from Type Ia supernovae (SN) and baryon acoustic oscillations (BAO) with the WMAP data find an estimate of $h = 0.701 \pm 0.013$ (see Komatsu et al. [47]).

The different contributions to the contents of the Universe are commonly given in terms of relative energy densities. For this reason, one defines the *critical density*

$$\rho_\mathrm{crit} \equiv \frac{3H^2}{8\pi G_\mathrm{N}}\,, \tag{1.14}$$

which corresponds to the Friedmann equation (1.10) for a flat Universe (i.e. with $K = 0$). Its value today can be determined via the Hubble constant and is

$$\rho_\mathrm{crit,0} = \frac{3H_0^2}{8\pi G_\mathrm{N}} = 1.879 \times 10^{-29} h^2 \mathrm{g\ cm^{-3}}\,, \tag{1.15}$$

where the subscript 0 indicates present day values. The relative contribution of a single species to the total energy density ρ_crit is customarily given in terms of the ratio

$$\Omega_i \equiv \frac{\rho_i}{\rho_\mathrm{crit}}\,, \tag{1.16}$$

where the sum of all species in the Universe defines the *total energy density parameter* Ω. Introducing this quantity allows to rewrite the Friedmann equation (1.10) as:

$$\frac{Kc^2}{H^2a^2} = \Omega - 1\,. \tag{1.17}$$

In this way, the complete energy content of the Universe is directly related to its underlying geometry. A density parameter $\Omega = 1$ identifies a flat Universe with $K = 0$. Accordingly, $\Omega < 1$ corresponds to an open geometry, while $\Omega > 1$ characterizes a closed geometry. If one now defines

$$\Omega_K \equiv -\frac{Kc^2}{a^2H^2} \tag{1.18}$$

for the curvature, the density parameters must satisfy the consistency relation

$$\Omega_K + \Omega = 1\,. \tag{1.19}$$

1.1.5 Solutions to the Friedmann equations

Assuming a flat Universe the Friedmann equations (1.10) and (1.11) can be solved explicitly. For this it is useful to characterize different species by their equations of state

$$w_i \equiv \frac{p_i}{\rho_i c^2} \tag{1.20}$$

where the index i refers to the considered species and w_i is usually independent of time (see Tab. 1.1). Combining the Friedmann equations (1.10) and (1.11) for a flat Universe with the previous definition (1.20) yields

$$\dot{\rho} = -3H\,(1+w)\,\rho\ . \tag{1.21}$$

To guarantee the uniqueness of the solution, one assumes the Universe to have a single dominant species (i.e. $\rho_i \simeq \rho$). Then equation (1.21) is solved by

$$\rho \propto a^{-3(1+w)}\,. \tag{1.22}$$

Accordingly, we find for the scale factor the following solutions:

$$a \propto t^{\frac{2}{3(1+w)}} \qquad \text{for } w \neq -1\ , \tag{1.23}$$

$$a \propto \mathrm{e}^{Ht} \qquad \text{for } w = -1\ . \tag{1.24}$$

For the most standard species we have summarized the results in Tab. 1.1. From the scale factor dependence we see that the energy density of matter decreases more slowly than the one of radiation as the Universe expands. This indicates that the radiation was dominant in the past and has been overtaken by the matter energy density at a certain redshift which we denote as *matter-radiation equality* z_{eq}. For the same reason

Table 1.1: *Equation of state and scale factor for the most standard components.*

	w	p	ρ	a
Radiation	$\frac{1}{3}$	$\frac{\rho}{3}$	$\propto a^{-4}$	$\propto t^{\frac{1}{2}}$
Matter	0	0	$\propto a^{-3}$	$\propto t^{\frac{2}{3}}$
Vacuum Energy	-1	$-\rho$	$\propto$ const	$\propto \mathrm{e}^{Ht}$

a cosmological constant will finally dominate over matter and radiation as the Universe expands.

Among these three species, only the vacuum energy density can be responsible for an accelerating Universe. In general, combining equations (1.11) and (1.20) yields

$$\frac{\ddot{a}}{a} = -\frac{4\pi G_{\mathrm{N}}}{3}\left(1+3w\right)\rho \tag{1.25}$$

and only if there is a dominant component with an equation of state $w < -\frac{1}{3}$, $\ddot{a}$ becomes positive and the Universe accelerates.

With the derived equations, one can finally rewrite the first Friedmann equation (1.10) in terms of the density parameters:

$$\left(\frac{H}{H_0}\right)^2 = \sum_i \Omega_i \left(\frac{a}{a_0}\right)^{-3(1+w_i)} + \Omega_K \left(\frac{a}{a_0}\right)^{-2} , \tag{1.26}$$

where the sum goes over all species i contributing to the total density parameter at the considered time a and the subscript 0 indicates present day values of the quantities. Combined measurements from BAO, SN and WMAP [47] find the following values for the present density parameters:

- Baryons: $\Omega_{\mathrm{b},0} = 0.0462 \pm 0.0015$,
- Dark matter: $\Omega_{\mathrm{d},0} = 0.233 \pm 0.013$,
- Dark energy: $\Omega_{\Lambda,0} = 0.721 \pm 0.015$,

which are consistent with a flat ΛCDM Universe. A contribution from radiation at the present time is usually neglected, since $\Omega_{\mathrm{r},0} \simeq 5 \times 10^{-5}$. It is common to define a matter density parameter $\Omega_{m,0} \equiv \Omega_{\mathrm{d},0} + \Omega_{\mathrm{b},0}$.

1.1.6 Distance measures

In a curved space-time, where physical lengths change according to the scale factor, distance measures are not unique. A central quantity from which all relevant measures

can be obtained is the *comoving distance* between a source at redshift z_2 and an observer situated at $z_1 < z_2$:

$$w(z_1, z_2) = \int_{z_1}^{z_2} \frac{c\,\mathrm{d}z'}{H(z')}\,. \tag{1.27}$$

It corresponds to the distance of a radial light ray propagating along the null-geodesics, where $\mathrm{d}s = 0$. In a flat ΛCDM Universe the comoving distance takes the form:

$$w(z_1, z_2) = \frac{c}{H_0} \int_{a(z_2)}^{a(z_1)} \mathrm{d}a \left[a\Omega_\mathrm{m} + a^4\Omega_\Lambda\right]^{-1/2}\,. \tag{1.28}$$

Of particular importance is the comoving distance that light could have traveled since the beginning of the Universe at $z = \infty$:

$$d_\mathrm{H}(z) = \int_z^\infty \frac{c\,\mathrm{d}z'}{H(z')}\,. \tag{1.29}$$

Regions with a separation larger than this distance are causally disconnected, which is why d_H is also known as *horizon distance* (see also Sect. 1.2.2). In specific cases, one can express the horizon distance in terms of the scale factor: in a radiation-dominated Universe we obtain $d_\mathrm{H}(a) \propto a$, whereas in a matter-dominated Universe we have $d_\mathrm{H}(a) \propto a^{1/2}$ instead.

In order to construct consistent distance measures for observations, cosmologists generalize the classic measures to an expanding Universe. Usually a distance to an object can be defined from its known physical size l and its apparent angular diameter θ according to $d_A = l/\theta$. In an expanding Universe the object has a comoving size of l/a and the comoving angular distance out to this object is given by $f_K(w(z))$ as defined in Eq. (1.4). The apparent angular diameter is then $\theta = (l/a)/f_K(w(z))$ so that the *angular diameter distance* is given by

$$d_A(z) = a f_K(w(z))\,. \tag{1.30}$$

This result assumes the observer to be situated at $z = 0$. We can generalize it further to a situation of an object at redshift z_2 which is seen by an observer at $z_1 < z_2$. The angular distance is then

$$d_A(z_1, z_2) = a(z_2) f_K[w(z_1) - w(z_2)]\,, \tag{1.31}$$

where in general $d_A(z_1, z_2) \neq d_A(z_2, z_1)$.

Another way of inferring distances in cosmology is by measuring the flux F of an object with known luminosity L. If we were in a Euclidian Universe, the distance would be given by

$$d_L = \sqrt{\frac{L}{4\pi F}}\,, \tag{1.32}$$

since the total luminosity through a spherical shell with area $4\pi r^2$ is constant. Again, we need to consider what happens to the measured quantities in an expanding Universe. The luminosity of the source at a time a decreases in an expanding Universe to La^2 due to two effects: the photons loose energy on their way to the observer and the photons arrive less frequently at the telescope. With the comoving radial distance to that source being $f_K(w(z))$ the flux changes to

$$F = \frac{La^2}{4\pi f_K(w(z))}\,. \tag{1.33}$$

Comparing this with Eq. (1.32), the *luminosity distance* is defined as

$$d_L(z) = \frac{f_K(w(z))}{a} = (1+z) f_K(w(z))\,. \tag{1.34}$$

Between the angular diameter and the luminosity distance exists the following relation

$$d_L(z) = (1+z)^2 d_A = (1+z) f_K(w)\,. \tag{1.35}$$

1.1.7 Big-Bang nucleosynthesis

An essential part of the standard model is the Big-Bang nucleosynthesis (BBN) the theory which predicts the light element abundance in the Universe. As the production of nuclei requires high energies, nucleosynthesis began in the early radiation-dominated era of the Universe. More specifically, the essential processes to form light elements started at temperatures $T \lesssim 1$ MeV which corresponds to a time $t \gtrsim 1$s after the Big-Bang. At these temperatures neutrinos decouple from the rest of the Universe and weak interactions occur only slowly with respect to the expansion of the Universe. As weak interactions are responsible for converting nucleons into each other, the ratio of neutron to proton number density stays approximately constant at a value

$$\frac{n_\mathrm{n}}{n_\mathrm{p}} = \mathrm{e}^{-Q/T_\mathrm{f}} \simeq \frac{1}{6}\,, \tag{1.36}$$

where $Q = 1.293$ MeV is the neutron-proton mass difference which is responsible for the larger abundance of protons and $T_\mathrm{f} \simeq 1$ MeV is the neutrino freeze-out temperature. As neutrons have a finite lifetime they gradually decay into protons and leptons. When the temperature reaches a value around $T \simeq 100$ keV, the number of photons per nucleon is finally small enough that nuclei can form without immediately being photo-dissociated. At this time the neutron to proton ratio is approximately 1/7. Since the most stable light element is ^{4}He almost all free neutrons at this time are converted into it. Thus a good estimate of the primordial ^{4}He-abundance is

$$Y_\mathrm{p} \equiv \frac{2n}{n+p} = \frac{2(n_\mathrm{n}/n_\mathrm{p})}{1+n_\mathrm{n}/n_\mathrm{p}}\,, \tag{1.37}$$

where n and p are the number of available neutrons and protons at this time. Additionally, small amounts of other light elements are produced. Specified in number per protons, nuclear reactions generate D and ^{3}He ($\sim 10^{-5}$) and small traces of ^{7}Li ($\sim 10^{-7}$). Heavier nuclei do not form during Big-Bang nucleosynthesis due to the lack of stable nuclei with mass numbers 5 or 8.

All rates of the nuclear processes leading to the aforementioned elements, depend essentially on one parameter, namely the baryon-to-photon ratio

$$\eta \equiv \frac{n_\mathrm{B}}{n_\gamma}\,, \tag{1.38}$$

where n_B corresponds to the baryon number density and n_γ denotes photon number density. To be consistent with the primordial abundances of D and ^{3}He the baryon-to-photon ratio has to be in the range $2.6 \times 10^{-10} < \eta < 6.2 \times 10^{-10}$ [90].

1.1.8 Cosmic microwave background

At a temperature around $T \simeq 3000$ K, corresponding to a time 380 000 years after the Big-Bang, the ionized photon-electron plasma decouples. As a consequence, electrons and protons combine to form hydrogen and the Universe becomes neutral with photons propagating freely. Following the cosmic expansion, these photons are redshifted and their temperature drops with $T \propto a^{-1}$. Since before the decoupling matter was in approximate thermal equilibrium, the photons formed at early times a blackbody radiation spectrum. As its form is not altered with the expansion we can still observe a perfect black body radiation today at a redshifted temperature around $T \simeq 2.7$ K. This relic radiation better known as *cosmic microwave background* (CMB) was discovered by Penzias and Wilson in 1965. Its existence is one of the strongest proofs supporting the Hot Big-Bang model.

As it turned out the CMB is almost perfectly isotropic. Its temperature deviates from isotropy at the level of one part in 10^5. These tiny fluctuations are a conserved imprint

of the initial conditions at that time and serve as seeds for structure formation in the Universe. They originate from the time of decoupling when photons were released from different regions in space with slightly different gravitational potentials. Since photons redshift when they climb out of these potentials, temperature anisotropies emerged which reflect the initial distribution of density perturbations. Additionally, these perturbations gave rise to *acoustic waves* in the primordial electron-photon plasma which are still imprinted on the CMB today. The largest possible wavelength of these oscillations is given by the *sound horizon* which has a physical length of $d_{\rm s}(z_{\rm dec})$. It provides us with a ruler on the sky and the corresponding angular scale depends then on the underlying geometry of the Universe. For a flat Universe, we expect the peak of this acoustic wave at a wave-number $l \simeq 220$ which is in perfect agreement with observations.

1.2 Beyond the standard model

1.2.1 Flatness problem

In Sect. 1.1.4 we showed that the Friedmann equation can be rewritten in terms of the density parameter as

$$\Omega - 1 = \frac{Kc^2}{a^2 H^2}\,. \tag{1.39}$$

If the Universe is flat, i.e. $K = 0$, we can deduce from Eq. (1.39) that this will remain for all times. Otherwise the density parameter evolves according to the dominant species in the Universe. In a matter-dominated Universe the density parameter changes with $|\Omega - 1| \propto t^{2/3}$, whereas in a radiation-dominated Universe we have $|\Omega - 1| \propto t$. Thus, during most of the cosmic evolution, it is a function increasing with time. Since at the present time the Universe is close to a flat geometry this implies that the density parameter must have been extremely close to 1 at early times. As almost all initial conditions lead either to a closed or an open geometry, these fine-tuned initial conditions seem extremely unlikely. To understand the severeness of this *flatness problem*, one can consider for example the Universe around the time of nucleosynthesis where $t_{\rm nuc} = 1$s. Eq. (1.39) requires then that

$$|\Omega(t_{\rm nuc}) - 1| \lesssim 10^{-16}\,. \tag{1.40}$$

At earlier times, the total density parameter must be even closer to 1.

1.2.2 Horizon problem

The distance how far photons could have traveled during the lifetime of the Universe is given by the *horizon distance* as defined in Eq. (1.29). In terms of the Hubble scale factor, it can be rewritten as

$$d_{\mathrm{H}}(a) = \int_0^a \frac{c\,\mathrm{d}a'}{a'^2 H(a')}\,. \tag{1.41}$$

Depending on how $a^2 H(a)$ evolves this distance can be infinite or take a finite value. In order to find out about the horizon distance today, let us simplify the consideration by assuming a matter-dominated Universe with $K = \Lambda = 0$. As this is true during most of the time our Universe exists, we can find in this way a good estimate for the real horizon distance. In this case, Eq. (1.41) becomes

$$d_{\mathrm{H}}(a) = \int_0^a \frac{c\,\mathrm{d}a'}{\sqrt{a'\Omega_{\mathrm{m}}}H_0} = \frac{2c\sqrt{a}}{\sqrt{\Omega_{\mathrm{m}}}H_0}\,, \tag{1.42}$$

where we made use of $H = H_0 a^{-2/3}\sqrt{\Omega_{\mathrm{m}}}$ (use Eq. (1.26) and set $K = 0$). The distance found is equivalent to the physical size of the horizon distance since $a_0 = 1$ today. As the result is finite, light signals can only propagate a finite distance between the Big Bang and the present and thus information can only be passed along this distance. If we consider the explicit example of the horizon size at the formation time of the cosmic microwave background at $z_{\mathrm{dec}} \approx 1100$, we find

$$\theta_{\mathrm{dec}} = \frac{d_{\mathrm{H}}(z_{\mathrm{dec}})}{d_A} = \sqrt{\frac{\Omega_{\mathrm{m}}}{z_{\mathrm{dec}}}} \approx \sqrt{\Omega_{\mathrm{m}}}\,2^\circ\,, \tag{1.43}$$

where θ_{dec} is the apparent angular diameter at decoupling. Thus regions which are separated by more than $\sim 2^\circ$ are causally disconnected. The question arises then why we see on the sky almost the same temperature. Due to the consideration above regions on the opposite side of the sky never had the chance to get into causal contact. This is known as the *horizon problem*.

1.2.3 Relic particle abundances

Another problem with the Hot Big Bang model arises from modern particle physics. Grand Unified Theories (GUT) which aim in unifying the fundamental forces, predict a high abundance of magnetic monopoles in the early Universe. As they are produced at very high energies, they are predicted to be very massive (around 10^{16} GeV). Such particles would be non-relativistic during most of the evolution of the Universe and thus have plenty of time to dominate over radiation. Since so far not a single magnetic

monopole has been observed, theories predicting them are in contradiction with the standard model. Recent particle physics models propose the existence of other such relic particles as e.g. gravitinos or moduli fields and have to deal with the same abundance problem.

1.2.4 Inflation

Especially the horizon problem provides a severe problem of the standard model as it concerns causality. A solution to most of the problems addressed here was proposed by Alan Guth in 1981 and became famous under the name of *inflation*. The basic idea of inflationary models is to assume that the early Universe had undergone a phase of accelerated expansion where $\ddot{a} > 0$. The effect of this acceleration is that the size of the Universe is hugely increased and its geometry is flattened leading to $K \simeq 0$. Additionally, the horizon size is extremely increased such that a much larger region of the Universe was in causal contact and the unwanted relics as e.g. magnetic monopoles are extremely diluted. As a bonus, these models also predict a scale-free spectrum of density perturbations which are required to form structure in the Universe.

In order to achieve a phase of accelerated expansion, Eq. (1.25) demands a dominant species of negative pressure which fulfills

$$p < -\frac{\rho c^2}{3} \Leftrightarrow w < -\frac{1}{3}\,. \tag{1.44}$$

The easiest way to construct such a model is by considering a Universe in which a cosmological constant is dominant at early times. It is characterized by a pressure $p = -\rho c^2$ and leads to a scale factor increasing exponentially in time

$$a(t) = \exp\left(\sqrt{\Lambda/3}t\right)\,. \tag{1.45}$$

After a certain amount of time, the inflationary expansion must come to an end and the energy of the cosmological constant has to be converted into radiation and matter. The usual picture here is that the particles which act as a cosmological constant decay into ordinary particles. This phase is referred to as *reheating*. Thereafter, the Universe can evolve according to the standard Big Bang model.

Up to now, a lot of different models of inflation have emerged. The most prominent class of models describes matter in the early Universe by one or more real scalar fields ϕ_i which are characterized by their potentials $V_i(\phi_i)$. The particles corresponding to such field are called *inflatons* and have not been detected yet.

If we specialize to a flat Universe with one dominant, homogeneous scalar field ϕ, the scalar field behaves as a perfect fluid with

$$\rho_\phi = \frac{1}{2}\dot{\phi}^2 + V(\phi)\,, \tag{1.46}$$

$$p_\phi = \frac{1}{2}\dot{\phi}^2 - V(\phi)\,. \tag{1.47}$$

The equation of motion for the scalar field is given by

$$\ddot{\phi} + 3H\dot{\phi} + \frac{\mathrm{d}V}{\mathrm{d}\phi} = 0\,, \tag{1.48}$$

which corresponds to an oscillator equation for a scalar field experiencing a friction due to the Hubble expansion of the Universe ($\dot{\phi}$-term). If the scalar field is dominant, the Friedmann equation is

$$H^2 = \frac{8\pi G_\mathrm{N}}{3}\left[\frac{1}{2}\dot{\phi} + V(\phi)\right]\,. \tag{1.49}$$

The accelerated phase in these models occurs if the potential energy of the scalar field becomes much larger than its kinetic energy, i.e. $\dot{\phi} \ll V(\phi)$. In this case the density and the pressure are dominated by the potential energy and the equation of state becomes $p_\phi \simeq -\rho_\phi$. As this behavior corresponds approximately to that of a cosmological constant, the expansion is accelerated. This scenario can be thought of as a scalar field slowly rolling down its potential. A more detailed description of the mechanism of inflation and specific examples are considered e.g. in Liddle and Lyth [48].

1.2.5 Dark matter

First hints for the existence of a *dark matter* component in the Universe were already found by Zwicky in 1933. He discovered that the orbital velocities of galaxies in the Coma Cluster were on average close to $1000\,\mathrm{km\,s^{-1}}$. This requires a cluster mass much larger than that from all the stars and gas contained in the Coma Cluster. A similar discrepancy between luminous and predicted mass was revealed in the rotation curves of spiral galaxies by Rubin and Ford in 1970. If the mass of the galaxy were to follow the observed light, one would expect a $\propto 1/\sqrt{r}$ Kepler decline in the rotation curves. Instead Rubin and Ford found the velocity to be constant up to the largest observable radii. The only explanation is the existence of a non-luminous and non-baryonic form of dark matter which forms a dark halo around galaxies and alters the rotation curves. Independent observations of other luminous objects as stars or globular clusters find the same discrepancy between luminous and predicted mass.

Although their has been no direct detection of dark matter up to now, particle physicists have several possible candidates for dark matter. In order to qualify as dark

matter, these candidates have to satisfy several conditions: they must interact very weakly with electromagnetic radiation in order to be non-luminous, they have to be stable on cosmic time-scales and they need the right relic density to explain the present dark matter abundance in the Universe. Possible candidates which fulfill these constraints are primordial black holes, axions and weakly interacting massive particles (WIMPs).

1.2.6 Cosmological constant problem(s)

Since Hubble, cosmologists had been trying to measure the slow-down of the Universe's expansion due to gravitational attraction. In 1998, two teams studying distant type Ia supernovae discovered independently that instead the opposite is true: the Universe is currently undergoing a phase of accelerated expansion. In order to explain this, we need a form of energy that 'counteracts gravity'. General relativity allows for the presence of such an energy with a sufficiently negative pressure. The simplest form for this energy is provided by a cosmological constant term that can be added to Einstein's equations. In this case it is important to notice that the effective energy of the ground state really is of importance. This is due to the fact that gravity couples directly to the vacuum energy. In most other physical contexts, we are mainly interested in the potential differences as e.g. in case of a falling object which loses potential energy.

A cosmological constant model is consistent with observations of the CMB and the clusters. Thus, it would seem natural to accept the presence of a cosmological constant. On the other hand, a contribution from the vacuum to the content of the Universe has to be explained judiciously from the theoretical point of view. There is even more a need for a good explanation since the measured value of $\Omega_{\Lambda,0}$ and the Hubble constant H_0 fix a possible vacuum energy contribution to

$$\rho_{\text{vac}} \approx 10^{-120} M_P^4 = (3 \times 10^{-3}\,\text{eV})^4 \ . \tag{1.50}$$

While, at the classical level, one can arbitrarily choose the value of the vacuum energy, one must remember that this value will receive new contributions considering quantum effects. The question is why all these contributions – which are of the order of typical fundamental scales in particle physics – should add up to this small value. If for example supersymmetry (SUSY) is broken at the TeV scale, one would expect $\rho_{\text{vac}} \sim (1\text{TeV})^4$.

Because of this big gap between the scales, a fine-tuning of the parameters is required to explain the non-zero vacuum expectation value (VEV). This is known as the *cosmological constant problem* (see e.g. Weinberg [87], Witten [88], Carroll et al. [14]).

A second problem associated with a cosmological constant is the so-called *coincidence problem*. The question is why matter and vacuum energy density are of the same order today. While the vacuum energy stays constant the other contributions decrease

rapidly during the whole evolution of the Universe. Therefore, their initial values have to be extraordinarily fine-tuned in order to have the energy densities at the same order today. Moreover, the fine-tuning required increases if the initial redshift of the Universe increases. This problem is also known as the *fine-tuning problem of the initial values*.

1.2.7 Dark energy

The simplest approach to 'solve' the fine-tuning problem of the initial values is to impose the vacuum energy to be exactly zero. Usually this is done by using a symmetry argument. One example would be unbroken supersymmetry (SUSY). However, since SUSY is broken, it can of course not help us in this specific case. We can be less demanding, simply assuming $\rho_{\text{vac}} = 0$ and then explaining the observed acceleration by some other form of dark energy. One possibility is to consider the dynamics of a sufficiently slowly rolling scalar field which mimics the behavior of a cosmological constant at the present time. This type of model is often referred to as *quintessence* and explains the present accelerating expansion of the Universe with the same mechanism as inflation . Such fields are motivated theoretically in models where supersymmetry is dynamically broken by gaugino condensation. Since the VEV of the condensate depends on the value of the dilaton field, it could mimic a quintessential potential.

Chapter 2

Perturbation theory

The cosmological principle holds only on scales larger than approximately $200\,h^{-1}$ Mpc. Below this scale, astronomers find the Universe to be inhomogeneous which shows in the richness of structure such as galaxies, galaxy clusters or filaments. According to the current paradigm of structure formation, the objects we observe originate from the gravitational collapse of small perturbations in a homogeneous, expanding Universe. The perturbations themselves were generated from quantum fluctuations of the inflaton field and are in the simplest inflationary models predicted to be adiabatic and Gaussian. Mainly due to gravitational instabilities, they started to grow in amplitude and form structure. The earliest traces of the primordial inhomogeneities are visible in the temperature anisotropies of the CMB in consistency with the prevailing theory.

In order to understand structure formation in more detail, we need a theory that describes the evolution of the underlying physical fields as e.g. the mass density $\rho(\boldsymbol{x}, t)$, the velocity $\boldsymbol{v}(\boldsymbol{x}, t)$ or the gravitational potential field $\phi(\boldsymbol{x}, t)$, which is valid for the matter density in the Universe. In this chapter, we focus mainly on pressureless, dark matter perturbations since they are the most important ingredient for forming bound objects. The effect of baryonic and radiation perturbations is only considered qualitatively. Since the resulting set of equations cannot be solved in general, we additionally restrict the consideration to perturbations on scales where the Newtonian theory of gravity can still be applied. After summarizing the equations of motions which govern the evolution of dark matter in Sect. 2.1, we solve them in the linear approximation in Sect. 2.2. This allows us to introduce the *growth factor* in Sect. 2.3 which describes the evolution of dark matter perturbations in terms of the scale factor. Using a perturbative approach for the density field, Sect. 2.4 deals with the non-linear solutions of the equations of motions and expands their validity to general ΛCDM cosmologies in Sect. 2.5. After a qualitative discussion of radiation perturbations and the suppression of perturbation growth due to pressure in Sect. 2.6, we summarize in Sect. 2.7 the effects on different scales in the *transfer function*. The results found can then be used to describe the power spectrum on large scales which we consider in the next chapter. A more detailed presentation of perturbation theory can be found in Peebles [65] or in the review by Bernardeau et al. [6].

2.1 The fluid equations

In the following, we study the evolution of the dark matter density $\rho \equiv \rho_{\mathrm{d}}(\boldsymbol{x}, t)$ in an expanding Universe. Furthermore, we restrict our consideration to the matter-dominated era such that we can neglect pressure. Pressure effects are only important for dark matter perturbations in the very early Universe where radiation dominates, and at the final stage of object formation when they counteract gravitational forces. Additionally, we consider scales, well below the Hubble radius $d_{\mathrm{H}} = cH^{-1}$. This allows to apply a non-relativistic Newtonian approach [65]. With these assumptions the dark matter content of the Universe can be treated as an *ideal fluid* where particles interact without friction and can be described completely by the stress energy-tensor $T_{\mu\nu}$ (see also Sect. 1.1.4). A more general derivation of the equations governing the evolution of dark matter can be found by solving the *collisionless Boltzmann equation* for the single stream approximation, which applies on large scales and for early times [18]. Nevertheless, one obtains in either of the considerations three coupled equations, which describe the evolution a pressureless fluid:

- Continuity equation: $$\dot{\rho} + 3H\rho + \frac{1}{a}\nabla \cdot (\rho \boldsymbol{v}) = 0\,, \tag{2.1}$$
- Euler equation: $$\dot{\boldsymbol{v}} + H\boldsymbol{v} + \frac{1}{a}(\boldsymbol{v}\nabla)\cdot\boldsymbol{v} = -\frac{1}{a}\nabla\phi\,, \tag{2.2}$$
- Poisson equation: $$4\pi G_N \rho a^2 + 3a\ddot{a} - \Lambda a^2 = \Delta\phi\,, \tag{2.3}$$

where the dot denotes a time-derivative, ∇ is the gradient and Δ the Laplace operator with respect to comoving coordinates. These *evolution equations* are highly non-linear and describe the evolution of the dark matter density $\rho = \rho(\boldsymbol{x}, t)$, the gravitational potential $\phi = \phi(\boldsymbol{x}, t)$ and the peculiar velocity field $\boldsymbol{v} = \boldsymbol{v}(\boldsymbol{x}, t)$ in an expanding Universe in terms of comoving coordinates. The peculiar velocity field is related to the proper velocity $\boldsymbol{u} = \boldsymbol{u}(\boldsymbol{r}, t)$ in physical (Eulerian) coordinates $\boldsymbol{r}$ by

$$\boldsymbol{v} = \boldsymbol{u} - H\boldsymbol{x}\,, \tag{2.4}$$

where the second term corresponds to the *Hubble flow*. Although the evolution equations (2.1)-(2.3) are approximate due to the assumptions made, they apply for the most important situations in an expanding Universe.

For a perturbative consideration it is necessary to rewrite the evolution equations in terms of the *density contrast* $\delta(\boldsymbol{x}, t)$. It is defined as the relative deviation of the local dark matter density field $\rho(\boldsymbol{x}, t)$ from the average density of the Universe $\bar{\rho}(t)$:

$$\delta(\boldsymbol{x}, t) = \frac{\rho(\boldsymbol{x}, t) - \bar{\rho}(t)}{\bar{\rho}(t)}\,, \tag{2.5}$$

where we omit in the following the dependencies in $\boldsymbol{x}$ and t for notational convenience. The mean density contrast is by definition 0 and we can expand all equations around the unperturbed, homogeneous Universe. In terms of the density contrast, the continuity equation becomes

$$\dot{\delta} + \frac{1}{a}\nabla\left[(1+\delta)\boldsymbol{v}\right] = 0\,, \tag{2.6}$$

and the Poisson equation can be rewritten using the second Friedmann equation (1.11) as

$$\Delta\phi = \frac{3H_0^2\Omega_\mathrm{m}}{2a}\,\delta\,. \tag{2.7}$$

Note that with the transformation to the density contrast the explicit dependence on a cosmological constant vanishes. Even so, a cosmological constant influences the growth of density perturbations since it enters the evolution equations via the Hubble parameter.

2.2 Linear solution

The set of evolution equations (2.2), (2.6) and (2.7) cannot be solved analytically. In order to solve them we have to restrict our consideration to small perturbations of the homogeneous and isotropic background cosmology, i.e. we assume $|\delta| \ll 1$. In this way, it is possible to linearize the Euler equation (2.2) and the continuity equation (2.6) which yields

$$\dot{\delta} + \frac{1}{a}\nabla\cdot\boldsymbol{v} \simeq 0\,, \tag{2.8}$$

$$\dot{\boldsymbol{v}} + H\boldsymbol{v} + \frac{1}{a}\nabla\phi \simeq 0\,. \tag{2.9}$$

Combining these two equations with the already linear Poisson equation (2.7), one obtains a second-order linear differential equation for the linear density contrast

$$\ddot{\delta} + 2H\dot{\delta} - \frac{3H_0^2\Omega_\mathrm{m}}{2a^3}\,\delta = 0\,. \tag{2.10}$$

Note that in this equation only derivatives with respect to time appear. Hence, the solutions can be factorized into a spatial and a time-dependent part. A general solution to Eq. (2.10) can then be constructed from two linearly independent solutions such that

$$\delta(\boldsymbol{x}, t) = D_+(t)\Delta_+(\boldsymbol{x}) + D_-(t)\Delta_-(\boldsymbol{x})\,. \tag{2.11}$$

To understand the physical meaning of these solutions, it is useful to solve the differential equation (2.10) for a specific cosmology. Choosing an Einstein-de Sitter (EdS) model, where $\Omega_\mathrm{m} = 1$ and $\Omega_\Lambda = 0$, as underlying model, we find

$$D_+(t) = a(t)\,, \tag{2.12}$$
$$D_-(t) = a^{-3/2}(t)\,. \tag{2.13}$$

The explicit dependence on the scale factor reveals that the first solution is increasing with the expansion of the Universe, while the second decays with an evolving scale factor. For this reason, the solutions are also known as *growing mode* $D_+(t)$ and *decaying mode* $D_-(t)$. As the decaying mode fades away quickly, it is not important for structure formation and we focus in the following on the growing mode.

2.3 Growth factor

For a general cosmology it is more difficult to obtain the two solutions for the corresponding second-order differential equation of the density contrast. Nevertheless, one finds in analogy to the EdS consideration still one growing and one decaying mode. In case of a ΛCDM Universe one can show that the growing mode takes the form

$$D_+(a, \Omega_\mathrm{m}, \Omega_\Lambda) \propto \frac{H(a)}{H_0} \int_0^a \mathrm{d}a' \left[1 + \Omega_\mathrm{m}\left(\frac{1}{a'} - 1\right) + \Omega_\Lambda\left(a'^2 - 1\right)\right]^{-3/2}\,. \tag{2.14}$$

In order to remove the missing proportionality constant, one normalizes by the growing mode today, such that

$$D(a) \equiv \frac{D_+(a)}{D_+(a=1)}\,. \tag{2.15}$$

The resulting quantity is referred to as *growth factor*. The density contrast at an arbitrary time $a(t)$ is then related to the present-day density contrast at $a_0 = a(t_0)$ by

$$\delta(a, \boldsymbol{x}) = D(a)\delta(a_0, \boldsymbol{x})\,. \tag{2.16}$$

Explicit expressions for the growth factor in ΛCDM cosmologies can be found in Hamilton [29]. In the following, we will omit the dependence on Ω_m and Ω_Λ.

2.4 Non-linear solution

In order to determine the non-linear evolution of the dark matter density field δ and the velocity field $\boldsymbol{v}$ in perturbation theory, one expands the fields around their linear solutions and assumes a curl free velocity field, i.e. $\nabla \times \boldsymbol{v} = \boldsymbol{0}$. We define a new quantity for the divergence of the field $\theta \equiv \nabla \boldsymbol{v}$. This will allow us to find a self-consistent description of the evolution equations in Fourier space. The ansatz for an expansion of the non-linear fields can then be chosen according to

$$\delta(\boldsymbol{x}, a) = \sum_{n=1}^{\infty} \delta_n(\boldsymbol{x}, a)\,, \qquad \theta(\boldsymbol{x}, a) = \sum_{n=1}^{\infty} \theta_n(\boldsymbol{x}, a)\,, \tag{2.17}$$

where n denotes the order in perturbation theory of the corresponding field. Note that the perturbative approach implies that δ_1 and θ_1 are linear in the initial density field, δ_2 and θ_2 are quadratic in the initial field, etc.

2.4.1 Fourier representation

On large scales, different Fourier modes evolve independently from each other and thus conserve the primordial statistics of the perturbations. It is therefore convenient to work in Fourier space and Fourier transform the fields as well as the non-linear equations of motions (2.2), (2.6) and (2.7). With the Fourier conventions as defined in Sect. 3.3, the non-linear continuity equation (2.6) takes the form:

$$a\dot{\tilde{\delta}}(\boldsymbol{k}, a) + \tilde{\theta}(\boldsymbol{k}, a) = -\int \mathrm{d}^3x\, \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}} \nabla(\boldsymbol{v}\delta)(\boldsymbol{x}, a)\,, \tag{2.18}$$

where the tilde sign indicates Fourier transformed quantities. Using integration by parts and rewriting $\boldsymbol{v}$ and δ as Fourier integrals, the right-hand side of Eq. (2.18) becomes

$$a\dot{\tilde{\delta}}(\boldsymbol{k}, a) + \tilde{\theta}(\boldsymbol{k}, a) = -\int \frac{\mathrm{d}^3k_1}{(2\pi)^3} \int \frac{\mathrm{d}^3k_2}{(2\pi)^3}\, \mathrm{i}\boldsymbol{k}\tilde{\boldsymbol{v}}(\boldsymbol{k}_1, a)\tilde{\delta}(\boldsymbol{k}_2, a) \int \mathrm{d}^3x\, \mathrm{e}^{\mathrm{i}\boldsymbol{x}\cdot(\boldsymbol{k}-\boldsymbol{k}_1-\boldsymbol{k}_2)}\,. \tag{2.19}$$

Transforming the last integral into Dirac's delta distribution and making use of $\tilde{\boldsymbol{v}}(\boldsymbol{k}_1) \propto \boldsymbol{k}_1$, which is valid for velocities exhibiting only a divergence part [6], one finds

$$a\dot{\tilde{\delta}}(\boldsymbol{k}, a) + \tilde{\theta}(\boldsymbol{k}, a) = -\int \frac{\mathrm{d}^3k_1}{(2\pi)^3} \int \mathrm{d}^3k_2 \delta_{\mathrm{D}}(\boldsymbol{k}-\boldsymbol{k}_1-\boldsymbol{k}_2)\alpha(\boldsymbol{k}_1, \boldsymbol{k}_2)\tilde{\theta}(\boldsymbol{k}_1, a)\tilde{\delta}(\boldsymbol{k}_2, a)\,, \tag{2.20}$$

where

$$\alpha(\boldsymbol{k}_1, \boldsymbol{k}_2) = \frac{(\boldsymbol{k}_1 + \boldsymbol{k}_2) \cdot \boldsymbol{k}_1}{k_1^2} \,. \tag{2.21}$$

Analogously, transforming the Euler equation (2.2) into Fourier space yields

$$a\dot{\tilde{\theta}}(\boldsymbol{k}, a) + \frac{3H_0^2\Omega_\mathrm{m}}{2a}\tilde{\delta}(\boldsymbol{k}, a) = -\int \mathrm{d}^3x \mathrm{e}^{\mathrm{i}\boldsymbol{x}\cdot\boldsymbol{k}} \left[\nabla \cdot (\boldsymbol{v}\nabla)\boldsymbol{v}\right](\boldsymbol{x}, a) \,, \tag{2.22}$$

if one combines it with the Fourier transformed expression of the Poisson equation (2.7). After some algebra, similar to the derivation of Eq. (2.20), one ends up with

$$a\dot{\tilde{\theta}}(\boldsymbol{k}, a) + \frac{3H_0^2\Omega_\mathrm{m}}{2a}\tilde{\delta}(\boldsymbol{k}, a) = -\int \frac{\mathrm{d}^3k_1}{(2\pi)^3} \int \mathrm{d}^3k_2 \delta_\mathrm{D}(\boldsymbol{k} - \boldsymbol{k}_1 - \boldsymbol{k}_2)\beta(\boldsymbol{k}_1, \boldsymbol{k}_2)\tilde{\theta}(\boldsymbol{k}_1, a)\tilde{\theta}(\boldsymbol{k}_2, a) \,, \tag{2.23}$$

where

$$\beta(\boldsymbol{k}_1, \boldsymbol{k}_2) = \frac{|\boldsymbol{k}_1 + \boldsymbol{k}_2|^2(\boldsymbol{k}_1 \cdot \boldsymbol{k}_2)}{2k_1^2k_2^2} \,. \tag{2.24}$$

The functions $\alpha(\boldsymbol{k}_1, \boldsymbol{k}_2)$ and $\beta(\boldsymbol{k}_1, \boldsymbol{k}_2)$ appear because of the non-linear terms in the continuity (2.6) and Euler equation (2.2) and encode the non-linear evolution of the density and velocity field. For this reason, they are also known as *fundamental mode coupling* functions. In order to solve Eqs. (2.20) and (2.23), one has to take into account all pairs of wave-vectors $(\boldsymbol{k}_1, \boldsymbol{k}_2)$ whose sum is $\boldsymbol{k}$ as imposed by Dirac's delta distribution. This reflects the properties of a spatially homogeneous Universe.

2.4.2 EdS cosmology

The non-linear equations of motions (2.20) and (2.23), which we derived in the previous section, cannot be solved analytically for an arbitrary cosmology. Restricting to an EdS Universe, we outline how these equations can be solved. The advantage of this choice is that we can remove the time-dependence from the equations.

In order to solve the equations of motions, one uses the following ansatz for the density and velocity field, in analogy to Eq. (2.17):

$$\tilde{\delta}(\boldsymbol{k}, a) = \sum_{n=1}^{\infty} D_n(a)\tilde{\delta}_n(\boldsymbol{k}) \,, \qquad \tilde{\theta}(\boldsymbol{k}, a) = -\dot{a}\sum_{n=1}^{\infty} D_n(a)\tilde{\theta}_n(\boldsymbol{k}) \,, \tag{2.25}$$

where n denotes the order of the fields and the growth factor. Note that D_n scales as a^n, while δ_n scales with the initial density field as $(\delta_1)^n$. The choice for this ansatz was made such that the dependence of the linear density field on the scale factor for $n = 1$ is reproduced. The extra factor $\dot{a}$ appears in order to fulfill the requirement $\tilde{\delta}_1 = \tilde{\theta}_1$, which is imposed by the linear Continuity equation (2.8). In particular, this implies that in the linear approximation, the initial perturbations are described completely by the initial density field δ_1.

2.4.3 Coupling functions

If one inserts this ansatz into the equations of motions (2.20) and (2.23) and compares equal powers of the scale factor a, one recovers the following solution for the n-th order of the density and the divergence of the velocity [6]:

$$\tilde{\delta}_n(\boldsymbol{k}) = \int \frac{\mathrm{d}^3 q_1 \mathrm{d}^3 q_2 \cdots \mathrm{d}^3 q_n}{(2\pi)^{n-1}} \, \delta_{\mathrm{D}} \left(\boldsymbol{k} - \boldsymbol{q}_1 - \ldots - \boldsymbol{q}_n\right) F_n(\boldsymbol{q}_1, \ldots, \boldsymbol{q}_n) \tilde{\delta}_1(\boldsymbol{q}_1) \cdots \tilde{\delta}_1(\boldsymbol{q}_n) \tag{2.26}$$

$$\tilde{\theta}_n(\boldsymbol{k}) = \int \frac{\mathrm{d}^3 q_1 \mathrm{d}^3 q_2 \cdots \mathrm{d}^3 q_n}{(2\pi)^{n-1}} \, \delta_{\mathrm{D}} \left(\boldsymbol{k} - \boldsymbol{q}_1 - \ldots - \boldsymbol{q}_n\right) G_n(\boldsymbol{q}_1, \ldots, \boldsymbol{q}_n) \tilde{\delta}_1(\boldsymbol{q}_1) \cdots \tilde{\delta}_1(\boldsymbol{q}_n) \,, \tag{2.27}$$

where the functions F_n and G_n are the n-th order *coupling functions*. Starting with the initial values $F_1 \equiv 1$ and $G_1 \equiv 1$, one recovers from Eqs. (2.26) and (2.27) the following recursion relations [26, 34]:

$$\begin{aligned} F_n(\boldsymbol{q}_1, \ldots, \boldsymbol{q}_n) = \sum_{m=1}^{n-1} \frac{G_m(\boldsymbol{q}_1, \ldots, \boldsymbol{q}_m)}{(2n+3)(n-1)} [&(2n+1)\alpha(\boldsymbol{k}_1, \boldsymbol{k}_2) F_{n-m}(\boldsymbol{q}_{m+1}, \ldots, \boldsymbol{q}_n) \\ &+ 2\,\beta(\boldsymbol{k}_1, \boldsymbol{k}_2) G_{n-m}(\boldsymbol{q}_{m+1}, \ldots, \boldsymbol{q}_n)], \end{aligned} \tag{2.28}$$

$$\begin{aligned} G_n(\boldsymbol{q}_1, \ldots, \boldsymbol{q}_n) = \sum_{m=1}^{n-1} \frac{G_m(\boldsymbol{q}_1, \ldots, \boldsymbol{q}_m)}{(2n+3)(n-1)} [&3\alpha(\boldsymbol{k}_1, \boldsymbol{k}_2) F_{n-m}(\boldsymbol{q}_{m+1}, \ldots, \boldsymbol{q}_n) \\ &+ 2n\,\beta(\boldsymbol{k}_1, \boldsymbol{k}_2) G_{n-m}(\boldsymbol{q}_{m+1}, \ldots, \boldsymbol{q}_n)] \end{aligned} \tag{2.29}$$

where we used $\boldsymbol{k}_1 \equiv \boldsymbol{q}_1 + \ldots + \boldsymbol{q}_m$ and $\boldsymbol{k}_2 \equiv \boldsymbol{q}_{m+1} + \ldots + \boldsymbol{q}_n$ to shorten the notation. $\alpha(\boldsymbol{k}_1, \boldsymbol{k}_2)$ and $\beta(\boldsymbol{k}_1, \boldsymbol{k}_2)$ are the fundamental mode coupling functions as defined in Eqs. (2.21, 2.24).

Since Eqs. (2.26) and (2.27) integrate over all possible q_n, we can equivalently study the symmetrized versions of the coupling functions:

$$F_n^{(s)}(\boldsymbol{q}_1,\ldots,\boldsymbol{q}_n) = \frac{1}{n!}\sum_{\pi} F_n(\boldsymbol{q}_{\pi(1)},\ldots,\boldsymbol{q}_{\pi(n)}), \tag{2.30}$$

$$G_n^{(s)}(\boldsymbol{q}_1,\ldots,\boldsymbol{q}_n) = \frac{1}{n!}\sum_{\pi} G_n(\boldsymbol{q}_{\pi(1)},\ldots,\boldsymbol{q}_{\pi(n)}), \tag{2.31}$$

where the sum is taken over all possible permutations of the set $\{1,\ldots,n\}$. In this way, subsequent calculations will simplify. For the symmetrized second and third order coupling functions as derived from the recursion relations in Eqs. (2.26, 2.27), we obtain

$$F_2^{(s)}(\boldsymbol{q}_1,\boldsymbol{q}_2) = \frac{5}{7} + \frac{2}{7}\frac{(\boldsymbol{q}_1\cdot\boldsymbol{q}_2)^2}{q_1^2 q_2^2} + \frac{1}{2}\frac{\boldsymbol{q}_1\cdot\boldsymbol{q}_2}{q_1 q_2}\left(\frac{q_1}{q_2}+\frac{q_2}{q_1}\right), \tag{2.32}$$

$$G_2^{(s)}(\boldsymbol{q}_1,\boldsymbol{q}_2) = \frac{3}{7} + \frac{4}{7}\frac{(\boldsymbol{q}_1\cdot\boldsymbol{q}_2)^2}{q_1^2 q_2^2} + \frac{1}{2}\frac{\boldsymbol{q}_1\cdot\boldsymbol{q}_2}{q_1 q_2}\left(\frac{q_1}{q_2}+\frac{q_2}{q_1}\right) \tag{2.33}$$

and

$$\begin{aligned} F_3^{(s)}(\boldsymbol{q}_1,\boldsymbol{q}_2,\boldsymbol{q}_3) = {} & \frac{7}{54}[\alpha(\boldsymbol{q}_1,\boldsymbol{q}_2+\boldsymbol{q}_3)F_2^{(s)}(\boldsymbol{q}_2,\boldsymbol{q}_3) + \alpha(\boldsymbol{q}_2,\boldsymbol{q}_1+\boldsymbol{q}_3)F_2^{(s)}(\boldsymbol{q}_1,\boldsymbol{q}_3) \\ & + \alpha(\boldsymbol{q}_3,\boldsymbol{q}_1+\boldsymbol{q}_2)F_2^{(s)}(\boldsymbol{q}_1,\boldsymbol{q}_2)] \\ & + \frac{4}{54}[\beta(\boldsymbol{q}_1,\boldsymbol{q}_2+\boldsymbol{q}_3)G_2^{(s)}(\boldsymbol{q}_2,\boldsymbol{q}_3) + \beta(\boldsymbol{q}_2,\boldsymbol{q}_1+\boldsymbol{q}_3)G_2^{(s)}(\boldsymbol{q}_1,\boldsymbol{q}_3) \\ & + \beta(\boldsymbol{q}_3,\boldsymbol{q}_1+\boldsymbol{q}_2)G_2^{(s)}(\boldsymbol{q}_1,\boldsymbol{q}_2)] \\ & + \frac{7}{54}[\alpha(\boldsymbol{q}_1+\boldsymbol{q}_2,\boldsymbol{q}_3)G_2^{(s)}(\boldsymbol{q}_1,\boldsymbol{q}_2) + \alpha(\boldsymbol{q}_1+\boldsymbol{q}_3,\boldsymbol{q}_2)G_2^{(s)}(\boldsymbol{q}_1,\boldsymbol{q}_3) \\ & + \alpha(\boldsymbol{q}_2+\boldsymbol{q}_3,\boldsymbol{q}_1)G_2^{(s)}(\boldsymbol{q}_2,\boldsymbol{q}_3)] \end{aligned} \tag{2.34}$$

Since in the following we make only use of the symmetrized version of the coupling functions, we drop the superscript from now on. The unsymmetrized fourth-order coupling function can be found in Goroff et al. [26].

2.5 ΛCDM cosmologies

For general cosmologies, finding a perturbative ansatz is more complicated as one cannot expect the solutions at each order to be separable functions in $\boldsymbol{k}$ and a. In particular, the n-th order growing mode does not necessarily scale as $D_1(a)$ (or $a^n(t)$ as for EdS cosmologies). However, it is possible to show that a simple approximation

to the Fourier space equations of motions for general ΛCDM cosmologies leads to separable solutions for arbitrary orders. The ansatz used in this case is

$$\tilde{\delta}(\boldsymbol{k}, a) = \sum_{n=1}^{\infty} D_n(a)\tilde{\delta}_n(\boldsymbol{k})\,. \tag{2.35}$$

$$\tilde{\theta}(\boldsymbol{k}, a) = -\dot{a}f(\Omega_\mathrm{m}, \Omega_\Lambda)\sum_{n=1}^{\infty} E_n(a)\tilde{\theta}_n(\boldsymbol{k})\,, \tag{2.36}$$

where $f(\Omega_\mathrm{m}, \Omega_\Lambda)$ takes into account the cosmology dependence of the solutions and $E_n(a)$ is a function depending on the growth factor. Astonishingly, the system of equations becomes separable if one chooses $f(\Omega_\mathrm{m}, \Omega_\Lambda) = \Omega_\mathrm{m}^{1/2}$ and $D_n = E_n = (D_1)^n$. For $\Omega_\mathrm{m} = 1$ and $\Omega_\Lambda = 0$, one recovers indeed the same recursion relations (2.26), (2.27) as for the EdS case [71]. As a consequence, all information on the cosmological information is then encoded in the linear growth factor $D_1(a)$ and the function $f(\Omega_\mathrm{m}, \Omega_\Lambda)$.

2.6 Growth suppression

So far, we have only considered the growth of cold dark matter perturbations in a matter dominated background and extended the consideration to general ΛCDM cosmologies. This suffices to describe the late Universe but shortly after the Big-Bang when the Universe is radiation dominated the growth factor must have evolved differently. Apart from this, we neglected pressure effects completely which are important for radiation and baryonic matter perturbations and only considered perturbations on length scales smaller than the Hubble radius $d_\mathrm{H} = cH^{-1}$. In the following, we give a qualitative overview of the additional effects which occur in more general scenarios. A more detailed treatment of these topics can be found e.g. in Peacock [61] or Dodelson [18].

We consider first the case of a cold dark matter perturbation of length scale λ. If $\lambda > d_\mathrm{H}$, a relativistic consideration predicts that perturbations grow with $\delta \propto a^2$ independent of the background cosmology. Once the length scale of the perturbation 'enters' the horizon, i.e. $\lambda < d_\mathrm{H}$, one has to distinguish between different epochs. If this happens during the radiation dominated phase of the Universe, the growth of cold dark matter perturbations stalls, i.e. $\delta \simeq$ const, due to the expansion rate in this epoch (*Meszaros effect*). As soon as $a > a_\mathrm{eq}$, the perturbation continues its growth with $\delta \propto a$. Thus, in contrast to perturbations which enter the horizon after matter-radiation equality, the growth is suppressed by a factor $(a_\mathrm{enter}/a_\mathrm{eq})^2$.

The situation becomes more complicated for baryonic perturbations of the same length scale λ. On superhorizon scales, where $\lambda > d_\mathrm{H}$, the perturbations grow like collisionless fluctuations with $\delta \propto a^2$. Once the perturbations enter the horizon, i.e. $\lambda < d_\mathrm{H}$, it

becomes important whether they are below a critical length scale called the *Jeans length* λ_J, which is the minimal length at which gravity exceeds the opposing pressure gradient. For $\lambda_\mathrm{J} < \lambda < \lambda_\mathrm{H}$, gravity still dominates over pressure but the growth is suppressed. For length scales smaller than the Jeans length, pressure exceeds the gravitational force and the baryonic perturbations begin to behave like *acoustic waves*, i.e. they oscillate due to the pressure in the coupled baryon-photon fluid. After recombination this process stops as baryons and photons decouple and the baryonic perturbations grow in a similar way than dark matter with $\delta \propto a$.

2.7 Transfer function

Although most of the scenarios discussed in the previous section can be treated analytically when they are considered isolated, it becomes difficult when one combines the effects on all scales. Additionally, a realistic treatment would require to consider the growth behavior of all components at once. All these effects are accounted for in the *transfer function* $T(k)$ which can be defined in the following way: consider a density perturbation $\delta(k, a_i)$ at a very early time a_i where all perturbations of interest are well outside the horizon. Additionally, choose a large-scale perturbation with wave-number $k_\mathrm{s} = 2\pi/\lambda_\mathrm{s}$ such that this fluctuation never suffered a period where the growth was inhibited. If we now compare the ratio of an arbitrary fluctuation $\delta(k)$ to a large scale fluctuation $\delta(k_\mathrm{s})$ at early times with the same ratio today, we can define

$$\frac{\delta(k, a=1)}{\delta(k_\mathrm{s}, a=1)} \equiv T(k)\,\frac{\delta(k, a_i)}{\delta(k_\mathrm{s}, a_i)}\,. \tag{2.37}$$

Thus the transfer function $T(k)$ accounts for the change in this ratio and is independent for all relevant scales on the initial behavior of the fluctuations. From the discussion in the previous section, we can infer the qualitative behavior of the transfer function for cold dark matter. For large-scale perturbations, i.e. small wave-numbers k the transfer function should approach 1, since fluctuations larger than the horizon grow unhindered. The Hubble radius at matter-radiation equality a_eq defines a characteristic length scale $\lambda_\mathrm{eq} \approx 16(\Omega_\mathrm{m}h^2)^{-1}\,h^{-1}\,\mathrm{Mpc}$ at which the transfer function decreases significantly. Perturbations smaller than this scale are suppressed by a factor $(a_\mathrm{enter}/a_\mathrm{eq})^2$. This results in a behavior $T(k) \propto k^{-2}$ for scales smaller than λ_eq.

As it is complicated to calculate the transfer function directly, there are fitting formulae available in the literature which cover the most important combinations of components. The one most commonly used is the fitting formula of Bardeen et al. [4], which is valid for cold dark matter perturbations

$$T(k) = \frac{\ln(1+2.34\,q)}{2.34\,q}\left[1+3.89\,q+(16.1\,q)^2+(5.46\,q)^3+(6.71\,q)^4\right]^{-1/4} \tag{2.38}$$

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

$$\Gamma = \Omega_\mathrm{m} h\,. \tag{2.39}$$

Peacock and Dodds extended the validity of the transfer function to small baryonic contributions by setting $\Gamma = \Omega_\mathrm{m} h \exp(-2\Omega_\mathrm{b})$. In this case, the overall shape of the transfer function is conserved but the amplitude is reduced at a certain length scale. General models with $\Omega_\mathrm{m} \neq 1$ can be described as well if one uses

$$\Gamma = \Omega_\mathrm{m} h \exp[-\Omega_\mathrm{b}(1 + \sqrt{2h}/\Omega_\mathrm{b})] \tag{2.40}$$

as discussed in Sugiyama [83]. If one seeks a more precise treatment of the baryonic content, one should make use of the fitting formula by Eisenstein and Hu [22]. They also include the oscillations of the photon-baryon plasma before decoupling. In order to obtain a transfer function for an arbitrary choice of cosmological parameters, one can make use of the publicly available code CMBFAST developed by Zaldarriaga and Seljak [91].

2.8 Power spectrum

With the ingredients of the forgoing sections, we can finally find a description of the density perturbation power spectrum in the linear regime. The only information missing is the actual form of the initial power spectrum. Supported by inflationary scenarios, one assumes the form of a power law, i.e.

$$P(k) \propto A k^{n_\mathrm{s}}\,, \tag{2.41}$$

where n_s is the spectral index. Most models prefer a choice $n_\mathrm{s} \lesssim 1$. For $n_\mathrm{s} = 1$ the spectrum is *scale-invariant*, since all density fluctuations which enter the horizon have the same amplitude independent of the expansion rate. This is also known as *Harrison-Zel'dovich spectrum* and has been predicted more than three decades ago. The amplitude A of the initial power spectrum has to be determined from observations.

If we multiply the initial power spectrum with the transfer function squared, we find the linear power spectrum today

$$P(k) = A k^{n_\mathrm{s}} T^2(k)\,. \tag{2.42}$$

Combining this result with the linear growth factor, we obtain for the evolution of the power spectrum in the linear approximation

$$P(k,a) = Ak^{n_s}T^2(k)D_1^2(a)\,, \tag{2.43}$$

which is valid for arbitrary times a.

Chapter 3

Cosmological random fields

Cosmology deals with understanding the formation and evolution of structure on large scales. Since such a model has to cover the enormous size and time-scale of the whole Universe, it cannot be expected to predict the exact configuration of the Universe at any given time. Instead cosmologists try to find a statistical description of the Universe in terms of a set of continuous random fields. These random fields behave stochastic in the sense that we assume that our Universe represents one particular realization from an ensemble of (all possible) Universes. They are characterized by a probability density function, which cosmologists seek to estimate. Since observing one realization of the Universe is not sufficient to determine the statistical properties of a random field, one makes use of the *Ergodic Theorem*. It states that for well-behaved fields, ensemble averages are equivalent to local volume averages if the number of independent regions of space is large enough (see Adler [2]). The ergodicity assumption is usually justified for cosmological fields, since measurements are taken in disconnected regions of the sky that can be assumed statistically independent (*Fair Sample Hypothesis*, Peebles [65]). As the Universe is described by a background geometry that assumes homogeneity and isotropy of the (dark) matter distribution on large scales, it is important to consider these properties for the corresponding fields, too. In this way the statistical quantities simplify considerably.

Of particular interest in cosmology are Gaussian random fields, since inflationary models predict the initial matter density field to be distributed randomly. Additionally, the importance of Gaussian fields is originated in the *Central Limit Theorem*, which states that any linear combination of a large number of independent random variables forms a Gaussian distribution independent of the probability distribution describing the single variable, as long as all variables contribute to the overall sum (Kendall and Stuart [44]).

In the following we will give a brief overview of the most important definitions for the statistics of random fields, consider the case of homogeneity, isotropy and Gaussianity for such fields and apply these properties to the dark matter density field of the Universe. A mathematical treatment of random fields is given in Adler [2], whereas a very detailed derivation of the statistical quantities is provided in Kendall and Stuart [44]. Applications for this to cosmology can be found in Martínez and Saar [49].

3.1 Basic definitions

Definition 3.1.1. A *random variable* $g(\boldsymbol{x})$ is an assignment of a real value to an element $\boldsymbol{x} \in \mathbb{R}^n$ that follows a *probability density function* (p.d.f.) $P[g(\boldsymbol{x})]$. $P[g(\boldsymbol{x})]\mathrm{d}g(\boldsymbol{x})$ corresponds to the probability that the assigned value lies in the infinitesimal cube between $g(\boldsymbol{x})$ and $g(\boldsymbol{x}) + \mathrm{d}g(\boldsymbol{x})$.

Definition 3.1.2. A *random field* g is a set of random variables $g(\boldsymbol{x}_i)$ for $\boldsymbol{x}_i \in \mathbb{R}^n$ with a *joint probability density function*

$$P[g(\boldsymbol{x}_1), g(\boldsymbol{x}_2), \ldots, g(\boldsymbol{x}_N)] \; \forall N \in \mathbb{N}\,. \tag{3.1}$$

One specific assignment of values for the random variables $g(\boldsymbol{x}_i)$ is referred to as a *realization* of the random field g.

Definition 3.1.3. If the joint probability density function factorizes to

$$P[g(\boldsymbol{x}_1), g(\boldsymbol{x}_2), \ldots, g(\boldsymbol{x}_N)] = P[g(\boldsymbol{x}_1)]P[g(\boldsymbol{x}_2)] \cdots P[g(\boldsymbol{x}_N)]\,, \tag{3.2}$$

the random variables $g(\boldsymbol{x}_i)$ are called *independent*.

Definition 3.1.4. The *raw moments* of a random field g with p.d.f. P are defined as expectations over products of field evaluations, if the integration is finite:

$$\langle g_1^{n_1} g_2^{n_2} \ldots g_N^{n_N} \rangle \equiv \int \mathrm{d}g_1 \mathrm{d}g_2 \ldots \mathrm{d}g_N \, g_1^{n_1} g_2^{n_2} \ldots g_N^{n_N} \, P\left[g_1, g_2, \ldots, g_N\right]\,, \tag{3.3}$$

where g_i is the short notation of $g(\boldsymbol{x}_i)$. The sum over the powers $m = \sum n_i$ is the *order* of the moment. The brackets $\langle \ldots \rangle$ are often referred to as *ensemble average*.

The statistical properties of a random field g are described completely, if all its m-order moments are known for all $m \in \mathbb{N}$. The most common statistical quantities are the first two moments, which are associated with specific names.

Definition 3.1.5. The *mean* or *expectation value* of a random field $g(\boldsymbol{x})$ is:

$$\mu(\boldsymbol{x}) = \langle g(\boldsymbol{x}) \rangle \equiv \int \mathrm{d}g(\boldsymbol{x}) g(\boldsymbol{x}) P[g(\boldsymbol{x})]\,. \tag{3.4}$$

Definition 3.1.6. The n-th order *central moment* is defined as n-th order raw moment around the mean μ:

$$\langle (g(\boldsymbol{x}) - \langle g(\boldsymbol{x}) \rangle)^n \rangle = \langle (g(\boldsymbol{x}) - \mu(\boldsymbol{x}))^n \rangle = \int \mathrm{d}g(\boldsymbol{x}) (g(\boldsymbol{x}) - \mu(\boldsymbol{x}))^n P[(g(\boldsymbol{x}))]\,. \tag{3.5}$$

In the later sections we consider only random fields with zero-mean, i.e. $\langle g(\boldsymbol{x}) \rangle = 0$. This can be done without loss of generality, since every field g can be written as $g = g' + \langle g \rangle$, where $\langle g' \rangle = 0$ and $\langle g \rangle$ is a deterministic function.

Definition 3.1.7. The second-order central moment or *variance* of a random field $g(\boldsymbol{x})$ is:

$$\sigma_g^2(\boldsymbol{x}) = \langle (g(\boldsymbol{x}) - \mu(\boldsymbol{x}))^2 \rangle \equiv \int \mathrm{d}g(\boldsymbol{x})\, (g(\boldsymbol{x}) - \mu(\boldsymbol{x}))^2 P[g(\boldsymbol{x})]\,. \tag{3.6}$$

Its square root $\sigma_g(\boldsymbol{x})$ is denoted as *standard deviation*.

Definition 3.1.8. If one considers the second-order central moment of a random field g at two locations $\boldsymbol{x}$,$\boldsymbol{y} \in \mathbb{R}^n$, the corresponding quantity is called the *covariance*

$$\begin{aligned} \mathrm{Cov}[g(\boldsymbol{x}), g(\boldsymbol{y})] &\equiv \langle (g(\boldsymbol{x}) - \langle g(\boldsymbol{x}) \rangle)(g(\boldsymbol{y}) - \langle g(\boldsymbol{y}) \rangle) \rangle \\ &= \int \mathrm{d}g(\boldsymbol{x}) \mathrm{d}g(\boldsymbol{y})\, (g(\boldsymbol{x}) - \mu(\boldsymbol{x}))(g(\boldsymbol{y}) - \mu(\boldsymbol{y})) P[g(\boldsymbol{x}), g(\boldsymbol{y})] \end{aligned} \tag{3.7}$$

and is a measure of the dependency between $g(\boldsymbol{x})$ and $g(\boldsymbol{y})$. Likewise one can define the covariance of two different random fields.

Definition 3.1.9. The dimensionless measure of the dependency of $g(\boldsymbol{x})$ and $g(\boldsymbol{y})$ is given by the *correlation coefficient*

$$\rho(\boldsymbol{x}, \boldsymbol{y}) = \frac{\mathrm{Cov}[g(\boldsymbol{x}), g(\boldsymbol{y})]}{\sigma_g(\boldsymbol{x}) \sigma_g(\boldsymbol{y})}\,, \tag{3.8}$$

where $\sigma_g(\boldsymbol{x})$ and $\sigma_g(\boldsymbol{y})$ denote the corresponding standard deviations. It can be shown that $-1 \leq \rho(\boldsymbol{x}, \boldsymbol{y}) \leq 1$ is valid, where a positive value indicates a correlation of the variables and a negative value describes an anti-correlation. From the definition it follows that $\rho(\boldsymbol{x}, \boldsymbol{x}) = \rho(\boldsymbol{y}, \boldsymbol{y}) = 1$.

An alternative way of describing the statistical properties of a random field is by its *connected moments*:

Definition 3.1.10. The n-th *connected* moment of a field g is defined recursively by

$$\langle g_1 \rangle_c \equiv \langle g_1 \rangle \tag{3.9}$$

$$\begin{aligned} \langle g_1 g_2 \ldots g_N \rangle_c \equiv\ & \langle g_1 g_2 \ldots g_N \rangle - \langle g_1 \rangle_c \langle g_2 \rangle_c \ldots \langle g_N \rangle_c \\ & - \langle g_1 \rangle_c \langle g_2 \ldots g_N \rangle_c + \text{perms} \\ & - \langle g_1 g_2 \rangle_c \langle g_3 \ldots g_N \rangle_c + \text{perms} \\ & - \ldots \\ & - \langle g_1 \ldots g_{N-1} \rangle_c \langle g_N \rangle_c + \text{perms}\,, \end{aligned} \tag{3.10}$$

where 'perms' indicates a cyclic permutation of indices.

Definition 3.1.11. In cosmology it is common to denote the n-th order connected moment as *n-point correlation function* (NPCF) with the special symbols:

$$\xi(\boldsymbol{x}_1, \boldsymbol{x}_2) \equiv \langle g(\boldsymbol{x}_1) g(\boldsymbol{x}_2) \rangle_c \,, \tag{3.11}$$

$$\zeta(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3) \equiv \langle g(\boldsymbol{x}_1) g(\boldsymbol{x}_2) g(\boldsymbol{x}_3) \rangle_c \,, \tag{3.12}$$

$$\eta(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \boldsymbol{x}_4) \equiv \langle g(\boldsymbol{x}_1) g(\boldsymbol{x}_2) g(\boldsymbol{x}_3) g(\boldsymbol{x}_4) \rangle_c \,. \tag{3.13}$$

Definition 3.1.12. If we consider the recursion relation (3.9) for the particular case that all points are situated at the same location, we can easily calculate the first few moments:

$$\begin{aligned}
\langle g \rangle_c &= \langle g \rangle \,, \\
\langle g^2 \rangle_c &= \sigma^2 = \langle g^2 \rangle - \langle g \rangle_c^2 \,, \\
\langle g^3 \rangle_c &= \langle g^3 \rangle - 3 \langle g^2 \rangle_c \langle g \rangle_c - \langle g \rangle_c^3 \,, \\
\langle g^4 \rangle_c &= \langle g^4 \rangle - 4 \langle g^3 \rangle_c \langle g \rangle_c - 3 \langle g^2 \rangle_c^2 - 6 \langle g^2 \rangle_c \langle g \rangle_c^2 - \langle g \rangle_c^4 \,,
\end{aligned} \tag{3.14}$$

where the connected moments $\langle g^n \rangle_c$ are referred to as *cumulants*. If g is a mean-zero random field, the above equations simplify considerably.

A powerful quantity for calculating statistical moments is provided by the *characteristic function* ϕ.

Definition 3.1.13. The *characteristic function* of a random field g is defined as the mean of the exponentiated random field:

$$\phi_g(t) \equiv \langle \exp(\mathrm{i}tg) \rangle = \int P[g] \, \mathrm{e}^{\mathrm{i}tg} \mathrm{d}g \,. \tag{3.15}$$

It is also referred to as *moment-generating function*.

Theorem 1 (Inversion Theorem)**.** *If the characteristic function ϕ_g of a random field g is known, it uniquely determines its probability density functions and vice versa.*

A proof of this statement can be found in [44].

The n-th moment of a p.d.f can now be generated by calculating the n-th derivative of $\phi(t)$ at the point $t = 0$:

$$\phi_g^{(n)}(0) \equiv \left[\frac{\mathrm{d}^n \phi_g}{\mathrm{d}t^n} \right]_{t=0} = \mathrm{i}^n \langle g^n \rangle \,, \tag{3.16}$$

which follows from the power series expansion of the characteristic function $\phi_g(t) = \sum_{n=0}^{\infty} \frac{(\mathrm{i}t)^n}{n!} \langle g^n \rangle$.

In a similar way, we can also generate the cumulants with the help of the characteristic function via the expansion

$$\ln \phi(t)_g = \sum_{n=0}^{\infty} \frac{(\mathrm{i}t)^n}{n!} \langle g^n \rangle_c \,. \tag{3.17}$$

A proof can be found in Kendall and Stuart [44].

3.2 Homogeneous and isotropic random fields

Definition 3.2.1. A random field g is *homogeneous* if all joint p.d.f.s are invariant under a common spatial translation, i.e.

$$P[g(\boldsymbol{x}_1 + \boldsymbol{r}), g(\boldsymbol{x}_2 + \boldsymbol{r}), \ldots, g(\boldsymbol{x}_N + \boldsymbol{r})] = P[g(\boldsymbol{x}_1), g(\boldsymbol{x}_2), \ldots, g(\boldsymbol{x}_N)] \tag{3.18}$$

for all $\boldsymbol{r} \in \mathbb{R}^n$.

Note that if a random field is homogeneous this property is also true for all its moments. This implies that e.g. the 2PCF depends only on the difference between the locations, i.e. $\xi(\boldsymbol{x}_1, \boldsymbol{x}_2) = \xi(\boldsymbol{x}_1 - \boldsymbol{x}_2)$.

Definition 3.2.2. A random field g is *isotropic* if all joint p.d.f.s are invariant under the same spatial rotation in all locations, i.e.

$$P[g(\mathcal{R}\boldsymbol{x}_1), g(\mathcal{R}\boldsymbol{x}_2), \ldots, g(\mathcal{R}\boldsymbol{x}_N)] = P[g(\boldsymbol{x}_1), g(\boldsymbol{x}_2), \ldots, g(\boldsymbol{x}_N)] \tag{3.19}$$

for all rotation matrices $\mathcal{R}$.

Again, this property holds for all the moments of a random field. If a random field g is both homogeneous and isotropic, the 2PCF simplifies to $\xi(\boldsymbol{x}, \boldsymbol{x} + \boldsymbol{r}) = \xi(|\boldsymbol{r}|)$.

3.3 Fourier description for random fields

It is useful to study the properties of random fields in Fourier space. Throughout this work we use the following definition for the Fourier transformations[1]:

Definition 3.3.1. The *forward* transformation $\mathcal{F}$ of a function $g : \mathbb{R}^n \to \mathbb{R}$ is defined as

$$\mathcal{F}[g(\boldsymbol{x})] = \tilde{g}(\boldsymbol{k}) = \int_{\mathbb{R}^n} \mathrm{d}^n x \, g(\boldsymbol{x}) \, \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}} \,. \tag{3.20}$$

[1] This convention can be remembered easily by interpreting $\mathrm{d}^n k/(2\pi)^n$ as the volume element in Fourier space.

Definition 3.3.2. Analogously, the *inverse* transformation $\mathcal{F}^{-1}$ of a function $\tilde{g} : \mathbb{R}^n \to \mathbb{C}$ is given by

$$\mathcal{F}^{-1}[\tilde{g}(\boldsymbol{k})] = g(\boldsymbol{x}) = \int_{\mathbb{R}^n} \frac{\mathrm{d}^n k}{(2\pi)^n}\, \tilde{g}(\boldsymbol{k})\, \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}\,. \tag{3.21}$$

Functions $g(\boldsymbol{x})$ and $\tilde{g}(\boldsymbol{k})$ that fulfill these definitions (3.20) and (3.21) are referred to as *Fourier pairs*. In the following, we will also write $g(\boldsymbol{k})$ instead of $\tilde{g}(\boldsymbol{k})$, where no confusion arises.

The advantage of working in Fourier space are the properties of its transformations. The following ones can be applied for all functions $f, g : \mathbb{R}^n \to \mathbb{R}$ with a defined Fourier transformation:

- The Fourier transformation is linear, i.e.

$$\mathcal{F}[af(\boldsymbol{x}) + bg(\boldsymbol{x})] = a\mathcal{F}[f(\boldsymbol{x})] + b\mathcal{F}[g(\boldsymbol{x})]\,. \tag{3.22}$$

- The Fourier transformation of a convolution transforms into a pointwise product

$$\mathcal{F}[f(\boldsymbol{x}) * g(\boldsymbol{x})] = \mathcal{F}[f(\boldsymbol{x})] \cdot \mathcal{F}[g(\boldsymbol{x})]\,. \tag{3.23}$$

- The Fourier transformation of the n-th derivative of a function f simplifies to

$$\mathcal{F}[f^{(n)}(\boldsymbol{x})](\boldsymbol{k}) = (2\pi\mathrm{i}\boldsymbol{k})^n \mathcal{F}[f(\boldsymbol{x})](\boldsymbol{k})\,. \tag{3.24}$$

Since the Dirac delta-distribution is used frequently in this thesis, we will summarize its most important properties briefly.

Definition 3.3.3. We define *Dirac's delta-distribution* via its fundamental property

$$\int \mathrm{d}^n x\, f(\boldsymbol{x})\, \delta_{\mathrm{D}}\left(\boldsymbol{x} - \boldsymbol{a}\right) = f(\boldsymbol{a})\,. \tag{3.25}$$

Definition 3.3.4. Making use of Eq. (3.25) the formal forward Fourier transformation of Dirac's delta-distribution yields

$$\mathcal{F}[\delta_{\mathrm{D}}\left(\boldsymbol{x}\right)] = \int \mathrm{d}^n x\, \delta_{\mathrm{D}}\left(\boldsymbol{x}\right) \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}} = 1\,, \tag{3.26}$$

whereas its inverse Fourier transformation results in

$$\mathcal{F}^{-1}[\delta_{\mathrm{D}}\left(\boldsymbol{k}\right)] = \int \frac{\mathrm{d}^n k}{(2\pi)^n}\, \delta_{\mathrm{D}}\left(\boldsymbol{k}\right)\, \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}} = \frac{1}{(2\pi)^n}\,. \tag{3.27}$$

We will therefore treat $\delta_{\mathrm{D}}\left(\boldsymbol{k}\right)$ and $1/(2\pi)^n$, as well as $\delta_{\mathrm{D}}\left(\boldsymbol{x}\right)$ and 1, as Fourier pairs.

As in real space (see Sect. 3.1) it is also possible to consider statistical moments in Fourier space.

Definition 3.3.5. The most important one is the *power spectrum* $P_g(\boldsymbol{k})$, which can be defined for any complex, homogeneous random field $\tilde{g}$ as

$$\langle \tilde{g}(\boldsymbol{k})\tilde{g}^*(\boldsymbol{k}')\rangle \equiv (2\pi)^n \delta_{\mathrm{D}}\left(\boldsymbol{k}-\boldsymbol{k}'\right) P_g(\boldsymbol{k})\,, \tag{3.28}$$

where $\delta_{\mathrm{D}}\left(\boldsymbol{k}\right)$ denotes the n-dimensional Dirac delta distribution.

Its occurrence is a direct consequence of the homogeneity of the random field and means that different Fourier modes are uncorrelated. Note that one can show easily that the power spectrum is always real-valued, if one calculates the complex conjugate of Eq. (3.28) and exchanges $\boldsymbol{k}$ and $\boldsymbol{k}'$.

Theorem 2 (Wiener-Khienchin)**.** *The 2PCF or auto-correlation function of a random field g is related to the power spectrum in the following way:*

$$P_g(\boldsymbol{k}) = \int \mathrm{d}^3 y\, \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{y}} \xi(\boldsymbol{y})\,, \tag{3.29}$$

if we restrict our consideration to homogeneous random fields g in three-dimensional space. In other words the power spectrum is the (forward) Fourier transformation of the 2PCF.

Proof. Applying the forward Fourier transformation as defined in Eq. (3.20) to the 2PCF yields

$$\begin{aligned}
\langle \tilde{g}(\boldsymbol{k})\tilde{g}(\boldsymbol{k}')\rangle &= \int \mathrm{d}^3 x\, \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}} \int \mathrm{d}^3 x'\, \mathrm{e}^{\mathrm{i}\boldsymbol{k}'\cdot\boldsymbol{x}'} \xi(\boldsymbol{x}-\boldsymbol{x}') \\
&= \int \mathrm{d}^3 x\, \mathrm{e}^{\mathrm{i}\boldsymbol{x}(\boldsymbol{k}+\boldsymbol{k}')} \int \mathrm{d}^3 y\, \mathrm{e}^{-\mathrm{i}\boldsymbol{k}'\cdot\boldsymbol{y}} \xi(\boldsymbol{y}) \\
&= (2\pi)^3\, \delta_{\mathrm{D}}\left(\boldsymbol{k}+\boldsymbol{k}'\right) P_g(\boldsymbol{k})\,,
\end{aligned} \tag{3.30}$$

where $\boldsymbol{y} = \boldsymbol{x}-\boldsymbol{x}'$ and since $\langle \tilde{g}(\boldsymbol{k})\tilde{g}(\boldsymbol{k}')\rangle \equiv (2\pi)^3 \delta_{\mathrm{D}}\left(\boldsymbol{k}+\boldsymbol{k}'\right) P_g(\boldsymbol{k})$ holds for real-valued fields $g(\boldsymbol{x})$. □

For the inverse transformation, the corresponding equation is valid, i.e.

$$\xi(\boldsymbol{y}) = \frac{1}{(2\pi)^3} \int \mathrm{d}^3 k\, \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{y}} P_g(\boldsymbol{k})\,. \tag{3.31}$$

If one assumes additionally isotropy of the random field g, $P_g(\boldsymbol{k}) = P_g(|\boldsymbol{k}|)$ and $\xi(\boldsymbol{y}) = \xi(|\boldsymbol{y}|)$ are valid.

This calculation can be extended to any order of n-spectra. Omitting the details, we summarize the results for the third- and fourth-order moments.

Definition 3.3.6. For any real, homogeneous random field g, we define the *bispectrum* $B_g(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3)$ and the *trispectrum* $T_g(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3, \boldsymbol{k}_4)$ as

$$\begin{aligned} \langle \tilde{g}(\boldsymbol{k}_1)\tilde{g}(\boldsymbol{k}_2)\tilde{g}(\boldsymbol{k}_3)\rangle_c &= (2\pi)^3\, \delta_{\mathrm{D}}\left(\boldsymbol{k}_{123}\right) B_g(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3)\,, \\ \langle \tilde{g}(\boldsymbol{k}_1)\tilde{g}(\boldsymbol{k}_2)\tilde{g}(\boldsymbol{k}_3)g(\boldsymbol{k}_4)\rangle_c &= (2\pi)^3\, \delta_{\mathrm{D}}\left(\boldsymbol{k}_{1234}\right) T_g(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3, \boldsymbol{k}_4)\,, \end{aligned} \tag{3.32}$$

where $\delta_{\mathrm{D}}\left(\boldsymbol{k}\right)$ denotes the three-dimensional delta-distribution and $\boldsymbol{k}_{i\ldots j} = \boldsymbol{k}_i + \cdots + \boldsymbol{k}_j$. The delta-distribution appears due to the homogeneity of the field and causes that only closed configurations in Fourier space contribute.

Definition 3.3.7. In analogy to Eq. (3.29) the bi- and the trispectrum are related to their real space counterparts by the following equations:

$$\begin{aligned} B_g(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) &= \int \mathrm{d}^3x\, \mathrm{d}^3y\, \mathrm{e}^{\mathrm{i}\boldsymbol{k}_1\boldsymbol{x}}\, \mathrm{e}^{\mathrm{i}\boldsymbol{k}_3\boldsymbol{y}}\, \zeta(\boldsymbol{x}, \boldsymbol{y})\,, \\ T_g(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3, \boldsymbol{k}_4) &= \int \mathrm{d}^3x\, \mathrm{d}^3y\, \mathrm{d}^3z\, \mathrm{e}^{\mathrm{i}\boldsymbol{k}_1\boldsymbol{x}}\, \mathrm{e}^{\mathrm{i}\boldsymbol{k}_3\boldsymbol{y}}\, \mathrm{e}^{\mathrm{i}\boldsymbol{k}_4\boldsymbol{z}}\, \eta(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})\,, \end{aligned} \tag{3.33}$$

where $\boldsymbol{x} = \boldsymbol{x}_1 - \boldsymbol{x}_2$, $\boldsymbol{y} = \boldsymbol{x}_3 - \boldsymbol{x}_2$ and $\boldsymbol{z} = \boldsymbol{x}_4 - \boldsymbol{x}_3$. The additional assumption of isotropy for the fields would make the correlation functions depend only on the magnitude of the defined differences, i.e. $\zeta(\boldsymbol{x}, \boldsymbol{y}) = \zeta(x, y, \phi_{xy})$ and $\eta(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) = \eta(x, y, z, \varphi_{xy}, \varphi_{yz})$, respectively.

3.4 Gaussian random fields

Definition 3.4.1. A mean-zero random field g is called *multivariate Gaussian* if the joint probability distribution function follows a multivariate Gaussian distribution of the form

$$P[g_1, g_2, \ldots, g_N]\, \mathrm{d}g_1 \mathrm{d}g_2 \ldots \mathrm{d}g_N = \frac{\exp\left[-\frac{1}{2}\sum_{ij} g_i C_{ij}^{-1} g_j\right]}{\sqrt{(2\pi)^N \det C}}\, \mathrm{d}g_1 \mathrm{d}g_2 \ldots \mathrm{d}g_N\,, \tag{3.34}$$

where we used for convenience the notation $g_i = g(x_i)$. C_{ij} is denoted as the *covariance matrix* of g, which can be identified with $C_{ij} = \langle g_i g_j \rangle$.

Definition 3.4.2. If each variable g_i of a multivariate Gaussian random field g is independent, then the covariance matrix C_{ij} is diagonal, i.e. $C_{ij} = \delta_{ij}\sigma_i^2$ holds. In this case Eq. (3.34) is the product of the p.d.f.s of N Gaussian distributions

$$P[g_1]P[g_2]\cdots P[g_N]\, \mathrm{d}g_1 \mathrm{d}g_2 \ldots \mathrm{d}g_N = \prod_{i=1}^{N} \frac{\exp\left(-g_i^2/2\sigma_{g_i}^2\right)}{\sqrt{2\pi\sigma_{g_i}^2}}\, \mathrm{d}g_1 \mathrm{d}g_2 \ldots \mathrm{d}g_N\,. \tag{3.35}$$

An alternative characterization of a Gaussian random field g is provided by the following theorem.

Theorem 3 (Wick Theorem)**.** *A random field g is multivariate Gaussian, if the moments of its elements g_i decompose to:*

$$\begin{aligned} \langle g_1, g_2, \ldots, g_{2p} \rangle &= \sum_{all\ perms.} \langle g_1 g_2 \rangle_c \langle g_3 g_4 \rangle_c \cdot \ldots \cdot \ldots \cdot \langle g_{2p-1} g_{2p} \rangle_c \,, \\ \langle g_1, g_2, \ldots, g_{2p+1} \rangle &= 0 \end{aligned} \tag{3.36}$$

and vice versa. In other words, for Gaussian random fields any even correlation function can be expressed in terms of a product of all permutations of 2PCFs, whereas the odd correlation functions vanish.

A direct consequence of the theorem (3.36) is that all higher-order connected moments (apart from $N = 2$) are equal to zero for Gaussian fields.

Example 3.4.3. Consider a zero-mean random variable g with a Gaussian p.d.f.

$$P[g]\,\mathrm{d}g = \frac{1}{\sqrt{2\pi\sigma_g^2}} \exp\left(-g^2/2\sigma_g^2\right) \mathrm{d}g \tag{3.37}$$

and let us calculate its first few moments. This can be done easily with the help of the characteristic function as defined in Eq. (3.15):

$$\begin{aligned} \phi_g(t) &= \int_{-\infty}^{\infty} \mathrm{d}g\, \frac{1}{\sqrt{2\pi\sigma_g^2}} \exp\left(-g^2/2\sigma_g^2\right) \exp\left(\mathrm{i}tg\right) \\ &= \int_{-\infty}^{\infty} \mathrm{d}g\, \frac{1}{\sqrt{2\pi\sigma_g^2}} \exp\left(-\frac{1}{2\sigma_g^2}\left[(g - \mathrm{i}t\sigma_g^2)^2 + t^2\sigma_g^4\right]\right) \\ &= \frac{1}{\sqrt{2\pi\sigma_g^2}} \exp\left(-\frac{t^2\sigma_g^2}{2}\right) \int_{-\infty}^{\infty} \mathrm{d}g\, \exp\left[\frac{1}{2\sigma_g^2}(g - \mathrm{i}t\sigma_g^2)^2\right] \\ &= \exp\left(-\frac{t^2\sigma_g^2}{2}\right). \end{aligned} \tag{3.38}$$

The n-th raw moment can now be determined by evaluating the nth-derivative of the characteristic function at $t = 0$ (see Eq. (3.16)). In this way, we obtain the following results:

$$\begin{aligned} \langle g \rangle &= 0\,, \\ \langle g^2 \rangle &= \sigma_g^2 \\ \langle g^4 \rangle &= 3\sigma_g^4 \\ \langle g^{2n+1} \rangle &= 0\,, \end{aligned} \tag{3.39}$$

which is consistent with the prediction of Wick's theorem (3.36).

Example 3.4.4. Consider zero-mean random variables g_i for $i = 1, \ldots, 4$ with a Gaussian p.d.f.. The 4-th moments are then

$$\begin{aligned} \langle g_1 g_2 g_3 g_4 \rangle &= \langle g_1 g_2 \rangle \langle g_3 g_4 \rangle + \langle g_1 g_3 \rangle \langle g_2 g_4 \rangle + \langle g_1 g_4 \rangle \langle g_2 g_3 \rangle \\ \langle g_1 g_2 g_3 g_4 \rangle_c &= \langle g_1 g_2 g_3 g_4 \rangle - \langle g_1 g_2 \rangle \langle g_3 g_4 \rangle - \langle g_1 g_3 \rangle \langle g_2 g_4 \rangle - \langle g_1 g_4 \rangle \langle g_2 g_3 \rangle \\ &= 0\,. \end{aligned} \tag{3.40}$$

Gaussian distributions exhibit several appealing properties, which makes their consideration important:

- If the random fields f and g are independent and normally distributed, any linear combination $h = af + bg$ of the two also follows a Gaussian distribution for $a, b \in \mathbb{R}$.
- The Fourier transformation of a Gaussian p.d.f. P is again of a Gaussian shape, since $\mathcal{F}[P(g)] = \int \mathrm{d}g\, P(g) \mathrm{e}^{\mathrm{i}kg} = \phi_g(k) = \exp\left(-\frac{k^2\sigma_g^2}{2}\right)$ according to Eq. (3.38).
- A Gaussian mean-zero random field is completely determined by its power spectrum.

Theorem 4 (Central Limit Theorem)**.** *If the random fields $f_1, \ldots, f_n$ are independent and distributed according to* any *p.d.f. with finite mean and variance, then the sum $g = \sum_{i=1}^{n} f_n$ will have a p.d.f. that converges to a Gaussian for large n.*

A proof of this can be found in [2] or [44].

3.5 The density contrast field

The central quantity for analyzing galaxy and mass clustering statistically is the *density contrast* or *density perturbation field* $\delta(\boldsymbol{x}, t)$ of dark matter. It is defined as the relative deviation of the local dark matter density field $\rho(\boldsymbol{x}, t)$ from the average density of the Universe $\bar{\rho}(t)$:

$$\delta(\boldsymbol{x}, t) = \frac{\rho(\boldsymbol{x}, t) - \bar{\rho}(t)}{\bar{\rho}(t)}\,, \tag{3.41}$$

where the t-dependence is usually omitted. In this way, the mean density contrast is by definition zero and we can make use of the properties for mean-zero Gaussian random fields if we consider the density contrast at early times.

Its connected moments are used to describe and measure the statistical properties of structure in the Universe. For our consideration we need all orders up to the fourth

moment, namely the power-, bi- and trispectrum:

$$\begin{aligned}
\langle\tilde{\delta}(\boldsymbol{k}_1)\rangle &= 0\,,\\
\langle\tilde{\delta}(\boldsymbol{k}_1)\tilde{\delta}(\boldsymbol{k}_2)\rangle_{\mathrm{c}} &= (2\pi)^3\,\delta_{\mathrm{D}}\left(\boldsymbol{k}_{12}\right)P(\boldsymbol{k}_1)\,,\\
\langle\tilde{\delta}(\boldsymbol{k}_1)\tilde{\delta}(\boldsymbol{k}_2)\tilde{\delta}(\boldsymbol{k}_3)\rangle_{\mathrm{c}} &= (2\pi)^3\,\delta_{\mathrm{D}}\left(\boldsymbol{k}_{123}\right)B(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3)\,,\\
\langle\tilde{\delta}(\boldsymbol{k}_1)\tilde{\delta}(\boldsymbol{k}_2)\tilde{\delta}(\boldsymbol{k}_3)\tilde{\delta}(\boldsymbol{k}_4)\rangle_{\mathrm{c}} &= (2\pi)^3\,\delta_{\mathrm{D}}\left(\boldsymbol{k}_{1234}\right)T(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\boldsymbol{k}_4)\,,
\end{aligned} \tag{3.42}$$

where $\delta_{\mathrm{D}}\left(\boldsymbol{k}\right)$ denotes the three-dimensional Dirac delta-distribution and $\boldsymbol{k}_{i\ldots j} = \boldsymbol{k}_i + \ldots + \boldsymbol{k}_j$. If one considers moments of the energy density $\rho(\boldsymbol{k})$ instead, one can obtain them from the density contrast moments, as can be seen for the example of the real space 2PCF:

$$\begin{aligned}
\xi(\boldsymbol{r}_a,\boldsymbol{r}_b) &\equiv \left\langle\delta(\boldsymbol{r}_a)\delta(\boldsymbol{r}_b)\right\rangle\\
&= \left\langle(\rho(\boldsymbol{r}_a)-\bar{\rho})(\rho(\boldsymbol{r}_b)-\bar{\rho})/\bar{\rho}^2\right\rangle\\
&= \left\langle\rho(\boldsymbol{r}_a)\rho(\boldsymbol{r}_b)\right\rangle/\bar{\rho}^2 - 1\,.
\end{aligned} \tag{3.43}$$

Since the power spectrum determines a Gaussian p.d.f. completely, it is one of the most important quantities to measure in cosmology. Due to the homogeneity and isotropy of the Universe, the power spectrum depends only on the magnitude of the wave-vector $\boldsymbol{k}$ and allows one to perform the angular integration in Eq. (3.31) immediately, so that

$$\xi(r) = \frac{1}{(2\pi)^3}\int \mathrm{d}k\,4\pi k^2 P(k)\frac{\sin(kr)}{kr}\,. \tag{3.44}$$

With this equation it is possible to obtain a useful expression for the variance of the fluctuation field $\delta(\boldsymbol{x})$ in real space:

$$\begin{aligned}
\sigma^2 \equiv \langle\delta(\boldsymbol{x})^2\rangle &= \xi(0)\\
&= \frac{1}{(2\pi)^3}\int 4\pi k^2 P(k)\mathrm{d}k\\
&= \frac{1}{(2\pi)^3}\int 4\pi k^3 P(k)\mathrm{d}\ln k\,.
\end{aligned} \tag{3.45}$$

It is used to determine the *dimensionless* power spectrum $\Delta^2(k)$, which is defined as the variance per logarithmic wave-number ($\Delta^2(k) \equiv \mathrm{d}\sigma^2/\mathrm{d}\ln k$):

$$\Delta^2(k) = \frac{k^3}{2\pi^2}P(k)\,. \tag{3.46}$$

In this way $\Delta^2(k) = 1$ obtains a clear interpretation[2]. It means that there are order unity density fluctuations from modes in the logarithmic bin around wave-number $\boldsymbol{k}$ (see e.g. Peacock [60]).

[2]In the literature there are two conventions used for the Fourier transformations (3.20), (3.21) leading to different relations between $P(k)$ and $\Delta(k)$: Our definition leads to a factor $k^3/2\pi^2$ following the book by Peebles [65], whereas the alternative one has a factor $4\pi k^3$ as e.g. used in Bernardeau et al. [6].

3.5.1 Smoothed density contrast field

If we want to apply the formalism for random fields to real objects, we have to take into account the finite size of the observed objects. This means that we are not interested in the behavior of the density field at scales smaller than a typical length R. A common practice in cosmology is therefore to consider the *smoothed* perturbation field

$$\delta_R(\boldsymbol{x}) = \int \mathrm{d}^3x'\, W_R(\boldsymbol{x}' - \boldsymbol{x})\delta(\boldsymbol{x}')\,, \tag{3.47}$$

which is defined as the convolution of the original field $\delta(\boldsymbol{x})$ with some filter function $W_R(x)$. Typically, two different types of filter are used: Either the top-hat filter

$$W_R^{\mathrm{TH}}(r) = \begin{cases} \frac{3}{4\pi R^3}\,, & \text{for} \quad r \leq R, \\ 0\,, & \text{otherwise}\,, \end{cases} \tag{3.48}$$

which smooths over a finite spherical volume of radius R or the Gaussian filter

$$W_R^{\mathrm{G}}(r) = \frac{1}{\sqrt{(2\pi)^3}R^3} \exp\left(-r^2/2R^2\right). \tag{3.49}$$

Note that both filters are normalized such, that $\int d^3r W_R(r) = 1$ holds.

The Fourier transformation of the smoothed perturbation field is related to the unfiltered field according to

$$\tilde{\delta}_R(\boldsymbol{k}) = \tilde{W}_R(\boldsymbol{k})\delta(\boldsymbol{k})\,, \tag{3.50}$$

where the Fourier counterparts of the top-hat and Gaussian filter are calculated to be

$$\tilde{W}_R^{\mathrm{TH}}(k) = \frac{3}{k^3R^3}\left[\sin(kR) - kR\cos(kR)\right] \tag{3.51}$$

and

$$\tilde{W}_R^{\mathrm{G}}(k) = \exp(-k^2R^2/2)\,, \tag{3.52}$$

respectively. With the above considerations it is now possible to find an expression for the smoothed variance:

$$\sigma_R^2(r) = \int \Delta^2(k)|\tilde{W}_R(k)|^2 \mathrm{d}\ln k\,. \tag{3.53}$$

3.5.2 Bispectrum

With the ingredients from perturbation theory (see Chapter 2), we can calculate arbitrary moments of the density field in the linear approximation. This description

is only valid on large scales, where the linear density contrast $\delta_1(\boldsymbol{k})$ follows the initial Gaussian distribution. The basic procedure is to expand the non-linear density contrast field at the considered location according to

$$\delta(\boldsymbol{k}) = \delta_1(\boldsymbol{k}) + \delta_2(\boldsymbol{k}) + \dots \,, \tag{3.54}$$

where the $\delta_n(\boldsymbol{k})$ are coupled to the n-th power of the linear density contrast $\delta_1(\boldsymbol{k})$ via the integral equation (2.26). If one calculates a specific moment of the density contrast, one has to find the non-vanishing contribution of lowest order in the $\delta_n(\boldsymbol{k})$ to that moment. This order is also known as *tree-level* contribution. A detailed description of this construction can be found in Fry [25]. In the following, we will derive the perturbation theory expressions of the bi- and trispectrum expressions explicitly.

The lowest non-vanishing order of the perturbation theory bispectrum has the form $\langle\delta_1\delta_1\delta_2\rangle_\mathrm{c}$, since the contribution $\langle\delta_1\delta_1\delta_1\rangle_\mathrm{c}$ is 0 for Gaussian random fields due to Wick's theorem (see Eq. 3.36). Thus, we have to expand the density field to second-order to find the tree-level bispectrum:

$$\begin{aligned}
(2\pi)^3\,\delta_\mathrm{D}\left(\boldsymbol{k}_{123}\right) B_\mathrm{pt}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3) &\equiv \langle\delta(\boldsymbol{k}_1)\delta(\boldsymbol{k}_2)\delta(\boldsymbol{k}_3)\rangle_\mathrm{tree} \\
&= \langle[\delta_1(\boldsymbol{k}_1)+\delta_2(\boldsymbol{k}_1)][\delta_1(\boldsymbol{k}_2)+\delta_2(\boldsymbol{k}_2)] \\
&\quad \times [\delta_1(\boldsymbol{k}_3)+\delta_2(\boldsymbol{k}_3)]\rangle \\
&\simeq \langle\delta_1(\boldsymbol{k}_1)\delta_1(\boldsymbol{k}_2)\delta_2(\boldsymbol{k}_3)\rangle + \langle\delta_1(\boldsymbol{k}_1)\delta_2(\boldsymbol{k}_2)\delta_1(\boldsymbol{k}_3)\rangle \\
&\quad + \langle\delta_2(\boldsymbol{k}_1)\delta_1(\boldsymbol{k}_2)\delta_1(\boldsymbol{k}_3)\rangle \,,
\end{aligned} \tag{3.55}$$

where we neglect higher-order contributions in the third step. To relate the second-order density field to the linear one, we use integral equation (2.26) in the form

$$\delta_2(\boldsymbol{k}) = \int \frac{\mathrm{d}^3 q_1}{(2\pi)^3} \int \mathrm{d}^3 q_2\, \delta_\mathrm{D}\left(\boldsymbol{k}-\boldsymbol{q}_1-\boldsymbol{q}_2\right) F_2(\boldsymbol{q}_1,\boldsymbol{q}_2)\delta_1(\boldsymbol{q}_1)\delta_1(\boldsymbol{q}_2)\,, \tag{3.56}$$

where F_2 denotes the second-order coupling function as defined in Sect. 2.4.3. Inserting this into Eq. (3.55), we finally find

$$\begin{aligned}
B_\mathrm{pt}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3) &= 2\,[F_2(\boldsymbol{k}_1,\boldsymbol{k}_2)P(k_1)P(k_2) + F_2(\boldsymbol{k}_2,\boldsymbol{k}_3)P(k_2)P(k_3) \\
&\quad + F_2(\boldsymbol{k}_3,\boldsymbol{k}_1)P(k_3)P(k_1)] \\
&= 2F_2(\boldsymbol{k}_1,\boldsymbol{k}_2)P(k_1)P(k_2) + 2\,\mathrm{perms.}\,,
\end{aligned} \tag{3.57}$$

where the permutations are considered with respect to the wave-vectors $\boldsymbol{k}_i$ and $P(k_i)$ denotes the linear power spectrum as defined in Eq. (2.42).

3.5.3 Trispectrum

For the lowest non-vanishing order of the perturbation theory trispectrum we have to expand the density contrast up to third-order, since $\langle\delta_1\delta_1\delta_1\delta_1\rangle_\mathrm{c}$ is 0 for Gaussian random fields due to Wick's theorem (see Eq. 3.36). The tree-level trispectrum splits into two types of terms $\langle\delta_1\delta_1\delta_2\delta_2\rangle_\mathrm{c}$ and $\langle\delta_1\delta_1\delta_1\delta_3\rangle_\mathrm{c}$ and one finds:

$$\begin{aligned}(2\pi)^3\,\delta_\mathrm{D}\left(\boldsymbol{k}_{1234}\right)T_\mathrm{pt}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\boldsymbol{k}_4) &\equiv \langle\delta(\boldsymbol{k}_1)\delta(\boldsymbol{k}_2)\delta(\boldsymbol{k}_3)\delta(\boldsymbol{k}_4)\rangle_\mathrm{tree}\\ &= \langle[\delta_1(\boldsymbol{k}_1)+\delta_2(\boldsymbol{k}_1)+\delta_3(\boldsymbol{k}_1)][\delta_1(\boldsymbol{k}_2)+\delta_2(\boldsymbol{k}_2)+\delta_3(\boldsymbol{k}_2)]\\ &\quad\times[\delta_1(\boldsymbol{k}_3)+\delta_2(\boldsymbol{k}_3)+\delta_3(\boldsymbol{k}_3)]\rangle\\ &\simeq \langle\delta_1(\boldsymbol{k}_1)\delta_1(\boldsymbol{k}_2)\delta_2(\boldsymbol{k}_3)\delta_2(\boldsymbol{k}_4)\rangle + 5\,\text{perms.}\\ &\quad+\langle\delta_1(\boldsymbol{k}_1)\delta_1(\boldsymbol{k}_2)\delta_1(\boldsymbol{k}_3)\delta_3(\boldsymbol{k}_4)\rangle + 3\,\text{perms.}\,, \quad (3.58)\end{aligned}$$

where 'n perms.' denotes that we have to consider n permutations of the forgoing expression. In case of the first term the permutations are with respect to δ_2, in the second they are with respect to δ_3. Substituting all non-linear densities with the integral equation (2.26), one obtains after some algebra for the perturbation theory trispectrum

$$T_\mathrm{pt}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\boldsymbol{k}_4) = 4T_a + 6T_b \quad (3.59)$$

with

$$\begin{aligned}T_a = {}& P(k_1)P(k_2)\left[P(k_{13})F_2(\boldsymbol{k}_1,-\boldsymbol{k}_{13})F_2(\boldsymbol{k}_2,\boldsymbol{k}_{13}) + P(k_{14})F_2(\boldsymbol{k}_1,-\boldsymbol{k}_{14})F_2(\boldsymbol{k}_2,\boldsymbol{k}_{14})\right]\\ &+ P(k_1)P(k_3)\left[P(k_{12})F_2(\boldsymbol{k}_1,-\boldsymbol{k}_{12})F_2(\boldsymbol{k}_3,\boldsymbol{k}_{12}) + P(k_{14})F_2(\boldsymbol{k}_1,-\boldsymbol{k}_{14})F_2(\boldsymbol{k}_3,\boldsymbol{k}_{14})\right]\\ &+ P(k_1)P(k_4)\left[P(k_{12})F_2(\boldsymbol{k}_1,-\boldsymbol{k}_{12})F_2(\boldsymbol{k}_4,\boldsymbol{k}_{12}) + P(k_{13})F_2(\boldsymbol{k}_1,-\boldsymbol{k}_{13})F_2(\boldsymbol{k}_4,\boldsymbol{k}_{14})\right]\\ &+ P(k_2)P(k_3)\left[P(k_{21})F_2(\boldsymbol{k}_2,-\boldsymbol{k}_{21})F_2(\boldsymbol{k}_3,\boldsymbol{k}_{21}) + P(k_{24})F_2(\boldsymbol{k}_2,-\boldsymbol{k}_{24})F_2(\boldsymbol{k}_3,\boldsymbol{k}_{24})\right]\\ &+ P(k_2)P(k_4)\left[P(k_{21})F_2(\boldsymbol{k}_2,-\boldsymbol{k}_{21})F_2(\boldsymbol{k}_4,\boldsymbol{k}_{21}) + P(k_{23})F_2(\boldsymbol{k}_2,-\boldsymbol{k}_{23})F_2(\boldsymbol{k}_4,\boldsymbol{k}_{23})\right]\\ &+ P(k_3)P(k_4)\left[P(k_{31})F_2(\boldsymbol{k}_3,-\boldsymbol{k}_{31})F_2(\boldsymbol{k}_4,\boldsymbol{k}_{31}) + P(k_{32})F_2(\boldsymbol{k}_3,-\boldsymbol{k}_{32})F_2(\boldsymbol{k}_4,\boldsymbol{k}_{32})\right]\,,\end{aligned}$$

where we use the short form $\boldsymbol{k}_{ij} = \boldsymbol{k}_i + \boldsymbol{k}_j$ and

$$\begin{aligned}T_b = {}& F_3(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3)P(k_1)P(k_2)P(k_3) + F_3(\boldsymbol{k}_2,\boldsymbol{k}_3,\boldsymbol{k}_4)P(k_2)P(k_3)P(k_4)\\ &+ F_3(\boldsymbol{k}_3,\boldsymbol{k}_4,\boldsymbol{k}_1)P(k_3)P(k_4)P(k_1) + F_3(\boldsymbol{k}_4,\boldsymbol{k}_1,\boldsymbol{k}_2)P(k_4)P(k_1)P(k_2)\,.\end{aligned}$$

Note that F_2 and F_3 are the mode coupling functions as defined in Sect. 2.4.3 and $P(k_i)$ denotes the linear power spectrum (see Eq. 2.42). For notational convenience we use in the following for the tree-level trispectrum the short form

$$T_a = P(k_1)P(k_2)\left[P(k_{13})F_2(\boldsymbol{k}_1, -\boldsymbol{k}_{13})F_2(\boldsymbol{k}_2, \boldsymbol{k}_{13}) + P(k_{14})F_2(\boldsymbol{k}_1, -\boldsymbol{k}_{14})F_2(\boldsymbol{k}_2, \boldsymbol{k}_{14})\right] + 5\,\text{perms.}\,, \tag{3.60}$$

and

$$T_b = F_3(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3)P(k_1)P(k_2)P(k_3) + 3\,\text{perms.}\,. \tag{3.61}$$

3.6 The projected density field

In weak lensing we are usually not interested in the three-dimensional polyspectra, but in their two-dimensional projection on the sky. For this reason we consider the projection of the density contrast $\delta(\boldsymbol{x})$ along the backwards directed light cone of the observer. If we choose our coordinate system such that $\boldsymbol{x} = (f_K(w)\boldsymbol{\theta}, w)$, where $\boldsymbol{\theta}$ resembles a two-dimensional vector on the sky and the comoving distance w gives the depth along the line of sight of the observation, the projected field is defined as

$$g_i(\boldsymbol{\theta}) = \int_0^{w_\mathrm{H}} \mathrm{d}w\, q_i(w)\delta(f_K(w)\boldsymbol{\theta}, w)\,, \tag{3.62}$$

where $q_i(w)$ is a weight function, which varies only slowly with w and w_H is the comoving distance of the horizon.

For weak lensing applications the weight functions are usually of the form:

$$q_i(w) = \frac{3}{2}\left(\frac{H_0}{c}\right)^2 \frac{\Omega_\mathrm{m}}{a(w)} \int_w^{w_\mathrm{H}} \mathrm{d}w' p(w') \frac{f_K(w - w')}{f_K(w')}\, f_K(w) \equiv G(w) f_K(w)\,, \tag{3.63}$$

where $p(w)$ denotes the distribution function of observed source galaxies. For this work, we make use of two types of distributions: the first type resembles a realistic redshift distribution

$$p(w)\mathrm{d}w = p(z)\mathrm{d}z = \frac{\beta}{z_0\Gamma(3/\beta)}\left(\frac{z}{z_0}\right)^2 \mathrm{e}^{-(z/z_0)^\beta}\mathrm{d}z\,, \tag{3.64}$$

as suggested by Brainerd et al. [12], where Γ denotes Euler's gamma function. The parameter β specifies how fast the distribution falls off towards high redshifts, while z_0 is related to the mean redshift $\bar{z}$ of the distribution. A common choice is $\beta = 1.5$ and $\bar{z} \approx 1.5 z_0$. The second type assumes that all observed galaxies are located at a single redshift plane at z_s. This corresponds to setting the galaxy distribution in Eq. (3.63) to a Dirac delta-distribution, which is the typical choice for ray-tracing simulations.

A detailed derivation of the weak lensing weight function in Eq. (3.63) can be found in Bartelmann and Schneider [5]. If we insert this weight function into Eq. (3.62), we obtain an expression for the projected density contrast field or *effective convergence*

$$\kappa(\boldsymbol{\theta}) = \int_0^{w_\mathrm{H}} \mathrm{d}w\, G(w) f_K(w) \delta(f_K(w)\boldsymbol{\theta}, w)\,. \tag{3.65}$$

As for the three-dimensional case (see Sect. 3.5), we can calculate the first few moments of the convergence field in Fourier space:

$$\begin{aligned}
\langle\kappa(\boldsymbol{l}_1)\rangle &= 0\,,\\
\langle\kappa(\boldsymbol{l}_1)\kappa(\boldsymbol{l}_2)\rangle_\mathrm{c} &= (2\pi)^2\,\delta_\mathrm{D}\left(\boldsymbol{l}_{12}\right) P_\kappa(\boldsymbol{l}_1)\,,\\
\langle\kappa(\boldsymbol{l}_1)\kappa(\boldsymbol{l}_2)\kappa(\boldsymbol{l}_3)\rangle_\mathrm{c} &= (2\pi)^2\,\delta_\mathrm{D}\left(\boldsymbol{l}_{123}\right) B_\kappa(\boldsymbol{l}_1,\boldsymbol{l}_2,\boldsymbol{l}_3)\,,\\
\langle\kappa(\boldsymbol{l}_1)\kappa(\boldsymbol{l}_2)\kappa(\boldsymbol{l}_3)\kappa(\boldsymbol{l}_4)\rangle_\mathrm{c} &= (2\pi)^2\,\delta_\mathrm{D}\left(\boldsymbol{l}_{1234}\right) T_\kappa(\boldsymbol{l}_1,\boldsymbol{l}_2,\boldsymbol{l}_3,\boldsymbol{l}_4)\,,
\end{aligned} \tag{3.66}$$

where $\delta_\mathrm{D}(\boldsymbol{x})$ denotes the two-dimensional delta-distribution and $\boldsymbol{l}_{i\ldots j} = \boldsymbol{l}_i + \ldots + \boldsymbol{l}_j$.

If one performs the projection as described in Eq. (3.65) of the density contrast $\delta(\boldsymbol{x})$ and calculates the Fourier transformation of $\langle\kappa(\boldsymbol{\theta}_1)\kappa(\boldsymbol{\theta}_2)\rangle$, one obtains the effective convergence power spectrum

$$P_\kappa(l) = \int_0^{w_\mathrm{H}} \mathrm{d}w\, G^2(w) P_\delta\left(\frac{l}{f_K(w)}, w\right)\,, \tag{3.67}$$

which relates the 2PCF of the projected field to that of the three-dimensional field. This important result is one form of the *Limber equation*. A detailed calculation of this can be found e.g. in Bartelmann and Schneider [5] or Kaiser [41].

Similarly, the projected versions of the bi- and trispectrum can be derived:

$$B_\kappa(\boldsymbol{l}_1,\boldsymbol{l}_2,\boldsymbol{l}_3) = \int_0^{w_\mathrm{H}} \frac{\mathrm{d}w}{f_K}\, G^3(w) B_\delta\left(\frac{\boldsymbol{l}_1}{f_K}, \frac{\boldsymbol{l}_2}{f_K}, \frac{\boldsymbol{l}_3}{f_K}, w\right)\,, \tag{3.68}$$

$$T_\kappa(\boldsymbol{l}_1,\boldsymbol{l}_2,\boldsymbol{l}_3,\boldsymbol{l}_4) = \int_0^{w_\mathrm{H}} \frac{\mathrm{d}w}{f_K^2}\, G^4(w) T_\delta\left(\frac{\boldsymbol{l}_1}{f_K}, \frac{\boldsymbol{l}_2}{f_K}, \frac{\boldsymbol{l}_3}{f_K}, \frac{\boldsymbol{l}_4}{f_K}, w\right)\,, \tag{3.69}$$

where $f_K \equiv f_K(w)$ denotes the angular diameter distance and B_δ and T_δ the three-dimensional matter bi- and trispectrum as defined in Eqs. (3.42).

In analogy to the three-dimensional case in Eq. (3.46), we define the *dimensionless convergence power spectrum* $\Delta^2_\kappa(l)$ as the variance per logarithmic wave-number ($\Delta^2_\kappa(l) \equiv \mathrm{d}\sigma^2/\mathrm{d}\ln l$):

$$\Delta^2_\kappa(l) = \frac{l^2}{2\pi} P_\kappa(l)\,. \tag{3.70}$$

This quantity is convenient for displaying the convergence power spectrum in Sect. 6.2.

3.7 Moment Estimation

In the previous sections, we gave an introduction to the theory of cosmological random fields. Observables such as the density contrast can be understood as a set of random variables assigned with a joint probability function. The nature of the probability function itself can be determined with the help of all the moments of the random variable. We now estimate some of these moments from measurements.

3.7.1 Properties of estimators

Definition 3.7.1. Let $x^1, x^2, \ldots, x^N$ be N measurements of a random variable x and let its p.d.f. $P = P(x^1, x^2, \ldots, x^N; \theta)$ depend additionally on some unknown parameter θ. An *estimator* $\hat{\theta}_N = \hat{\theta}_N(x^1, x^2, \ldots, x^N)$ is a function of the observed data, whose value is a guess of the unknown parameter θ. Estimators are in the following denoted with a hat, while the number of measurements N in the index is usually omitted.

There is no unique way on how to construct an estimator for a specific parameter. But it is important that the estimator inhibits certain properties. The most important ones are: *consistency*, *unbiasedness*, *efficiency* and *robustness*, which will be defined in the subsequent definitions.

Definition 3.7.2. An estimator $\hat{\theta}_N$ is said to be *consistent*, if its value converges in probability to the true value θ as the number of measurements N increases. More formally, it means that any estimator $\hat{\theta}$ determined from N measurements is consistent, if for all $\epsilon > 0$ and $\eta > 0$ there is some n such that the probability that

$$|\hat{\theta}_N - \theta| < \epsilon \tag{3.71}$$

is greater than $1 - \eta$ for all $N > n$.

Definition 3.7.3. The difference between the expectation value of $\hat{\theta}_N$ and its true value θ is called the *bias* b of an estimator, i.e. $\langle \hat{\theta}_N \rangle - \theta = b$. If $b = 0$, the estimator $\hat{\theta}_N$ is said to be *unbiased*.

Example 3.7.4. Let us consider the *sample mean*

$$\hat{\mu} = \frac{1}{N} \sum_{i=1}^{N} x^i \,, \tag{3.72}$$

where N is the number of observed data points. Eq. (3.72) is an example of an unbiased estimator of the true mean μ, since $\langle \hat{\mu} \rangle = \frac{1}{N} \sum_{i=1}^{N} \langle x^i \rangle = \langle x \rangle = \mu$.

Example 3.7.5. A biased estimator for the variance σ^2 is given by

$$\widehat{\sigma^2} = \frac{1}{N} \sum_{i=1}^{N} \left(x^i - \hat{\mu} \right)^2 \,, \tag{3.73}$$

where $\hat{\mu}$ is the sample mean as defined in Eq. (3.72). By calculating the mean of $\widehat{\sigma^2}$ to be

$$\langle\widehat{\sigma^2}\rangle = \frac{N-1}{N}\left(\langle x^2\rangle - \langle x\rangle^2\right) = \frac{N-1}{N}\sigma^2\,, \tag{3.74}$$

where we considered different data points x^i to be independent, we find that the bias is $b = \sigma^2/N$. Multiplying Eq. (3.73) with a factor $N/(N-1)$ yields an unbiased estimator for the variance, which we refer to as *sample variance*.

Definition 3.7.6. If the sample variance of an estimator $\hat{\theta}_N$ is lower than the one of some other estimator of the same parameter, it is called *more efficient*. If a *minimum variance unbiased estimator* $\sigma^2_{\min}$ exists, the efficiency of an estimator $\hat{\theta}_N$ is defined as

$$\mathrm{eff}(\hat{\theta}_N) = \frac{\sigma^2_{\min}}{\sigma^2_{\hat{\theta}}}\,, \tag{3.75}$$

where $\sigma^2_{\hat{\theta}}$ denotes the variance of the estimator. The estimator $\hat{\theta}$ is said to be *efficient*, if $\mathrm{eff}(\hat{\theta}_N) = 1$.

Remark. Under certain regularity conditions (see [43] for details) the minimum variance of a biased estimator is given by the *Cramér-Rao-bound*:

$$\sigma^2_{\hat{\theta}} \leq \sigma^2_{\min} = \left(1 + \frac{\partial b}{\partial\theta}\right)^2 \Big/ I(\theta)\,, \tag{3.76}$$

where

$$I(\theta) = \left\langle \left(\frac{\partial}{\partial\theta}\sum_{i=1}^{N} \ln P(x^i;\theta)\right)^2 \right\rangle \tag{3.77}$$

is the *Fisher information*, b the bias and $P(x^i;\theta)$ the p.d.f. of the data point x^i. For unbiased estimators the nominator in Eq. (3.76) is one.

3.7.2 Estimation of covariance matrices

Lemma 3.7.7. Let x, y be two random variables, which have been measured N times in independent pairs (x^i, y^i) and let $\hat{\mu}_x, \hat{\mu}_y$ be their estimated sample means as defined in Eq. (3.72). We define now two different estimators for the covariance of the random variables x and y. The *maximum likelihood* estimator is defined as

$$\widehat{\mathrm{Cov}}[x,y]_{\mathrm{MLE}} \equiv \frac{1}{N}\sum_{i=1}^{N}(x^i - \hat{\mu}_x)(y^i - \hat{\mu}_y) \tag{3.78}$$

and is biased, whereas the *sample covariance* estimator is given by

$$\widehat{\mathrm{Cov}}[x,y]_{\mathrm{SE}} \equiv \frac{1}{N-1}\sum_{i=1}^{N}(x^i - \hat{\mu}_x)(y^i - \hat{\mu}_y) \tag{3.79}$$

and is unbiased.

Proof. We calculate the expected value of $\sum_{i=1}^{N}(x^i - \hat{\mu}_x)(y^i - \hat{\mu}_y)$:

$$\begin{aligned}
&\Big\langle \sum_{i=1}^{N} \left(x^i - \hat{\mu}_x^N\right)\left(y^i - \hat{\mu}_y^N\right) \Big\rangle \\
= &\Big\langle \sum_{i=1}^{N} \Big(x^i y^i - x^i \frac{1}{N}\sum_{j=1}^{N} y^j - y^i \frac{1}{N}\sum_{j=1}^{N} x^j + \frac{1}{N}\sum_{j=1}^{N} x^j \frac{1}{N}\sum_{k=1}^{N} y^k \Big) \Big\rangle \\
= &\Big\langle \sum_{i=1}^{N} x^i y^i - \frac{1}{N}\sum_{i=1}^{N} x^i \sum_{j=1}^{N} y^j \Big\rangle \\
= &\; N\langle xy\rangle - \frac{1}{N}\Big\langle \sum_{i=1}^{N} x^i \Big\rangle \Big\langle \sum_{j\neq i}^{N} y^j \Big\rangle - \frac{1}{N}\Big\langle \sum_{i=1}^{N} x^i y^i \Big\rangle \\
= &\; (N-1)\left(\langle xy\rangle - \langle x\rangle\langle y\rangle\right) = (N-1)\,\mathrm{Cov}[x,y]\,. \qquad (3.80)
\end{aligned}$$

Multiplying Eq. (3.80) with $1/N$ gives the biased estimator for the covariance (Eq. (3.78)), whereas with $1/(N-1)$ results in the unbiased estimator of Eq. (3.79). Note that in the limit of large N both estimators yield the same result. □

3.7.3 Errors of estimators

We are now interested in the expected error when we estimate the mean and variance of a random variable x from samples.

Definition 3.7.8. We measure the performance of an estimator $\hat{\theta}$ by its *mean square error*

$$\mathrm{MSE} \equiv \langle(\hat{\theta} - \theta)^2\rangle = \sigma_{\hat{\theta}}^2 + b^2\,, \qquad (3.81)$$

which combines the uncertainty due to variance and bias.

Lemma 3.7.9. Let $\hat{\mu}$ be the sample mean of some random variable x that has been determined from N measurements. The standard deviation of $\hat{\mu}$ is called *sample error of the mean* and is given by

$$\hat{\sigma}_{\mathrm{SEM}} = \frac{\sigma_x}{\sqrt{N}}\,, \qquad (3.82)$$

where σ_x is the standard deviation of the random variable x.

Proof.

$$\begin{aligned}
\hat{\sigma}^2_{\mathrm{SEM}} &= \sigma^2_{\hat{\mu}_x} = \langle \hat{\mu}_x^2 \rangle - \langle \hat{\mu}_x \rangle^2 \\
&= \left\langle \frac{1}{N} \sum_{i=1}^{N} x^i \frac{1}{N} \sum_{j=1}^{N} x^j \right\rangle - \mu^2 \\
&= \left\langle \frac{1}{N} \sum_{i=1}^{N} x^i \right\rangle \left\langle \frac{1}{N} \sum_{j \neq i}^{N} x^j \right\rangle \left\langle \frac{1}{N^2} \sum_{i=1}^{N} (x^i)^2 \right\rangle - \mu^2 \\
&= \frac{1}{N} \langle x \rangle (N-1) \langle x \rangle + \frac{1}{N} \langle x^2 \rangle - \mu^2 \\
&= \frac{1}{N} \left(\langle x^2 \rangle - \langle x \rangle^2 \right) = \frac{\sigma_x^2}{N}
\end{aligned} \tag{3.83}$$

□

Lemma 3.7.10. Let $\hat{\sigma}_x^2$ be the sample variance of some random variable x, which has been determined from N independent measurements. The *sample error of the variance* is then given by

$$\sigma_{\mathrm{SEV}} = \langle \hat{\sigma}_x^2 \rangle = \frac{1}{\sqrt{N}} \left(m_4 - \frac{N-3}{N-1} \sigma_x^4 \right)^{-1/2} , \tag{3.84}$$

where m_4 is the fourth central moment of the random variable x as defined in Eq. (3.5) and σ_x its standard deviation.

For a proof see e.g. Kendall and Stuart [44].

Remark. In the limit of large N, Eq. (3.84) becomes

$$\sigma_{\mathrm{SEV}} = \frac{1}{\sqrt{N}} \left(m_4 - \sigma_x^4 \right)^{-1/2} . \tag{3.85}$$

If the random variable x follows a Gaussian distribution, Eq. (3.84) simplifies to

$$\sigma_{\mathrm{SEV}} = \frac{\sigma_x^2}{\sqrt{2(N-1)}} , \tag{3.86}$$

since $m_4 = 3\sigma_x^4$ holds due to Eq. (3.38).

Chapter 4

Halo model description of dark matter

In the previous chapters we introduced the notion of random fields and how to describe the statistical properties of the dark matter density field with their help. As long as we are in the linear regime, where $|\delta| \ll 1$, perturbation theory provides a reliable basis to derive analytical expressions for the N-point correlation functions (N-PCF). However, a lot of interesting phenomena in cosmology, such as gravitational clustering, occur on highly non-linear scales, where the perturbative description long since broke down. In order to explore the statistics of the density field in the non-linear regime, we can make use of numerical N-body simulations. Although simulations give us some important insight into structure formation, their disadvantage is that they are computationally very costly and do not offer a physical understanding of the results found. An alternative way to investigate the non-linear regime is provided by the semi-analytic *halo model*, which gained a lot of interest in the literature recently [17, 51, 74]. It assumes that all dark matter in the Universe is bound in virialized halos. If one has an appropriate description of these halos, it is possible to find an analytic expression for the N-point correlation function of dark matter. This approach suggests a separate treatment on different scales: on small scales the statistics is dominated by the spatial distribution of dark matter *within* the halos and one makes use of results from numerical simulations yielding typical halo mass profiles. On large scales the statistics depends on the *abundance of the halos* and their *clustering properties* and is described by perturbation theory. As a consequence, the power spectrum splits into two terms

$$P(k) = P_{1\mathrm{h}}(k) + P_{2\mathrm{h}}(k)\,,$$

which indicate the different behavior in the two regimes as illustrated in Fig. 4.1. On small scales the power spectrum is dominated by the *1-halo*-term $P_{1\mathrm{h}}$, which describes correlations of dark matter within *one* halo, whereas on large scales the largest contribution comes from the *2-halo*-term $P_{2\mathrm{h}}$, which accounts for correlations between *two* halos.

This chapter is organized in the following way: section 4.1 covers the analytic treatment of the halo model and provides a recipe to calculate any N-PCF for a continuous mass distribution of halos. This consideration is completely general in the sense that the specific description of the halos can be included afterwards. It is known as

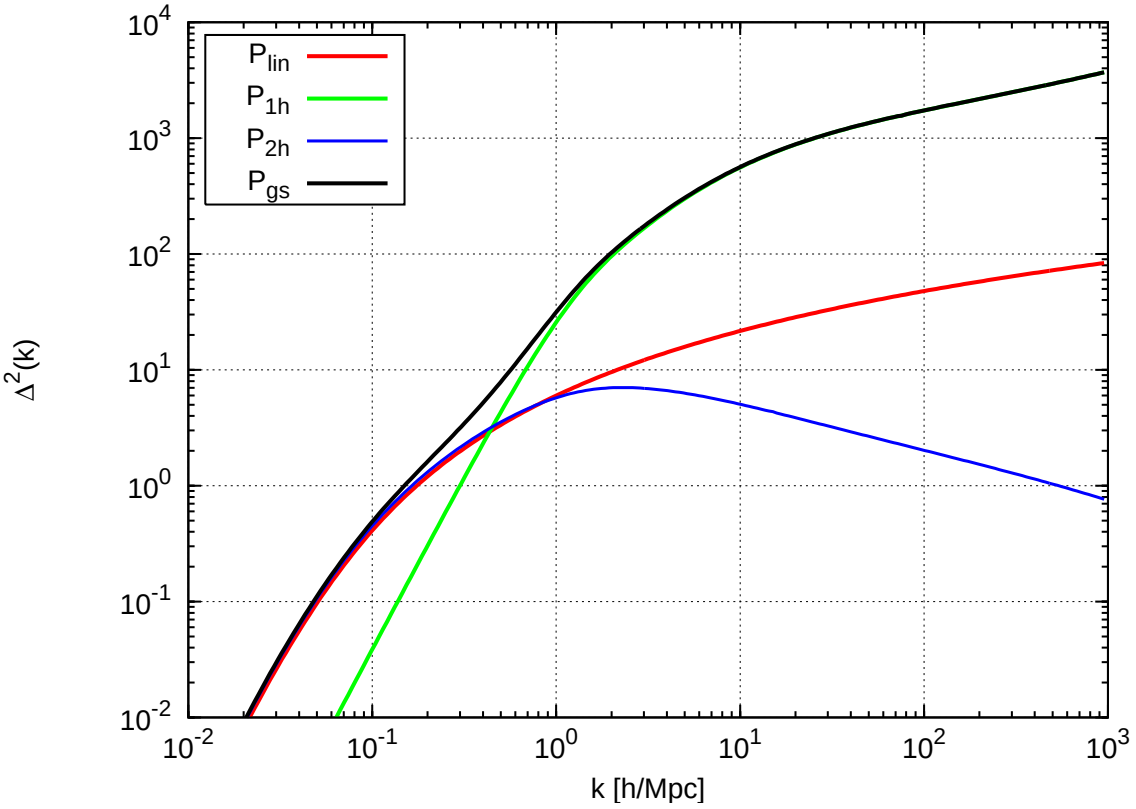

Figure 4.1: *Dimensionless power spectrum* $\Delta^2(k) \equiv k^3/2\pi^2 P(k)$ *against wave-number* k *as calculated with the halo model. The complete power spectrum (black line) has two contributions: One from the* 1*-halo-term (green line), which contributes mainly on small scales and another from the* 2*-halo-term (blue line), which is dominant on large scales. The linear power spectrum (red line) is appropriate only for wave-numbers below* $k = 0.1\,h\,\mathrm{Mpc}^{-1}$.

Bertschinger-Scherrer-formalism [69]. The subsequent sections address the various ingredients of the halo model, such as the *mass function* of halos (Sect. 4.3), the *halo bias* (Sect. 4.4) or the *halo profile* (Sect. 4.5). Since the description of these ingredients is not unique, their choice has to be adjusted judiciously according to the addressed problem. The theoretical background for describing analytically when a bound object has formed is provided by the *spherical collapse model*, which we review in section 4.2. Its main result consists of two quantities, the linear density contrast of a virialized halo $\delta_{\rm sc}$ and the non-linear counterpart $\Delta_{\rm vir}$. The linear density contrast is used to approximate the time-evolution of a collapsing object from a small overdensity to a virialized object. It is important for defining a mass function of how frequent halos of a specific mass are and for the halo bias consideration. The non-linear density contrast gives the actual density contrast of an object that has just formed and is used to parametrize the halo profile. After considering the different ingredients of the halo model, we will define in section 4.6 a building block to simplify the notation for the N-PCFs considerably. Finally, in section 4.7, we will summarize the choice of ingredients used for this thesis.

4.1 Statistics for halo density fields

In the following, we will derive a general formalism to calculate the N-PCF of dark matter for continuous density fields. The treatment will be completely general. All specific properties of halos, such as their profile, abundance or spatial distribution can be incorporated into the description later. A detailed discussion of the choice of ingredients for the halo model is the main topic of the subsequent sections. Historically, this formalism goes back to Neyman and Scott [58] who analyzed the spatial distribution of galaxies in 1952. Their original description was only valid for discrete density fields and was formulated entirely in terms of real space coordinates. Additionally, they used a model for halo clustering which turned out to be not very realistic. The first to provide a general formalism to characterize the statistical properties of random fields were Scherrer and Bertschinger [69] in 1991. In contrast to Neyman and Scott, they considered the correlation functions also in Fourier space which simplifies the involved convolutions considerably. Furthermore, their formalism allows to incorporate more realistic halo-halo correlations into the model.

4.1.1 Average mass density

The basic assumption of the halo model is that all dark matter is bound in spherical halos which implies that the average density of the Universe is

$$\bar{\rho} = \langle \rho(\boldsymbol{x}) \rangle = \int_0^\infty \mathrm{d}m \, m \, n(m) \,, \tag{4.1}$$

where $n(m)$ is the comoving *halo mass function* denoting the abundance of halos with mass m (see Sect. 4.3 for a detailed consideration) and the brackets denote the ensemble average as defined in section 3.1. Additionally, this means that the density field at an arbitrary position is given as a superposition of all halo density profiles in the Universe at this point:

$$\rho(\boldsymbol{x}) = \sum_i f(m_i, \boldsymbol{x} - \boldsymbol{x}_i) \equiv \sum_i m_i \, u(m_i, \boldsymbol{x} - \boldsymbol{x}_i) \,, \tag{4.2}$$

where $f(m_i, \boldsymbol{x} - \boldsymbol{x}_i)$ denotes the density profile of the i-th halo with center of mass at $\boldsymbol{x}_i$ and $u \equiv f/m_i$ is the normalized profile, which fulfills

$$\int \mathrm{d}^3x \, u(m, \boldsymbol{x} - \boldsymbol{x}_1) = 1 \,. \tag{4.3}$$

By parameterizing the halo profile in this way, we assume that its shape at a position $\boldsymbol{x}$ depends only on the halo mass and distance to the center (see Sect. 4.5). Since we are seeking a description for continuous variables, we rewrite Eq. (4.2) with Dirac's delta distribution such that

$$\rho(\boldsymbol{x}) = \int \mathrm{d}m_1 \, \mathrm{d}^3x_1 \sum_i \delta_\mathrm{D} \left(m_1 - m_i\right) \delta_\mathrm{D} \left(\boldsymbol{x}_1 - \boldsymbol{x}_i\right) m_1 \, u(m_1, \boldsymbol{x} - \boldsymbol{x}_1) \,. \tag{4.4}$$

Equipped with these equations, we check whether averaging over the mass density, as given in Eq. (4.4), really results in the mean density of the Universe:

$$\begin{aligned}
\langle \rho(\boldsymbol{x}) \rangle &= \Big\langle \sum_i m_i \, u(m_i, \boldsymbol{x} - \boldsymbol{x}_i) \Big\rangle \\
&= \Big\langle \int \mathrm{d}m_1 \, \mathrm{d}^3x_1 \sum_i \delta_\mathrm{D} \left(m_1 - m_i\right) \delta_\mathrm{D} \left(\boldsymbol{x}_1 - \boldsymbol{x}_i\right) m_1 u(m_1, \boldsymbol{x} - \boldsymbol{x}_1) \Big\rangle \\
&= \int \mathrm{d}m_1 m_1 \Big\langle \sum_i \delta_\mathrm{D} \left(m_1 - m_i\right) \delta_\mathrm{D} \left(\boldsymbol{x}_1 - \boldsymbol{x}_i\right) \Big\rangle \int \mathrm{d}^3x_1 \, u(m_1, \boldsymbol{x} - \boldsymbol{x}_1) \\
&= \int \mathrm{d}m_1 m_1 \Big\langle \sum_i \delta_\mathrm{D} \left(m_1 - m_i\right) \delta_\mathrm{D} \left(\boldsymbol{x}_1 - \boldsymbol{x}_i\right) \Big\rangle \,.
\end{aligned} \tag{4.5}$$

Comparing this result with Eq. (4.1), we find that $\langle \rho(\boldsymbol{x}) \rangle = \bar{\rho}$ holds, if we identify

$$n(m) = \Big\langle \sum_i \delta_\mathrm{D} \left(m - m_i\right) \delta_\mathrm{D} \left(\boldsymbol{x} - \boldsymbol{x}_i\right) \Big\rangle \tag{4.6}$$

as the halo mass function. Note that the ensemble average goes of the position $\boldsymbol{x}$ and thus the halo mass function is only mass-dependent.

4.1.2 2-point correlation function

With the formalism developed in the foregoing section, we can calculate the 2-point correlation function (2PCF) of the halo density contrast $\xi(\boldsymbol{r}_a, \boldsymbol{r}_b)$ as defined in section 3.5. Since we worked so far in terms of the mass density, we will calculate first the 2PCF of the mass density in the halo model approach and then transform the resulting expression to a 2PCF in terms of the density contrast:

$$\begin{aligned}\left\langle \rho(\boldsymbol{r}_a)\rho(\boldsymbol{r}_b)\right\rangle = &\int \mathrm{d}m_1\, \mathrm{d}^3x_1 u(m_1, \boldsymbol{r}_a - \boldsymbol{x}_1) \int \mathrm{d}m_2\, \mathrm{d}^3x_2 u(m_2, \boldsymbol{r}_b - \boldsymbol{x}_2) \\ &\times \left\langle \sum_{i,j} \delta_\mathrm{D}\left(m_1 - m_i\right) \delta_\mathrm{D}\left(\boldsymbol{x}_1 - \boldsymbol{x}_i\right) \delta_\mathrm{D}\left(m_2 - m_j\right) \delta_\mathrm{D}\left(\boldsymbol{x}_2 - \boldsymbol{x}_j\right) \right\rangle , \quad (4.7)\end{aligned}$$

where we applied Eq. (4.4) twice. In order to simplify the expression, we distinguish between two cases. If $i = j$, we consider correlations within the same halo and obtain for the brackets in the last line of Eq. (4.7)

$$\begin{aligned}&\left\langle \sum_i \delta_\mathrm{D}\left(m_1 - m_i\right) \delta_\mathrm{D}\left(\boldsymbol{x}_1 - \boldsymbol{x}_i\right) \delta_\mathrm{D}\left(m_2 - m_i\right) \delta_\mathrm{D}\left(\boldsymbol{x}_2 - \boldsymbol{x}_i\right) \right\rangle \\ &= \left\langle \sum_i \delta_\mathrm{D}\left(m_1 - m_i\right) \delta_\mathrm{D}\left(m_2 - m_1\right) \delta_\mathrm{D}\left(\boldsymbol{x}_1 - \boldsymbol{x}_i\right) \delta_\mathrm{D}\left(\boldsymbol{x}_2 - \boldsymbol{x}_1\right) \right\rangle \\ &= \delta_\mathrm{D}\left(m_2 - m_1\right) \delta_\mathrm{D}\left(\boldsymbol{x}_2 - \boldsymbol{x}_1\right) \left\langle \sum_i \delta_\mathrm{D}\left(m_1 - m_i\right) \delta_\mathrm{D}\left(\boldsymbol{x}_1 - \boldsymbol{x}_i\right) \right\rangle \\ &= \delta_\mathrm{D}\left(m_2 - m_1\right) \delta_\mathrm{D}\left(\boldsymbol{x}_2 - \boldsymbol{x}_1\right) n(m_1) , \quad (4.8)\end{aligned}$$

where we used the properties of Dirac's delta distribution to rearrange the indices in the second step and Eq. (4.6) to obtain the final result.

The case $i \neq j$ corresponds to correlations between different halos, which leads to the definition of the *halo seed function* $\xi_\mathrm{hh}(m_1, m_2, \boldsymbol{x}_1, \boldsymbol{x}_2)$:

$$\begin{aligned}&\left\langle \sum_{i\neq j} \delta_\mathrm{D}\left(m_1 - m_i\right) \delta_\mathrm{D}\left(\boldsymbol{x}_1 - \boldsymbol{x}_i\right) \delta_\mathrm{D}\left(m_2 - m_j\right) \delta_\mathrm{D}\left(\boldsymbol{x}_2 - \boldsymbol{x}_j\right) \right\rangle \\ &\equiv n(m_1)n(m_2)\left(1 + \xi_\mathrm{hh}(m_1, m_2, \boldsymbol{x}_1, \boldsymbol{x}_2)\right) . \quad (4.9)\end{aligned}$$

If halos are independently distributed, then $\xi_\mathrm{hh} = 0$. Otherwise it measures the excess due to the clustering of halos. An explicit form of the halo seed function can be derived in the frame of a model for halo clustering and is discussed in Sect. 4.4.

By inserting the results of the two cases, i.e. Eq. (4.8) and Eq. (4.9), into Eq. (4.7) and transforming them to the 2PCF of halo density contrasts making use of $\xi(\boldsymbol{r}_a, \boldsymbol{r}_b) = \langle\rho(\boldsymbol{r}_a)\rho(\boldsymbol{r}_b)\rangle/\bar{\rho}^2 - 1$, one obtains

$$\xi_{1\mathrm{h}}(\boldsymbol{r}_a, \boldsymbol{r}_b) = \frac{1}{\bar{\rho}^2}\int \mathrm{d}m_1\,\mathrm{d}^3x_1 m_1^2 n(m_1) u(m_1, \boldsymbol{r}_a - \boldsymbol{x}_1) u(m_1, \boldsymbol{r}_b - \boldsymbol{x}_1)\,, \tag{4.10}$$

$$\begin{aligned}\xi_{2\mathrm{h}}(\boldsymbol{r}_a, \boldsymbol{r}_b) = \frac{1}{\bar{\rho}^2}\int \mathrm{d}m_1\,\mathrm{d}m_2\,\mathrm{d}^3x_1\,\mathrm{d}^3x_2 m_1 m_2 n(m_1) n(m_2) \\ \times u(m_1, \boldsymbol{r}_a - \boldsymbol{x}_1)\, u(m_2, \boldsymbol{r}_b - \boldsymbol{x}_2)\, \xi_{\mathrm{hh}}(m_1, m_2, \boldsymbol{x}_1, \boldsymbol{x}_2)\,.\end{aligned} \tag{4.11}$$

Note that the $1/\bar{\rho}^2$ factor appears due to the change of variables from mass density to density contrast. The 1 of the previous Equation cancels with the first term of Eq. (4.9), which then only contains the seeded halo term. The two equations above are denoted the *1-halo* and the *2-halo term* of the 2PCF, since they are associated with correlations within one halo or between two halos, respectively and sum up to the complete 2PCF:

$$\xi(\boldsymbol{r}_a, \boldsymbol{r}_b) = \xi_{1\mathrm{h}}(\boldsymbol{r}_a, \boldsymbol{r}_b) + \xi_{2\mathrm{h}}(\boldsymbol{r}_a, \boldsymbol{r}_b)\,. \tag{4.12}$$

4.1.3 Power spectrum

It is possible to simplify the derived expressions further by transforming them to Fourier space, resulting in

$$\xi(\boldsymbol{r}_a, \boldsymbol{r}_b) = \int \frac{\mathrm{d}^3k}{(2\pi)^3} P(\boldsymbol{k}) \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{r}_a - \boldsymbol{r}_b)}\,. \tag{4.13}$$

The kernel can be identified with the power spectrum, since $\xi(\boldsymbol{r})$ and $P(\boldsymbol{r})$ are a Fourier pair (see Sect. 3.3). One way to find the expression (4.13) is by starting with the decomposition (4.12) of the 2PCF. We first calculate the 1-halo term of the power spectrum $P_{1\mathrm{h}}$ explicitly, starting from Eq. (4.10):

$$
\begin{aligned}
\xi_{1\mathrm{h}}(\boldsymbol{r}_a, \boldsymbol{r}_b) &= \frac{1}{\bar{\rho}^2} \int \mathrm{d}m\, \mathrm{d}^3x_1 m^2 n(m) u(\boldsymbol{r}_a - \boldsymbol{x}_1) u(\boldsymbol{r}_b - \boldsymbol{x}_1) \\
&= \int \mathrm{d}m \left(\frac{m}{\bar{\rho}}\right)^2 n(m) \int \mathrm{d}^3x\, u(\boldsymbol{x}) u(\boldsymbol{x} - \boldsymbol{r}_a + \boldsymbol{r}_b) \\
&= \int \mathrm{d}m \left(\frac{m}{\bar{\rho}}\right)^2 n(m) \int \mathrm{d}^3r\, \delta_{\mathrm{D}}\left(\boldsymbol{r}_a - \boldsymbol{r}_b - \boldsymbol{r}\right) \\
&\quad \times \int \mathrm{d}^3x\, u(\boldsymbol{x}) u(\boldsymbol{x} - \boldsymbol{r}) \\
&= \int \mathrm{d}m \left(\frac{m}{\bar{\rho}}\right)^2 n(m) \int \mathrm{d}^3r \int \frac{\mathrm{d}^3k}{(2\pi)^3}\, \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{r}_a - \boldsymbol{r}_b - \boldsymbol{r})} \\
&\quad \times \int \mathrm{d}^3x\, u(\boldsymbol{x}) u(\boldsymbol{x} - \boldsymbol{r}) \\
&= \int \mathrm{d}m \left(\frac{m}{\bar{\rho}}\right)^2 n(m) \int \frac{\mathrm{d}^3k}{(2\pi)^3}\, \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{r}_a - \boldsymbol{r}_b)} \\
&\quad \times \int \mathrm{d}^3r\, \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} \int \mathrm{d}^3x\, u(\boldsymbol{x})\, v(\boldsymbol{r} - \boldsymbol{x})\,, \qquad (4.14)
\end{aligned}
$$

where we dropped the mass dependency of the profile and performed a coordinate transformation in the second step. In the final step, we substituted $v(\boldsymbol{r}-\boldsymbol{x}) := u(\boldsymbol{x}-\boldsymbol{r})$ in order to obtain the convolution $(u * v)(\boldsymbol{r}) = \int \mathrm{d}^3x\, u(\boldsymbol{x})\, v(\boldsymbol{r} - \boldsymbol{x})$. The last line of equation (4.14) can now be simplified with the convolution theorem to

$$
\begin{aligned}
\int \mathrm{d}^3r \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} \int \mathrm{d}^3x\, u(\boldsymbol{x})\, v(\boldsymbol{r} - \boldsymbol{x}) &= \int \mathrm{d}^3r\, \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\, (u * v)(\boldsymbol{r}) \\
&= \mathcal{F}(u * v)(\boldsymbol{k}) = \tilde{u}(\boldsymbol{k}, m)\, \tilde{v}(\boldsymbol{k}, m) \\
&= \tilde{u}(\boldsymbol{k}, m)\, \tilde{u}(-\boldsymbol{k}, m) = |\tilde{u}(\boldsymbol{k}, m)|^2\,, \qquad (4.15)
\end{aligned}
$$

where $\mathcal{F}$ and the tilde sign both denote the forward Fourier transform as defined in Sect. 3.3. With the result of Eq. (4.15), we can continue our calculation of the 2PCF and arrive at

$$
\begin{aligned}
\xi(\boldsymbol{r}_a, \boldsymbol{r}_b) &= \int \mathrm{d}m \left(\frac{m}{\bar{\rho}}\right)^2 n(m) \int \frac{\mathrm{d}^3k}{(2\pi)^3}\, \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{r}_a - \boldsymbol{r}_b)} |\tilde{u}(\boldsymbol{k}, m)|^2 \\
&= \int \frac{\mathrm{d}^3k}{(2\pi)^3} \int \mathrm{d}m \left(\frac{m}{\bar{\rho}}\right)^2 n(m) |\tilde{u}(\boldsymbol{k}, m)|^2 \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{r}_a - \boldsymbol{r}_b)}\,. \qquad (4.16)
\end{aligned}
$$

From Eq. (4.13) we can deduce the 1-halo term of the power spectrum to be

$$P_{1\mathrm{h}}(\boldsymbol{k}) = \frac{1}{\bar{\rho}^2} \int \mathrm{d}m\, n(m)\, m^2\, |\tilde{u}(\boldsymbol{k}, m)|^2 \tag{4.17}$$

Analogously, we can derive the Fourier transformation of the 2-halo term, so that we obtain the full power spectrum

$$P(\boldsymbol{k}) = P_{1\mathrm{h}}(\boldsymbol{k}) + P_{2\mathrm{h}}(\boldsymbol{k})\,, \tag{4.18}$$

where equations (4.10) and (4.11) are Fourier transformed to

$$P_{1\mathrm{h}}(\boldsymbol{k}) = \frac{1}{\bar{\rho}^2} \int \mathrm{d}m\, n(m)\, m^2\, |\tilde{u}(\boldsymbol{k}, m)|^2\,, \tag{4.19}$$

$$\begin{aligned} P_{2\mathrm{h}}(\boldsymbol{k}) = \frac{1}{\bar{\rho}^2} \int \mathrm{d}m_1 \int \mathrm{d}m_2\, m_1 n(m_1)\, m_2\, n(m_2) \\ \times\, \tilde{u}^*(\boldsymbol{k}, m_1)\, \tilde{u}(\boldsymbol{k}, m_2)\, P_{\mathrm{hh}}(\boldsymbol{k}, m_1, m_2)\,, \end{aligned} \tag{4.20}$$

where $P_{\mathrm{hh}}(\boldsymbol{k}, m_1, m_2)$ is the Fourier transform of the halo seeds correlation function ξ_{hh} and $\tilde{u}(\boldsymbol{k}, m)$ is the Fourier transform of the halo density profile u.

4.1.4 Bispectrum

In a similar way, one can derive expressions for the 3-point correlation function (3PCF) and its Fourier transformation, the bispectrum. Since working in Fourier space is more convenient, we will only summarize the Fourier expressions. Analogous calculations as in the previous subsections show that the dark matter bispectrum splits into three terms

$$B(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) = B_{1\mathrm{h}}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) + B_{2\mathrm{h}}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) + B_{3\mathrm{h}}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3)\,, \tag{4.21}$$

where each term is given as

$$B_{1\mathrm{h}}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) = \frac{1}{\bar{\rho}^3} \int \mathrm{d}m_1\, m_1^3\, n(m_1) \prod_{i=1}^{3} \tilde{u}(\boldsymbol{k}_i, m_1)\,, \tag{4.22}$$

$$\begin{aligned} B_{2\mathrm{h}}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) = \frac{1}{\bar{\rho}^3} \int \mathrm{d}m_1\, m_1\, n(m_1)\, \tilde{u}(\boldsymbol{k}_1, m_1) \int \mathrm{d}m_2\, m_2^2\, n(m_2) \tilde{u}(\boldsymbol{k}_2, m_2)\,, \\ \times\ \tilde{u}(\boldsymbol{k}_3, m_3)\, P_{\mathrm{hh}}(m_1, m_2, \boldsymbol{k}_1) + 2\,\text{perms.} \end{aligned} \tag{4.23}$$

$$B_{3\mathrm{h}}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) = \frac{1}{\bar{\rho}^3} \left[\prod_{i=1}^{3} \int \mathrm{d}m_i\, m_i\, \tilde{u}(\boldsymbol{k}_i, m_i) \right] B_{\mathrm{hhh}}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3, m_1, m_2, m_3)\,, \tag{4.24}$$

where $B_{\mathrm{hhh}}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3, m_1, m_2, m_3)$ is the Fourier transformed excess for three halos in analogy to P_{hh} and '2 perms.' means that additionally two permutations of the forgoing expression have to be considered. An explicit expression for B_{hhh} will be derived in Sect. 4.4.3.

4.1.5 Trispectrum

Since this thesis deals with higher-order correlation functions and more specifically with the covariance of the power spectrum, we finally want to list the terms of the trispectrum, which is the Fourier transform of the 4PCF (see Sect. 3.5). It can be written as the sum of four terms

$$T(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3, \boldsymbol{k}_4) = T_{1\mathrm{h}} + T_{2\mathrm{h}} + T_{3\mathrm{h}} + T_{4\mathrm{h}}\,, \tag{4.25}$$

where we dropped the $\boldsymbol{k}$-dependency of the halo terms for the sake of readability, and each term is

$$T_{1\mathrm{h}} = \frac{1}{\bar{\rho}^4} \int \mathrm{d}m_1\, m_1^4\, n(m_1) \prod_{i=1}^{4} \tilde{u}(\boldsymbol{k}_i, m_i) \tag{4.26}$$

$$\begin{aligned} T_{2\mathrm{h}} = \frac{1}{\bar{\rho}^4} \Bigg[& \int \mathrm{d}m_1\, m_1\, n(m_1)\, \tilde{u}(\boldsymbol{k}_1, m_1) \int \mathrm{d}m_2\, m_2^3\, n(m_2) \tilde{u}(\boldsymbol{k}_2, m_2) \\ & \times \tilde{u}(\boldsymbol{k}_3, m_3)\, \tilde{u}(\boldsymbol{k}_4, m_4) P_{\mathrm{hh}}(\boldsymbol{k}_1, m_1, m_2) \Bigg] + 3\,\mathrm{perms.} \\ & + \frac{1}{\bar{\rho}^4} \Bigg[\int \mathrm{d}m_1\, m_1^2\, n(m_1) \tilde{u}(\boldsymbol{k}_1, m_1)\, \tilde{u}(\boldsymbol{k}_2, m_2) \int \mathrm{d}m_2\, m_2^2\, n(m_2) \tilde{u}(\boldsymbol{k}_3, m_3)\, \tilde{u}(\boldsymbol{k}_4, m_4) \\ & \times P_{\mathrm{hh}}(|\boldsymbol{k}_1 + \boldsymbol{k}_2|, m_1, m_2) \Bigg] + 2\,\mathrm{perms.} \end{aligned} \tag{4.27}$$

$$\begin{aligned} T_{3\mathrm{h}} = \frac{1}{\bar{\rho}^4} \Bigg[& \int \mathrm{d}m_1\, m_1\, n(m_1)\, \tilde{u}(\boldsymbol{k}_1, m_1) \int \mathrm{d}m_2\, m_2\, n(m_2)\, \tilde{u}(\boldsymbol{k}_2, m_2) \\ & \times \int \mathrm{d}m_3\, m_3^2\, n(m_3)\, \tilde{u}(\boldsymbol{k}_3, m_3) \tilde{u}(\boldsymbol{k}_4, m_4) B_{\mathrm{hhh}}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3 + \boldsymbol{k}_4) \Bigg] + 5\,\mathrm{perms.} \end{aligned} \tag{4.28}$$

$$T_{4\mathrm{h}} = \frac{1}{\bar{\rho}^4} \left[\prod_{i=1}^{4} \int \mathrm{d}m_i\, m_i\, \tilde{u}(\boldsymbol{k}_i, m_i) \right] T_{\mathrm{hhhh}}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3, \boldsymbol{k}_4, m_1, m_2, m_3, m_4) + 3\,\mathrm{perms.}\,, \tag{4.29}$$

where $T_{\mathrm{hhhh}}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3, \boldsymbol{k}_4, m_1, m_2, m_3, m_4,)$ is the Fourier transformed excess for four halos. An explicit expression for this is derived in Sect. 4.4.3.

4.1.6 Polyspectra

Even higher-order terms can be constructed, but their expressions are tedious to obtain and to write down. Instead, a good approximation is given by considering only the 1-halo and the n-halo terms, which have the largest contributions on large and on small scales, respectively:

$$T^n(\boldsymbol{k}_1, \ldots, \boldsymbol{k}_n) \simeq T^n_{1\mathrm{h}} + T^n_{n\mathrm{h}}\,, \tag{4.30}$$

where

$$T^n_{1\mathrm{h}} = \frac{1}{\bar{\rho}^n} \int \mathrm{d}m_1\, m_1^n\, n(m_1) \prod_{i=1}^{n} \tilde{u}(\boldsymbol{k}_i, m_i) \tag{4.31}$$

$$T^n_{n\mathrm{h}} = \frac{1}{\bar{\rho}^n} \left[\prod_{i=1}^{n} \int \mathrm{d}m_i\, m_i\, \tilde{u}(\boldsymbol{k}_i, m_i) \right] T^n_{n\mathrm{h}}(\boldsymbol{k}_1, \boldsymbol{k}_2, \ldots, \boldsymbol{k}_n, m_1, m_2, \ldots, m_n)\,, \tag{4.32}$$

where $T^n_{n\mathrm{h}}(\boldsymbol{k}_1, \boldsymbol{k}_2, \ldots, \boldsymbol{k}_n, m_1, m_2, \ldots, m_n)$ denotes the Fourier transformed excess between all n halos.

4.2 The spherical collapse model

The difficulty of finding an analytic model for the statistics of dark matter halos is that it has to deal with highly non-linear objects. The usual approach is to use numerical simulations in order to describe the formation of halos. However, there exists a special case, in which the non-linear evolution of the density contrast can be solved explicitly. If one considers the evolution of a spherical matter overdensity embedded in a homogeneous Universe, it is possible to calculate the density contrast, at which an object can be considered as virialized. This description is referred to as *spherical collapse model* and was studied first by Gunn and Gott [28] in 1972, who considered the collapse from an initially top-hat density perturbation. In addition to estimating the value for the virial density contrast, the spherical collapse model in the linear regime is used to determine a characteristic time-scale, at which bound objects form. In the following, we will derive the two important quantities of the spherical collapse model, namely the virial density contrast Δ_{vir} and the linear density contrast of a collapsed object δ_{sc}.

We consider the evolution of an overdense spherical perturbation of comoving scale R in a matter-dominated Universe. The evolution of the overdense sphere will occur independently of the rest of the Universe. At some initial time t_i the spherical

perturbation of radius R_i will have the average density

$$\rho(R_i, t_i) = \bar{\rho}\left[1 + \bar{\delta}(R_i, t_i)\right] , \tag{4.33}$$

where $\bar{\delta}(R_i, t_i) = \frac{3}{R_i^3}\int \mathrm{d}R\, R^2 \delta(R, t_i)$ is the average density contrast within the volume V_i and $\bar{\rho}$ is the comoving background density of the Universe. The dynamics of the collapse is governed by the mass M enclosed in the sphere, which is conserved as long as different spheres do not cross each other[1], i.e.

$$M(R, t) = M(R_i, t_i) = \frac{4}{3}\pi R_i^3 \bar{\rho}[1 + \bar{\delta}(R_i, t_i)] . \tag{4.34}$$

If one additionally assumes $\bar{\delta}(R_i) \ll 1$, one finds a simple relation between the initial Lagrangian radius of the sphere R_i and the evolved Eulerian radius R:

$$\left(\frac{R_i}{R}\right)^3 = (1 + \bar{\delta}(R)) . \tag{4.35}$$

Eq. (4.35) will be used, when we want to relate the linear density contrast to the non-linear one and whenever a transformation from Langrangian to Eulerian coordinates is performed. Note that in the following we will denote the density contrast $\bar{\delta}(R)$ simply by δ for notational convenience.

Changing to physical coordinates, the dynamics of a gravitating sphere with radius R and mass M can be described with Newtonian physics and thus follows

$$\frac{\mathrm{d}^2 R}{\mathrm{d}t^2} = -\frac{GM}{R^2} , \tag{4.36}$$

where we restrict ourselves to one dimension due to the spherical symmetry of the problem. The total energy of the sphere determines its actual fate and is found by integrating the above equation:

$$\frac{1}{2}\left(\frac{\mathrm{d}R}{\mathrm{d}t}\right)^2 - \frac{GM}{R} = E \tag{4.37}$$

For $E \geq 0$ the shell will expand forever, whereas for $E < 0$ it will first expand until it reaches a maximal radius $R_{\max}$ and then collapse[2]. Since we are interested in the formation of objects, we restrict our calculation to $E < 0$. In this case, Eq. (4.36) has the following parametric solution:

[1] Within the spherical collapse model one assumes that shell crossing happens only after the actual collapse

[2] This consideration is also valid for the whole Universe and provides a way to derive the Friedmann equations in the Newtonian limit.

$$R = A\,(1 - \cos\theta)\ , \tag{4.38}$$
$$t = B\,(\theta - \sin\theta)\ , \tag{4.39}$$

where $A^3 = GMB^2$ and θ parametrizes the evolution in time. With the parametric solution, it is possible to relate θ to the corresponding radius R and time t for each stage of the evolution. At the beginning of the process, one has $\theta_i = 0$ and $R_i = 0$. At turnaround, where the sphere reaches its maximal radius $R_{\max}$, one has $\theta_{\rm ta} = \pi$, and the sphere is collapsed completely, when $\theta_{\rm coll} = 2\pi$. An overview can be found in tab. 4.1.

4.2.1 Linear regime

In the initial stage, where $\theta \ll 1$, we can expand the parametric solution, i.e. Eqs. (4.38) and (4.39), into a power series of θ:

$$R \approx \frac{A}{2}\,\theta^2 \left(1 - \frac{\theta^2}{12}\right)\ , \tag{4.40}$$
$$t \approx \frac{B}{6}\,\theta^3 \left(1 - \frac{\theta^2}{20}\right)\ . \tag{4.41}$$

Combining Eqs. (4.40) and (4.41) and neglecting higher-order terms, the radius becomes

$$R(t) \approx \frac{A}{2}\left(\frac{6t}{B}\right)^{2/3}\left[1 - \frac{1}{20}\left(\frac{6t}{B}\right)^{2/3}\right]\ , \tag{4.42}$$

whereas the mean density of the perturbation is

$$\rho(t) = \frac{3M}{4\pi R^3} \approx \frac{1}{6\pi G t^2}\left[1 + \frac{3}{20}\left(\frac{6t}{B}\right)^{2/3}\right]\ . \tag{4.43}$$

For an EdS-Universe, the background density of the Universe is

$$\bar{\rho} = \frac{1}{6\pi G t^2}\,, \tag{4.44}$$

thus, from $\rho = \bar{\rho}\,(1 + \delta)$, we can identify

$$\delta_{\rm lin} = \frac{3}{20}\left(\frac{6t}{B}\right)^{2/3} \tag{4.45}$$

Table 4.1: *Three stages of the spherical collapse of a perturbation with their characteristic values for the linear and non-linear density contrast.*

Stage	θ	r	t	$\delta_{\rm lin}$	$\Delta_{\rm nl}$
Initial time	0	0	0	0	∞
Turnaround	π	$2A$	πB	1.06	$9\pi^2/16$
Collapse	2π	0	$2\pi B$	1.686	∞

as the linear density contrast of a spherical object. The scaling $t^{2/3}$ of the density contrast is exactly what one would expect in case of an EdS-Universe. Although the linear solution is a poor approximation for the real density contrast as soon as $\theta \gtrsim 1$ (see Fig. 4.2), its value at collapse can be used to define a characteristic time-scale for virialized objects. This becomes e.g. important when deriving the halo abundance at different redshifts in section 4.3. The exact value for the linear density contrast is obtained by inserting the time of collapse $t_{\rm coll} = 2\pi B$ into Eq. (4.45):

$$\delta_{\rm sc} = \frac{3}{20}\,(12\pi)^{2/3} \approx 1.686\,, \tag{4.46}$$

where we the use the subscript 'sc' in order to indicate the origin from the spherical collapse model.

The redshift dependence of the density contrast can also be incorporated into the consideration. If one assumes e.g. that an object has formed in the recent past, one can set the present time for the collapse epoch and calculate the initial linear density contrast at redshift z with the help of the growth factor according to

$$\delta_{\rm sc}(z) = G(z)\,\delta_{\rm sc}(z_{\rm coll} = 0) = \frac{1.686}{1+z}\,, \tag{4.47}$$

where the last equality is only valid for an EdS-Universe.

4.2.2 Non-linear regime

As stated in the beginning, the evolution of a gravitating sphere can also be solved analytically in the non-linear regime by Eqs. (4.38) and (4.39). In analogy to the linear case, one can use these solutions to determine the average density and density contrast:

$$\rho_{\rm nl}(t) = \frac{3M}{4\pi R_{\rm nl}^3(t)} = \frac{3}{4\pi G t^2}\frac{(\theta-\sin\theta)^2}{(1-\cos\theta)^3}\,, \tag{4.48}$$

$$\delta_{\rm nl}(t) = \bar{\rho}[1+\delta_{\rm nl}(t)] = \frac{9}{2}\frac{(\theta-\sin\theta)^2}{(1-\cos\theta)^3} - 1\,, \tag{4.49}$$

where $\bar{\rho}$ denotes the background density in an EdS cosmology as given in Eq. (4.44). We can now reconsider the most important stages of a collapsing sphere and focus on how the non-linear density contrast evolves with respect to the background density of the Universe. At first the spherical perturbation expands from a zero radius until it reaches a maximum at $R_{\rm ta}$. The turnaround marks the point of time, where the actual collapse begins. At this moment, the ratio of the density of the sphere to that of the background density is

$$\Delta_{\rm ta} \equiv \frac{\rho(t_{\rm ta})}{\bar{\rho}(t_{\rm ta})} = \frac{9\pi^2}{16} \approx 5.552\,, \tag{4.50}$$

which can be derived using the non-linear density as given in Eq. (4.48). For the time of the actual collapse, non-linear theory predicts the sphere to have a radius $R_{\rm coll} = 0$ and an infinite density contrast, whereas the linear consideration approximates $\delta_{\rm sc} \simeq 1.686$. In reality the spherical perturbation will not become singular, since this occurs only if the collapse is exactly symmetric. Instead the perturbation virializes in a process of violent relaxation to a finite radius. With the help of the virial theorem, one finds the radius of the relaxed state to be $R_{\rm vir} = R_{\rm ta}/2$. Thus, the sphere is in a virialized state if the radius reaches half of the maximum expansion at turnaround, which results in a density

$$\rho(t_{\rm vir}) = \frac{3M}{4\pi G R_{\rm vir}^3} = 8\,\frac{3M}{4\pi G R_{\rm ta}^3} = 8\,\rho(t_{\rm ta}). \tag{4.51}$$

Numerical simulations of the process suggest instead that the relaxation occurs at about the time of collapse. Assuming this time-scale, one finds

$$\bar{\rho}(t_{\rm vir}) = (6\pi G t_{\rm coll}^2)^{-1} = \bar{\rho}(t_{\rm ta})/4\,. \tag{4.52}$$

Combing these last two equations determined for the time of virialization with the result in Eq. (4.50), one finds the virial density contrast to be

$$\Delta_{\rm vir} \equiv \delta_{\rm vir} + 1 = 32\,\frac{\rho(t_{\rm ta})}{\bar{\rho}(t_{\rm ta})} = 18\pi^2 \approx 177.7 \tag{4.53}$$

in the non-linear regime. Note that this consideration is only valid for an EdS-Universe and the collapse of spherical objects. A more realistic treatment would assume an elliptical mass distribution of gravitating perturbation instead. Consequences of an elliptical collapse are discussed in detail e.g. in [8, 53, 78]. Fitting formulae for other cosmologies can be found in [23, 31, 55, 60].

4.2.3 Relating the linear to the non-linear density

The advantage of the spherical collapse model is that it provides a relation between the initial, linear density contrast δ_i and the evolved, non-linear density contrast δ. In order to derive this relation, one has to rewrite the original parametric solution of into

$$\frac{R(z)}{R_i} = \frac{3}{10}\frac{1-\cos\theta}{|\delta_i|}\,,$$
$$\frac{1}{1+z} = \frac{3\times 6^{2/3}}{20}\frac{(\theta-\sin\theta)^{2/3}}{|\delta_i|}\,,$$

as has been done by Peebles [65]. This solution holds for $\delta_i > 0$. For $\delta_i < 0$ one has to substitute the trigonometric functions by their hyperbolic counterparts. Without loss of generality one can now choose $z = 0$ as the point of time of our consideration. In this case, the initial density contrast δ_i depends only on the present density contrast, since $\delta = (R/R_i)^3 - 1$ holds, as we have shown earlier in this section. For $\delta \ll 1$, we can expand $\delta_i(\delta)$ into a power series

$$\delta_i = \sum_{k=0}^{\infty} a_k \delta^k = \delta - \frac{17}{21}\delta^2 + \frac{341}{567}\delta^3 - \frac{55805}{130977}\delta^4 + \dots\,, \tag{4.54}$$

where $a_0 = 0$. A derivation of the inverse power series $\delta(\delta_i)$ can be found in [68].

4.3 Halo mass function

The halo model approach assumes that all dark matter is distributed in spherical halos. An essential ingredient to understand the statistics of halos is therefore the *halo mass function*, which specifies how frequent halos of a certain mass are. More precisely, the halo mass function, which we denote in the following by[3]

$$n(m,z) \equiv \frac{\mathrm{d}N(m,z)}{\mathrm{d}m}\,, \tag{4.55}$$

is defined as the comoving number density of halos with a mass between m and $m+\mathrm{d}m$, which are collapsed at a redshift z. The original formalism for calculating the mass function was provided by Press and Schechter [66] in 1974 and will be summarized in the following, since it is the basis for most of the recent considerations of halo clustering.

[3]Note that in the literature $n(m,z)$ and $\mathrm{d}n(m,z)/\mathrm{d}m$ are often used interchangeably.

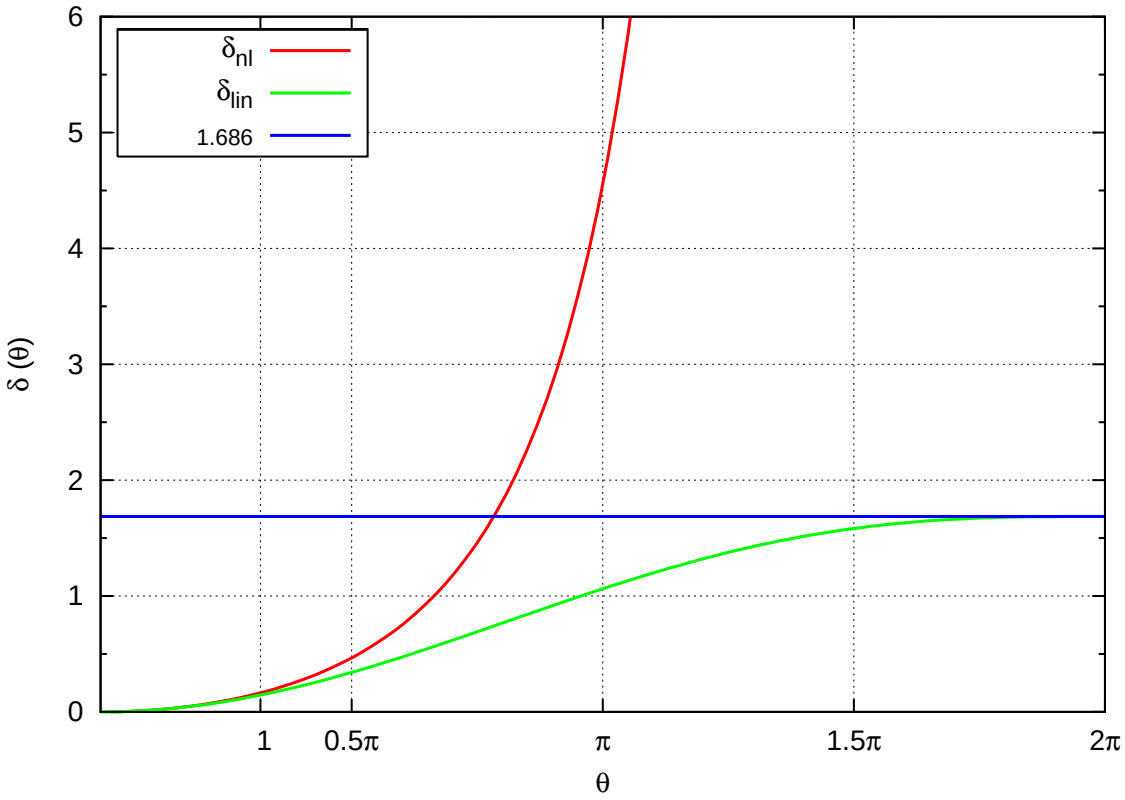

Figure 4.2: *Comparison of the linear with the non-linear density contrast. Around $\theta \approx 1$ the linear approximation breaks down. The blue line marks the threshold $\delta_{\rm sc} = 1.686$, when an object is considered as collapsed in the linear regime.*

4.3.1 Press-Schechter formalism

Let us consider some early time t_i at which the density field $\delta(\boldsymbol{x}, t_i)$ can still be described as a Gaussian random field. The smoothed density field $\delta_R(\boldsymbol{x}, t_i)$ as defined in Eq. (3.47) follows then a Gaussian distribution as well, since a convolution is a linear operation. Accordingly, the probability density of observing a value of $\delta_R(\boldsymbol{x}, t_i)$ between δ_R and $\delta_R + \mathrm{d}\delta_R$ is given by a Gaussian distribution:

$$P(\delta_R, t_i; R)\,\mathrm{d}\delta_R = \frac{1}{\sqrt{2\pi\sigma^2(m, t_i)}} \exp\left[-\delta_R^2 / 2\sigma^2(m, t_i)\right] \mathrm{d}\delta_R\,, \tag{4.56}$$

where $\sigma^2(m, t_i) = \langle \delta_R^2(\boldsymbol{x}, t_i) \rangle$ is the variance of the initial density field smoothed with a filter function $W_R(\boldsymbol{k})$ as defined in Eq. (3.53). The comoving smoothing scale R is determined by the mass enclosed in the volume of the filter function. In case of a top-hat filter it is related to the mass through $m = (4\pi/3)\,\bar{\rho} R^3$. From these equations, one can see that R, $\sigma^2(m, t_i)$ and m are equivalent variables for specifying an object or region in a given cosmology.

According to the Press and Schechter (PS) formalism, an object of some mass m will collapse if the smoothed density contrast exceeds a threshold δ_*. Consequently, the cumulative probability for a region of scale R to have a smoothed density contrast above the threshold is

$$F(m) = \int_{\delta_*}^{\infty} P(\delta_R, t_i; R)\, \mathrm{d}\delta_R = \frac{1}{2}\,\mathrm{erfc}(\nu/\sqrt{2})\,, \tag{4.57}$$

where erfc(x) is the complementary error function and $\nu \equiv \delta_*/\sigma(m)$ the height of the threshold in units of the smoothed variance. In other words, the fraction of virialized objects today with a mass larger than m is $F(m)$.

In order to determine the threshold δ_* for objects to be virialized, PS use the results from the spherical collapse model in the linear regime (see Sect. 4.2) and set $\delta_* = \delta_{\mathrm{sc}}$. For an EdS-Universe, we calculated $\delta_{\mathrm{sc}} = 1.686$. Within the frame of this specific cosmology, it is now possible to find the characteristic scale collapsing today. It can be determined by setting $\sigma(m_*) = \delta_{\mathrm{sc}}$, which yields $m_* \approx 10^{13} M_\odot$. This is in agreement with the observation that objects with masses above this threshold are rare.

Calculating $F(0)$ in Eq. (4.57) should give the fraction of all mass that is contained in virialized objects. Since we consider a hierarchical model of structure formation, implying that $\sigma(m) \to \infty$ if $R \to 0$, this fraction is $F(0) = \frac{1}{2}\,\mathrm{erfc}(0) = \frac{1}{2}$. Astonishingly, this means that, in the framework of the considered formalism, only half of the mass in the Universe is in virialized objects. This led Press and Schechter to introduce in the literature the famous *fudge factor* of 2.

4.3.2 Extended Press-Schechter theory

Subsequent studies revealed that the problem of the PS approach is that the correspondence between smoothing scale and threshold height is not unique. In other words, a region below the threshold δ_* on a particular scale, can be above the threshold on a larger scale. Including this effect would increase the fraction of virialized objects in the Universe. In the literature, it was shown independently by Peacock and Heavens [63] and Bond et al. [9] that this effect compensates the original discrepancy. They derived the factor of 2 formally by computing the largest value of the smoothing scale R at which the threshold is exceeded.

If we continue with the extra factor of 2 in Eq. (4.57), we can finally derive an expression for the halo mass function. The fraction of halos with a mass m between m and

$m + \mathrm{d}m$ is given by

$$\begin{aligned}\frac{\mathrm{d}F}{\mathrm{d}m} &= \frac{1}{\sqrt{2\pi\sigma^2}}\,\frac{\delta_{\rm sc}}{\sigma^2}\left|\frac{\mathrm{d}\sigma^2}{\mathrm{d}m}\right|\exp(-\delta_{\rm sc}^2/2\sigma^2)\\ &= \sqrt{\frac{2}{\pi}}\,\frac{\delta_{\rm sc}}{\sigma^2}\left|\frac{\mathrm{d}\sigma}{\mathrm{d}m}\right|\exp(-\delta_{\rm sc}^2/2\sigma^2)\\ &= \sqrt{\frac{2}{\pi}}\,\frac{\delta_{\rm sc}}{\sigma}\left|\frac{\mathrm{d}\ln\sigma}{\mathrm{d}m}\right|\exp(-\delta_{\rm sc}^2/2\sigma^2)\\ &= \sqrt{\frac{2}{\pi}}\,\nu\frac{\mathrm{d}\ln\nu}{\mathrm{d}m}\exp(-\nu^2/2)\,,\end{aligned} \tag{4.58}$$

where we made use of $\mathrm{d}\ln\sigma = -\mathrm{d}\ln\nu$ in the last step and omitted the dependencies of the variables to focus on the calculation. Formally, we can now define the comoving number density of halos through the fraction of halos with masses between m and $m + \mathrm{d}m$, which we just calculated, to be

$$\frac{\mathrm{d}N}{\mathrm{d}m} \equiv \frac{\bar{\rho}}{m}\left|\frac{\mathrm{d}F(m)}{\mathrm{d}m}\right|, \tag{4.59}$$

where $\bar{\rho}/m = V^{-1}$ is the inverse of the comoving volume associated with the mass inside the comoving smoothing scale R. Analytic mass functions are often compared to numerical simulations by considering the fraction of collapsed objects of mass m per logarithmic interval in ν. In this way they can be compared without referring to a specific initial power spectrum or cosmological model. For the cumulative probability found by Press and Schechter this is

$$\nu f(\nu) \equiv \frac{\mathrm{d}F}{\mathrm{d}\ln\nu} = \sqrt{\frac{2}{\pi}}\,\nu\exp(-\nu^2/2)\,. \tag{4.60}$$

For small masses, i.e. $m \ll m_*$, the mass function behaves as $\nu f(\nu) \propto \nu$, whereas for $m \gtrsim m_*$ it is exponentially cut-off with $\nu f(\nu) \propto \exp(-\nu^2/2)$.

4.3.3 Sheth-Tormen mass function

Numerical simulations have shown [77] that the PS mass function (with factor 2), overestimates the number density of halos with high masses and underestimates the abundance of halos with low masses, as illustrated in Fig. 4.3. As an alternative, Sheth and Tormen (ST) proposed a mass function of the form

$$\nu f(\nu) = A(p)\left(1 + (q\nu^2)^{-p}\right)\sqrt{\frac{2q}{\pi}}\,\nu\exp\left(-q\nu^2/2\right) \tag{4.61}$$

where the parameters that fit numerical simulations best are $p \approx 0.3$ and $q = 0.707$. The constant

$$A(p) = \left[1 + 2^{-p}\Gamma(1/2 - p)/\sqrt{\pi}\right]^{-1} \tag{4.62}$$

is defined such that it normalizes

$$\int_0^\infty f(\nu)\,\mathrm{d}\nu = \frac{1}{\bar{\rho}} \int_0^\infty \frac{\mathrm{d}n}{\mathrm{d}m} m\,\mathrm{d}m = 1, \tag{4.63}$$

which, in case of the best fit parameters, yields $A(0.3) \approx 0.322$. If the parameters in Eq. (4.61) are set to $p = 0$, $q = 1$ and $A(0) = 1/2$, one obtains the standard Press-Schechter form of equation (4.60). For $\nu \ll 1$, equation (4.61) scales as $\nu f(\nu) \propto \nu^{1-2p}$, whereas for $\nu \gtrsim 1$ it has an exponential cut-off, i.e. $\nu f(\nu) \propto \exp\left(-q\nu^2/2\right)$.

In a later work, it was possible to derive the ST mass function from ellipsoidal collapse in the excursion set approach (see Sheth et al. [78]). In contrast to the original formulation, the threshold height is then not constant, but depends on the ellipticity distribution of the halos. There are several other fitting formulae for $f(\nu)$ in the literature available (Jenkins et al. [37], Reed et al. [67]), but due to the good correspondence to numerical simulations we will use the ST mass function for our implementation of the halo model.

4.3.4 Redshift dependency of the mass function

So far, we considered the comoving mass function only at some early time t_i, where the density contrast can be considered as a Gaussian random field. Perturbation theory allows us to extrapolate the initial density contrast and its smoothed variance to their values today

$$\sigma^2(m, t_0) = D_+^2(t_0)\sigma^2(m, t_i) \tag{4.64}$$

$$\delta(t_0) = D_+(t_0)\delta(t_i)\,, \tag{4.65}$$

where $D_+(t_0)$ denotes the linear growth factor as defined in Sect. 2.3. This means that δ and σ grow in the same manner and that the mass functions keep the same form in linear perturbation. To determine the mass function of halos that already collapsed at a redshift z, we have to replace the threshold δ_{sc} by

$$\delta'_{\mathrm{sc}} = \frac{\delta_{\mathrm{sc}}}{D(z)} = \delta_{\mathrm{sc}}(1 + z)\,, \tag{4.66}$$

where $\delta_{\mathrm{sc}} = 1.686$. Note that the second equality is only valid for an EdS-Universe. The difference to Eq. (4.47) is the redshift at which an object is considered to be collapsed. Whereas in Eq. (4.47), the time of collapse is today, in Eq. (4.66) the object already virialized at redshift z.

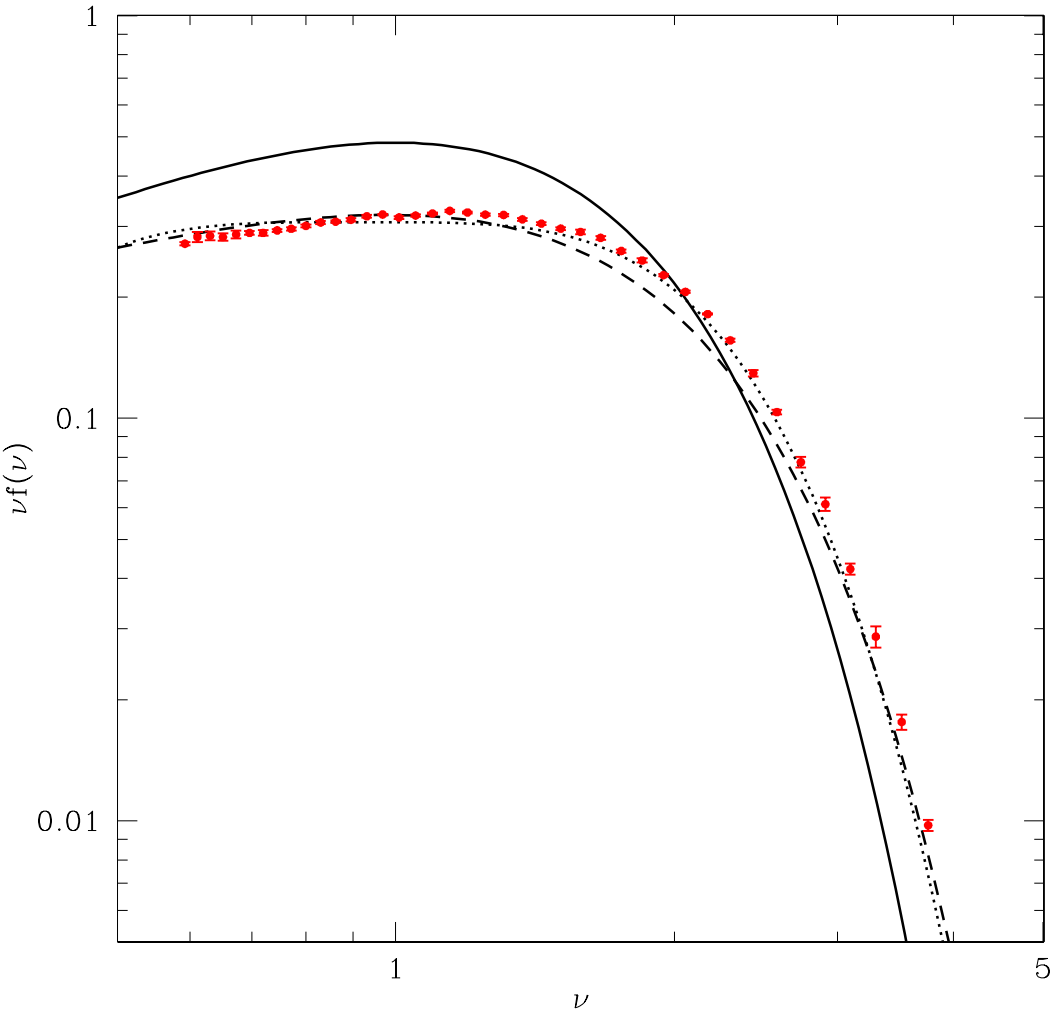

Figure 4.3: *Fraction of collapsed halos per logarithmic mass interval in ν (plot from Zentner [92]). The solid line represents the extended Press-Schechter mass function as given in Eq.* (4.60)*. The dashed line corresponds to the Sheth-Tormen mass function defined in Eq.* (4.61)*, whereas the dotted line corresponds to the Jenkins fitting formula [37]. The points illustrate N-body simulations from J. L. Tinker.*

4.4 Halo bias

In order to describe the large-scale behavior of halos, we need to understand the way halos are correlated. For this reason, we have to determine how the clustering properties of dark matter halos differ from that of dark matter. The different behavior in clustering is usually assumed to be of the form

$$\delta_{\mathrm{h}}(m, z) = b(m, z)\, \delta(m, z)\,. \tag{4.67}$$

where δ_{h} denotes the *halo density contrast*, δ the matter density contrast and $b(m, z)$ the *linear bias*. In the following we omit the dependencies where no confusion arises. If $b = 1$, halos are said to be exact or *unbiased* tracers of the underlying dark matter. In contrast to finding the average abundance of halos in the previous section, we now want to determine how the mass function changes, if the halos are embedded in an over- or underdense region. A powerful tool for modeling the clustering process is provided by the excursion set theory (see Zentner [92] for a review), which is an extension of the Press-Schechter formalism (see Sect. 4.3) for the halo mass function. The first to calculate the amount of bias within the framework of the excursion set theory were Mo and White in 1996 [51]. Before outlining their derivation, we will first discuss the *peak-background split* approach (see [15, 21, 40]), which provides a more intuitive understanding of the halo clustering problem and leads to a similar result as Mo and White's consideration.

4.4.1 Peak-background split

Let us suppose that one can split the density field into two independent random fields $\delta = \delta_b + \delta_p$. The decomposition is constructed such that δ_b corresponds to a smooth background density field, which describes a long wavelength component; and δ_p corresponds to a short wavelength component, namely the peaks, which will eventually form halos[4]. If we consider the superposition of the two density fields, as sketched in Figure 4.4, one can see that halos form more easily, when they are close to peaks of the background density component. In this case the halos are more biased with respect to the underlying dark matter. Analogously, they are less biased, if one is close to a region in the trough of the background field. A simple way to account for this effect is to shift the density threshold required to form an object according to the background density to $\delta_* = \delta_{\mathrm{sc}} - \delta_b$. The peak density δ_p has now to reach the new threshold δ_* to collapse. Accordingly, the threshold height ν changes then to

$$\nu_b = \frac{\delta_{\mathrm{sc}} - \delta_b}{\sigma(m)}\,, \tag{4.68}$$

[4]Note that background density refers in this case to the long wavelength component of the considered region and not to the background density of the Universe.

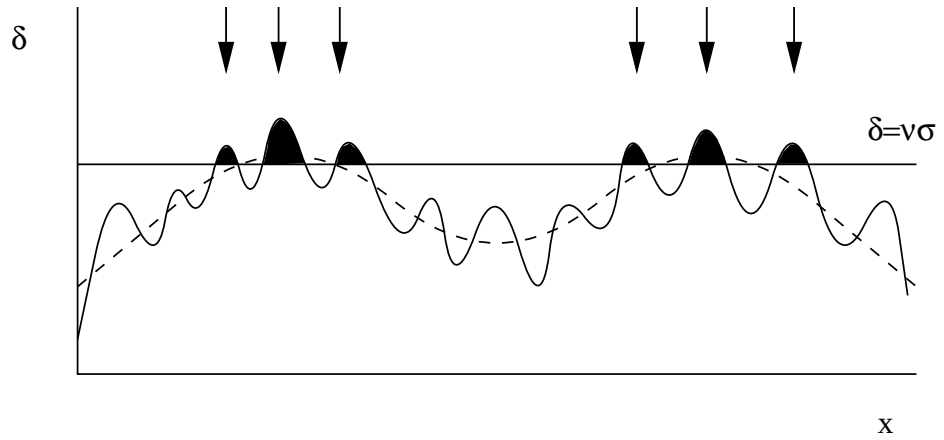

Figure 4.4: *Illustration of the biasing in the peak-background-split model (from [61]). The sketch shows the decomposition of the density field into a long wavelength component δ_b (dashed line) and a short wavelength component δ_p (black, oscillating line). The straight, solid line is the threshold δ_* a dense region has to exceed in order to form a virialized halo.*

where $\delta_{\rm sc}$ denotes the density contrast of a collapsed halo derived from the spherical collapse model. The simplest form one can assume for the new mass function, is then the Press-Schechter form, i.e.

$$n(\delta_b) \equiv \frac{{\rm d}N}{{\rm d}m} \propto \nu_b \exp(-\nu_b^2/2)\,, \tag{4.69}$$

with a threshold altered according to Eq. (4.68). We can now derive an expression for the bias by Taylor expanding the mass function to first order in δ_b around the unaltered threshold:

$$n(\delta_b) \approx n(0) + \frac{{\rm d}n(\delta_b)}{{\rm d}\delta_b}\Big|_{\delta_b=0} \delta_b\,. \tag{4.70}$$

It makes sense to perform first a derivation with respect to ν_b:

$$\frac{{\rm d}n(\delta_b)}{{\rm d}\nu_b} = \left(\frac{1}{\nu_b} - \nu_b\right) n(\delta_b) \tag{4.71}$$

and use then ${\rm d}\nu_b/{\rm d}\delta_b = -1/\sigma(m)$ to obtain

$$\frac{{\rm d}n(\delta_b)}{{\rm d}\delta_b} = \frac{{\rm d}n(\delta_b)}{{\rm d}\nu_b}\,\frac{{\rm d}\nu_b}{{\rm d}\delta_b} = \left(\frac{\nu_b^2 - 1}{\nu_b\,\sigma(m)}\right)\, n(\delta_b)\,. \tag{4.72}$$

If we insert this expression into Eq. (4.70), we find

$$n(\delta_b) \approx \left(1 + \frac{\nu^2 - 1}{\delta_{\rm sc}}\,\delta_b\right) n(0)\,, \tag{4.73}$$

where $n(\delta_b = 0)$ can be identified as the average mass function of halos. Making use of

$$\delta_{\rm h} = \frac{n(\delta_b) - n(0)}{n(0)} \tag{4.74}$$

the bias of halos relative to the underlying dark matter is

$$b^{\rm L}(m) = \frac{\delta_{\rm h}}{\delta_b} = \frac{\nu^2 - 1}{\delta_{\rm sc}}\,, \tag{4.75}$$

where we are still in the initial Langrangian space, since the background density contrast stayed constant. As we will show in detail later (see Sect. 4.4.2) there is a simple relation for transforming the bias to real, Eulerian space

$$b^{\rm E}(m) = 1 + b^{\rm L}(m) = 1 + \frac{\nu^2 - 1}{\delta_{\rm sc}}\,. \tag{4.76}$$

This means in particular that halos with a characteristic mass m_*, are unbiased tracers, since $\nu(m_*) = 1$. Accordingly, if halos have a mass $m > m_*$ they are biased, i.e. $b > 1$, with respect to the underlying dark matter, whereas if halos have a mass $m < m_*$ they are anti-biased, i.e. $b < 1$.

4.4.2 Results from excursion set theory

Conditional mass function

A more thorough derivation of the halo bias is offered by Mo and White [51] and is reproduced very detailed in the review of Zentner [92]. In order to describe the clustering in an overdense regions, Mo and White needed a formula which relates halo abundances to the underlying density field on large scales. For their consideration, they were able to make use of the results by Bower [11] and Bond et al. [9]. These authors extended the Press-Schechter theory (see subsection 4.3.1) to find the desired *conditional mass function*, which gives the number density of halos assuming an overdense background density. In the following, we will briefly summarize the most important results.

Let us consider a region of comoving scale R_1 at a redshift z_1 characterized by a smoothed density contrast $\delta_{R_1}(\boldsymbol{x})$ and a smoothed variance $\sigma^2(R_1)$. This region is

enclosed in a larger region of comoving scale R_0 defined by $\delta_{R_0}(\boldsymbol{x})$ and $\sigma^2(R_0)$. If we use a top-hat window as filter function, the average mass contained in a scale R is $m(R) = (4\pi/3)\bar{\rho}R^3 = V\bar{\rho}$, which implies for the considered regions that $m_1 < m_0$. Assuming that the density contrasts of these regions were initially Gaussian distributed, Bower and Bond et al. [9, 11] extended the results from the Press-Schechter formalism and found the conditional probability of attaining a density δ between δ and $\delta + \mathrm{d}\delta$ given a background density δ_0 to be

$$P(\delta|\delta_0)\,\mathrm{d}\delta = \frac{1}{\sqrt{2\pi(\sigma^2 - \sigma_0^2)}}\,\exp\left[-\frac{(\delta-\delta_0)^2}{2(\sigma^2-\sigma_0^2)}\right]\mathrm{d}\delta\,, \tag{4.77}$$

where we use here and in the following $\delta_i \equiv \delta_{R_i}(\boldsymbol{x})$ and $\sigma_i \equiv \sigma(R_i)$ to shorten the notation. As before, we assume that an object is collapsed, if its density contrast exceeds a threshold δ_{sc}. The fraction of mass in bound halos contained in a region of mass m_0 is then obtained by integrating the conditional probability density according to

$$F(m_1|\delta_0,\sigma_0) = \int_{\delta_{\mathrm{sc}}}^{\infty} P(\delta|\delta_0)\mathrm{d}\delta = \frac{1}{2}\,\mathrm{erfc}\left[\frac{\delta_{\mathrm{sc}}-\delta_0}{\sqrt{2(\sigma_1^2-\sigma_0^2)}}\right]\,, \tag{4.78}$$

where $\mathrm{erfc}(x)$ denotes the complementary error function. If the density of the surrounding region increases, the fraction F of masses in bound halos increases as well. In the case that $\delta_0 \to \delta_{\mathrm{sc}}$, one has $F \to 1$ and the entire region will be interpreted as one collapsed object. The fraction of mass in halos with mass in the range m_1 and $m_1 + \mathrm{d}m_1$ is then

$$\frac{\mathrm{d}F(m_1|\delta_0,\sigma_0)}{\mathrm{d}m_1}\,\mathrm{d}m_1 = \frac{1}{\sqrt{2\pi}}\frac{\delta_{\mathrm{sc}}-\delta_0}{(\sigma_1^2-\sigma_0^2)^{3/2}}\left|\frac{\mathrm{d}\sigma_1^2}{\mathrm{d}m_1}\right|\,\exp\left[-\frac{(\delta_{\mathrm{sc}}-\delta_0)^2}{2(\sigma_1^2-\sigma_0^2)}\right]\mathrm{d}m_1\,. \tag{4.79}$$

In the limit $\delta_0 \to 0$, when the surrounding region has the same smoothed density as the Universe, we obtain the unconditional solution from Eq. (4.58), remembering that $\sigma_1 > \sigma_0$. According to Mo et al. [52] the average comoving number density of halos in a region of mass m_0 can then be defined as

$$\begin{aligned} n(m_1|\delta_0,\sigma_0)\mathrm{d}m_1 &\equiv \frac{1}{V_0}\,\frac{m_0}{m_1}\,\frac{\mathrm{d}F(m_1|\delta_0,\sigma_0)}{\mathrm{d}m_1}\,\mathrm{d}m_1\,, \\ &\equiv \frac{N(\delta_1|\delta_0,\sigma_0)}{V_0}\,\mathrm{d}m_1\,. \end{aligned} \tag{4.80}$$

This expression is interpreted as the *conditional mass function* in a dense region with smoothed density contrast δ_0 for masses in a range m_1 and $m_1 + \mathrm{d}m_1$ at a redshift z_1,

while $N(\delta_1|\delta_0,\sigma_0)$ denotes the *average number* of halos with masses between m_1 and $m_1 + \mathrm{d}m_1$.

Transformation from Lagrangian to Eulerian space

With the conditional mass function we define the halo density contrast as

$$\delta_\mathrm{h}(m_1) \equiv \frac{n(m_1|\delta_0,\sigma_0) - n(m_1)}{n(m_1)} = \frac{n(m_1|\delta_0,\sigma_0)}{n(m_1)} - 1\,, \tag{4.81}$$

where $n(m)$ denotes the unconditional mass function as defined in Eq. (4.59) that corresponds to the average number density of halos. Up to now, we used only comoving coordinates for our consideration and ignored the dynamical evolution of the region R_0. To calculate the Eulerian bias relation, we need an expression for the halo density contrast in terms of the gravitationally evolved region R. The corresponding conditional mass function is then defined as

$$n^\mathrm{E}(m_1|\delta_0,\sigma_0)\mathrm{d}m_1 \equiv \frac{N(\delta_1|\delta_0,\sigma_0)}{V}\,\mathrm{d}m_1\,, \tag{4.82}$$

where we distinguish between Eulerian and comoving Lagrangian conditional mass function by introducing the superscripts 'E' and 'L'. In order to proceed, we need a relation between the initial region of scale R_0 and the gravitationally collapsed region R. This is provided by the spherical collapse model (see Sect. 4.2). Since the mass within a region about to collapse is assumed to be conserved in this model, the evolved volume of the region V is related to its initial volume V_0 according to $V_0 \simeq V(1+\delta)$ for small initial overdensities. Additionally, the spherical collapse model provides a relation between the initial and evolved matter density of the form $\delta = \delta(\delta_0)$. In the limit of small initial overdensities it yields $\delta \simeq \delta_0$. With these relations at hand, we can transform the desired quantities from Lagrangian to Eulerian space. Starting with the Eulerian conditional mass function of Eq. (4.82), we find

$$\begin{aligned} n^\mathrm{E}(m_1|\delta_0,\sigma_0)\mathrm{d}m_1 &= \frac{N(\delta_1|\delta_0,\sigma_0)}{V}\mathrm{d}m_1 \\ &\simeq \frac{N(\delta_1|\delta_0,\sigma_0)}{V_0}\,(1+\delta)\,\mathrm{d}m_1 \\ &= n^\mathrm{L}(m_1|\delta_0,m_0)\,(1+\delta)\,\mathrm{d}m_1\,, \end{aligned} \tag{4.83}$$

where $n^\mathrm{L}(m_1|\delta_0,m_0)$ denotes the conditional mass function as defined in Eq. (4.79).

Making use of this, the Eulerian density contrast transforms into

$$\begin{aligned}\delta_{\mathrm{h}}^{\mathrm{E}} &= \frac{n^{\mathrm{E}}(m_1|\delta_0,\sigma_0)}{n(m_1)} - 1\,,\\ &= \frac{n^{\mathrm{L}}(m_1|\delta_0,m_0)}{n(m_1)}\,(1+\delta) - 1\\ &= (\delta_{\mathrm{h}}^{\mathrm{L}}+1)(1+\delta) - 1\\ &= \delta_{\mathrm{h}}^{\mathrm{L}}(1+\delta) + \delta\,.\end{aligned}$$

Remembering the bias relation $\delta_{\mathrm{h}}^{\mathrm{L}} = b^{\mathrm{L}}\delta_0$ and that $\delta \simeq \delta_0$ on small scales, we obtain

$$\delta_{\mathrm{h}}^{\mathrm{L}} \simeq (b^{\mathrm{L}}+1)\delta_0\,, \tag{4.84}$$

where we only kept terms of first-order in δ and δ_0. Thus, Lagrangian and Eulerian bias are related to each other according to $b^{\mathrm{E}} = b^{\mathrm{L}} + 1$ for small density contrasts.

Press-Schechter bias

The introduced quantities allow us to calculate the bias relation for the Press-Schechter mass function in a perturbative approach. Following Mo, Jing, and White [52], we restrict the consideration to large scales in Eulerian space and assume that the halo density can be described by a smooth function $\mathcal{F}(\delta)$ that depends only on the matter density. If the function is finite for δ around 0, we can expand $\mathcal{F}$ in a Taylor series around δ, such that

$$\delta_{\mathrm{h}} = F(\delta) = \sum_{k=0}^{\infty} \frac{b_k}{k!}\,\delta^k\,, \tag{4.85}$$

where b_k are the bias parameters and δ the non-linear matter density. The bias relation from Eq. (4.67) corresponds then to the first-order approximation of $\mathcal{F}$ and the linear bias we introduced there is now equivalent to b_1. Since we remain during the subsequent considerations in Eulerian space, we omit the superscript 'E' here and in the following.

In order to find explicit expressions for the bias parameters b_k, we expand in the definition of the halo density (Eq. 4.81), the conditional mass function around the matter density δ_0 of the surrounding region and compare equal orders in δ_0 to the Taylor expansion of $\mathcal{F}(\delta)$. Since δ denotes the non-linear density, we additionally have to use the expansion $\delta(\delta_0)$ as found from the spherical collapse model.

To determine the first order or linear bias b_1, we have to calculate the first derivative

of the conditional mass function with respect to the density, since

$$\begin{aligned}\delta_\mathrm{h} &\simeq \frac{1}{n(m_1)}\frac{\mathrm{d}n(m_1|\delta_0,\sigma_0)}{\mathrm{d}\delta_0}\Big|_{\delta_0=0}\delta_0\,(1+\delta)-1 \\ &\simeq \frac{1}{n(m_1)}\frac{\mathrm{d}n(m_1|\delta_0,\sigma_0)}{\mathrm{d}\delta_0}\Big|_{\delta_0=0}\delta_0 \qquad (4.86)\\ &= b_1\delta_0\,. \qquad (4.87)\end{aligned}$$

Combining Eq. (4.80) and (4.79), we obtain for the conditional Press Schechter mass function on large scales

$$\begin{aligned}n(m_1|\delta_0,\sigma_0) &= \frac{\bar{\rho}}{m_1}\,(1+\delta)\,\frac{\mathrm{d}F(m_1|\delta_0,m_0)}{\mathrm{d}m_1}\,\mathrm{d}m_1 \qquad (4.88)\\ &\simeq \frac{\bar{\rho}}{m_1}\,(1+\delta)\,\frac{1}{\sqrt{2\pi}}\frac{\delta_1-\delta_0}{\sigma_1^3}\left|\frac{\mathrm{d}\sigma_1^2}{\mathrm{d}m_1}\right|\exp\left[-\frac{(\delta_1-\delta_0)^2}{2\sigma_1^2}\right]\mathrm{d}m_1\,, \qquad (4.89)\end{aligned}$$

where we assume $R_0 \gg R_1$ and thus have $\sigma_1^2 \gg \sigma_0^2$ in hierarchical structure formation. From this we can calculate the first derivative of the conditional mass function with respect to δ at the background density which gives two terms:

$$\frac{\mathrm{d}n(m_1|\delta_0,\sigma_0)}{\mathrm{d}\delta_0} = \frac{\bar{\rho}}{m_1}\,\frac{1}{\sqrt{2\pi}}\,\frac{\mathrm{d}\sigma_1^2}{\mathrm{d}m_1}\,\frac{1}{\sigma_1^3}\left[\frac{(\delta_1-\delta_0)^2}{\sigma_1^2}-1\right]\exp\left[-\frac{(\delta_1-\delta_0)^2}{2\sigma_1^2}\right]. \qquad (4.90)$$

Inserting this result into Eq. (4.86) and evaluating the found expression, we obtain finally the linear Eulerian bias

$$b(\nu) = 1+\frac{\nu_1^2-1}{\delta_\mathrm{sc}} \qquad (4.91)$$

in the extended Press-Schechter formalism, where $\nu_1 \equiv \delta_1/\sigma_1$. This result is equivalent to equation (4.76) that we obtained in the peak-background split consideration. Higher-orders are found in a similar way, but require tedious calculations, since the higher-order derivatives in the mass function have more terms and one additionally has to apply the expansion $\delta_0(\delta) = \sum_k a_k\delta^k$ (see Eq. 4.54). For this work, we need the first four orders of the bias expansion in Eulerian space which are

$$b_0 = 0\,, \tag{4.92}$$

$$b_1 = 1 + \frac{\nu_1^2 - 1}{\delta_1}\,, \tag{4.93}$$

$$b_2 = 2(1+a_2)\frac{\nu_1^2-1}{\delta_1} + \left(\frac{\nu_1}{\delta_1}\right)^2 (\nu_1^2 - 3)\,, \tag{4.94}$$

$$\begin{aligned} b_3 = {} & 6(a_2+a_3)\frac{\nu_1^2-1}{\delta_1} + 3(1+2a_2)\left(\frac{\nu_1}{\delta_1}\right)^2 (\nu_1^2-3) \\ & + \left(\frac{\nu_1}{\delta_1}\right)^2 \frac{\nu_1^4 - \nu_1^2 + 3}{\delta_1}\,, \end{aligned} \tag{4.95}$$

where $\nu_1 \equiv \delta_1/\sigma_1$ and $a_2 = -17/21$ and $a_3 = 341/567$ are coefficients of the expansion $\delta_0(\delta) = \sum_k a_k \delta^k$.

Sheth-Tormen mass function

The same consideration can be made for the Sheth-Tormen mass function, where one has to use the corresponding conditional mass function to calculate the different orders of the Taylor expansion. From the Sheth-Tormen mass function [76, 77], one finds the following fraction of halos with masses between m and $m + \mathrm{d}m$ in a background with density δ_0:

$$\begin{aligned} \frac{\mathrm{d}F(m_1|\delta_0,\sigma_0)}{\mathrm{d}m_1} = {} & \left(\frac{q}{2\pi}\right)^{1/2} A(p) \left(1 + \left[q\,\frac{(\delta_1-\delta_0)^2}{\sigma_1^2-\sigma_0^2}\right]^{-p}\right) \frac{\delta_1-\delta_0}{(\sigma_1^2-\sigma_0^2)^{3/2}} \\ & \times \exp\left[-\frac{q(\delta_1-\delta_0)^2}{2(\sigma_1^2-\sigma_0^2)}\right]\,, \end{aligned} \tag{4.96}$$

where $p, q, A(p)$ denote the constants that we already defined for the average Sheth-Tormen mass function in Sect. 4.3.3. In the large scale limit, where $\sigma_1^2 \gg \sigma_0^2$, this becomes

$$\frac{\mathrm{d}F(m_1|\delta_0,\sigma_0)}{\mathrm{d}m_1} = \left(\frac{q}{2\pi}\right)^{1/2} A(p) \left(1 + \left[q\,\frac{(\delta_1-\delta_0)^2}{\sigma_1^2}\right]^{-p}\right) \frac{\delta_1-\delta_0}{\sigma_1^3}\,. \tag{4.97}$$

Expanding the mass function in a Taylor series as in the case of the Press-Schechter formalism in the previous section, the Eulerian density contrast can be approximated. The first four orders give according to [73] the following bias parameters

$$b_0 = 0\,, \tag{4.98}$$
$$b_1 = 1 + \epsilon_1 + E_1\,, \tag{4.99}$$
$$b_2 = 2(1 + a_2)(\epsilon_1 + E_1) + \epsilon_2 + E_2\,, \tag{4.100}$$
$$b_3 = 6(a_2 + a_3)(\epsilon_1 + E_1) + 3(1 + 2a_2)(\epsilon_2 + E_2)\,, \tag{4.101}$$

where

$$\epsilon_1 = \frac{q\nu_1^2 - 1}{\delta_{\rm sc}}\,, \quad \epsilon_2 = \frac{q\nu_1^2}{\delta_{\rm sc}^2}(q\nu_1^2 - 3)\,, \tag{4.102}$$
$$\epsilon_3 = \frac{q\nu_1^2}{\delta_{\rm sc}^3}(q^2\nu_1^4 - 6q\nu_1^2 + 3)\,, \tag{4.103}$$
$$E_1 = \frac{2p/\delta_{\rm sc}}{1 + (q\nu_1^2)^p}\,, \quad \frac{E_1}{E_2} = \left(\frac{1 + 2p}{\delta_{\rm sc}} + 2\epsilon_1\right)\,, \tag{4.104}$$
$$\frac{E_3}{E_1} = \left[\frac{4(p^2 - 1) + 6pq\nu_1^2}{\delta_{\rm sc}^2} + 3\epsilon_1^2\right]\,, \tag{4.105}$$

where $\nu_1 \equiv \delta_1/\sigma_1$ and $a_2 = -17/21$ and $a_3 = 341/567$ are coefficients of the expansion $\delta_0(\delta) = \sum_k a_k\delta^k$ (see Eq. (4.54)). If one sets $p = 0$ and $q = 1$, the bias parameters reduce to the Press-Schechter bias as summarized in Sect. 4.4.2. In this case the E_n parameters are 0. If not stated otherwise, we will use the best fit parameters as found by Sheth and Tormen [77], i.e. the values $p = 0.3$ and $q = 0.707$.

4.4.3 Halo correlation functions

Equipped with a formalism that relates the halo density to the underlying matter density via a Taylor series $\mathcal{F}(\delta) = \frac{b_k}{k!}\delta^k$, we can finally find expressions for the correlation functions of halo densities. For higher-order correlations the calculations become lengthy, since we have to expand the non-linear density δ as well. As the consideration is valid on large scales, we can apply a perturbative approach and make use of the results from Sect. 3.5.

Power spectrum

The halo-halo correlation function of two halos with mass m_1 and m_2, we only require the linear bias parameter and find:

$$\begin{aligned}\xi_{\rm hh}(m_1, m_2, \boldsymbol{r}) &\equiv \langle\delta_{\rm h}(m_1, \boldsymbol{x})\delta_{\rm h}(m_2, \boldsymbol{x} + \boldsymbol{r})\rangle \\ &= b_1(m_1)b_1(m_2)\langle\delta(\boldsymbol{x})\delta(\boldsymbol{x} + \boldsymbol{r})\rangle \\ &\simeq b_1(m_1)b_1(m_2)\xi_{\rm lin}(\boldsymbol{r})\,,\end{aligned} \tag{4.106}$$

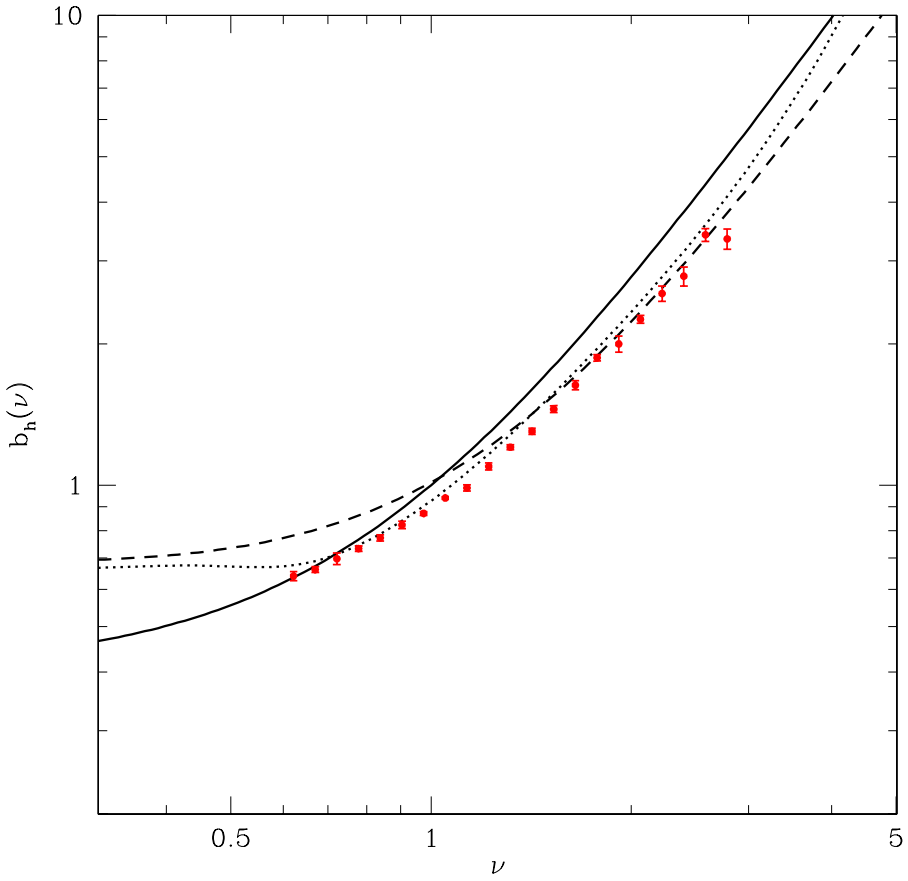

Figure 4.5: *Bias against ν (plot from Zentner [92]). The solid line is the extended Press-Schechter bias as calculated in Eq.* (4.60)*. The dashed line corresponds to the Sheth-Tormen bias given in Eq.* (4.61)*, whereas the dotted line corresponds to a fitting formula by Seljak and Warren [75]. The points illustrate results from N-body simulations of J. L. Tinker.*

where the last step holds for large scales. Fourier transforming the equation gives the halo-halo power spectrum:

$$P_{\rm hh}(m_1, m_2, k) = b(m_1)b(m_2)P_{\rm lin}(k)\,. \tag{4.107}$$

Bispectrum

For the halo bispectrum correlating three halos with masses m_1, m_2 and m_3, we have to expand $\mathcal{F}(\delta)$ to second-order, such that $\delta_{\rm h}(m,k) \simeq b_1\delta + b_2\delta^2$. As a result, one obtains

$$\begin{aligned}
(2\pi)^3\,\delta_{\rm D}\left(\boldsymbol{k}_{123}\right) B_{\rm hhh} &\equiv \langle\delta_{\rm h}(m_1,\boldsymbol{k}_1)\delta_{\rm h}(m_2,\boldsymbol{k}_2)\delta_{\rm h}(m_3,\boldsymbol{k}_3)\rangle_{\rm tree} \\
&= \langle[b_1(m_1)\delta(\boldsymbol{k}_1) + b_2(m_1)\delta(\boldsymbol{k}_1)][b_1(m_2)\delta(\boldsymbol{k}_2) + b_2(m_2)\delta(\boldsymbol{k}_2)] \\
&\quad \times [b_1(m_3)\delta(\boldsymbol{k}_3) + b_2(m_3)\delta(\boldsymbol{k}_3)]\rangle_{\rm tree} \\
&\simeq b_1(m_1)b_1(m_2)b_1(m_3)\langle\delta(\boldsymbol{k}_1)\delta(\boldsymbol{k}_2)\delta(\boldsymbol{k}_3)\rangle \\
&\quad + b_1(m_1)b_1(m_2)b_2(m_3)\langle\delta(\boldsymbol{k}_1)\delta(\boldsymbol{k}_2)\delta(\boldsymbol{k}_3)\rangle + 2\,\text{perms.}\,.
\end{aligned}$$

Considering the lowest order non-vanishing contribution in the density, one finds

$$\begin{aligned}
B_{\rm hhh} &= b_1(m_1)b_1(m_2)b_1(m_3)B_{\rm pt}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3) \\
&\quad + b_1(m_1)b_1(m_2)b_2(m_3)P_{\rm lin}(k_1)P_{\rm lin}(k_2) + 2\,\text{perms.}\,,
\end{aligned} \tag{4.108}$$

where $B_{\rm pt}$ denotes the lowest-order non-vanishing contribution from perturbation theory to the bispectrum and is defined in Sect. 3.5.2.

Trispectrum

Accordingly the trispectrum takes into account correlations of four halos in Fourier space and we need to expand the halo density up to the third order, such that $\delta_{\rm h}(m,k) \simeq b_1\delta + b_2\delta^2 + b_3\delta^3$. In a similar way as for the bispectrum we find

$$\begin{aligned}
&\langle\delta_{\rm h}(m_1,\boldsymbol{k}_1)\delta_{\rm h}(m_2,\boldsymbol{k}_2)\delta_{\rm h}(m_3,\boldsymbol{k}_3)\delta_{\rm h}(m_4,\boldsymbol{k}_4)\rangle_{\rm tree} \\
&\simeq \langle b_1(m_1)\delta(\boldsymbol{k}_1)b_1(m_2)\delta(\boldsymbol{k}_2)b_1(m_3)\delta(\boldsymbol{k}_3)b_1(m_4)\delta(\boldsymbol{k}_4)\rangle \\
&\quad + \langle b_1(m_1)\delta(\boldsymbol{k}_1)b_1(m_2)\delta(\boldsymbol{k}_2)b_1(m_3)\delta(\boldsymbol{k}_3)b_3(m_4)\delta^3(\boldsymbol{k}_4)\rangle + 3\,\text{perms.}
\end{aligned}$$

and obtain

$$\begin{aligned}
T_{\rm hhhh} &= b_1(m_1)b_1(m_2)b_1(m_3)b_1(m_4)T_{\rm pt}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\boldsymbol{k}_4) \\
&\quad + b_1(m_1)b_1(m_2)b_1(m_3)b_3(m_4)P_{\rm lin}(k_1)P_{\rm lin}(k_2)P_{\rm lin}(k_3) + 3\,\text{perms.}\,,
\end{aligned} \tag{4.109}$$

where $T_{\rm pt}$ denotes the lowest-order non-vanishing contribution from perturbation theory to the trispectrum and is defined in Sect. 3.5.3.

4.5 Halo density profile

The last ingredient we need for our halo model of dark matter is an accurate description of the inner structure of the halo, namely the halo density profile. Recent studies of numerical simulations seem to show (see Navarro et al. [56]) that the density profile of virialized halos on galaxy, group and cluster scale follows a universal function, depending only on the mass of the considered object. Furthermore, these simulations suggest more massive halos to be less peaked in the center region than less massive halos. Such profiles seem to be well described by functions of the form

$$\rho(r,m) = \frac{\rho_\mathrm{s}}{(r/r_\mathrm{s})^\alpha (1+r/r_\mathrm{s})^\beta} \tag{4.110}$$

or

$$\rho(r,m) = \frac{\rho_\mathrm{s}}{(r/r_\mathrm{s})^\alpha [1+(r/r_\mathrm{s})^\beta]}\,, \tag{4.111}$$

where $r_\mathrm{s} = r_\mathrm{s}(m)$ is the scale radius, which denotes the break between inner and outer core profile and $\rho_\mathrm{s} = \rho_\mathrm{s}(m)$ is the amplitude of the density profile[5]. The most prominent choice of parameters are $(\alpha,\beta) = (1,2)$ in Eq. (4.110) for the Navarro-Frenk-White (NFW) profile [56] and $(\alpha,\beta) = (3/2,3/2)$ in Eq. (4.111) for the Moore (M99) profile [54]. Both profiles differ on small scales for $r \lesssim r_\mathrm{s}$, but agree that the profile scales according to $\rho(r,m) \propto r^{-3}$ on large scales. Up to the present day, it is still unclear which of them provides the better description for virialized halos. One reason probably is that numerical simulations still cannot resolve the inner core of the halos to a high enough accuracy and that the power-law description in this region is too simplistic. A compromise between the two offers the Hernquist profile [32], which uses $(\alpha,\beta) = (1,2)$ in Eq. (4.110).

Although the density profile seems to dependent – apart from the halo mass – on the two quantities ρ_s and r_s, in practice it is only mass dependent. This can be seen if one parametrizes the profiles in Eq. (4.110) and (4.111) in an alternative way. The mass of a halo is defined as the mass within the virial radius r_vir, which leads to

$$m \equiv \int_0^{r_\mathrm{vir}} \mathrm{d}r\, 4\pi r^2 \rho(r,m) \tag{4.112}$$

assuming as usual a spherical form of the halo. The mass cut-off at r_vir is done to be consistent with the description of the spherical collapse model in section 4.2. An object is considered to be collapsed if

[5]Note that in general $\rho_\mathrm{s} \neq \rho(r_\mathrm{s})$. In case of the NFW-profile one has $\rho_\mathrm{s} = 4\rho(r_s)$, whereas for the M99-profile one obtains $\rho_\mathrm{s} = 2\rho(r_s)$.

$$\bar{\rho}(r_{\rm vir}) = \Delta_{\rm vir}\bar{\rho}\,, \tag{4.113}$$

which defines the mass of a bound object to be $m \equiv 4\pi/3 r_{\rm vir}^3 \Delta_{\rm vir}\,\bar{\rho}$ in the halo model description. Since for an EdS-Universe $\Delta_{\rm vir} \simeq 200$, the virial radius is often denoted in the literature as r_{200}. Changing the integration variable in Eq. (4.112) to $x \equiv r/r_{\rm vir}$ yields

$$\bar{\rho}(r_{\rm vir}) = 3\int_0^1 {\rm d}x\, x^2 \rho(x)\,, \tag{4.114}$$

where we additionally made use of $m = 4\pi/3\, r_{\rm vir}^3\,\bar{\rho}(r_{\rm vir})$. If we consider the special case of an NFW-profile (Eq. (4.110) with $(\alpha,\beta) = (1,2)$) and introduce the concentration parameter

$$c \equiv \frac{r_{\rm vir}}{r_{\rm s}}\,, \tag{4.115}$$

Eq. (4.114) can be rewritten as

$$\bar{\rho}(r_{\rm vir}) = 3\rho_{\rm s}\int_0^1 {\rm d}x\,\frac{x^2}{xc\,(1+xc)^2}\,.$$

Performing the integration and solving for $\rho_{\rm s}$, we find

$$\rho_{\rm s} = \frac{\bar{\rho}(r_{\rm vir})}{3}\,\frac{c^3}{\ln(1+c) - c/(1+c)}\,. \tag{4.116}$$

For the M99-profile (Eq. (4.111) with $(\alpha,\beta) = (3/2, 3/2)$), the analogous calculation yields

$$\rho_{\rm s} = \frac{\bar{\rho}(r_{\rm vir})}{2}\,\frac{c^3}{\ln(1+c^{3/2})}\,. \tag{4.117}$$

This means that the profile can be parametrized by $r_{\rm vir}$, which is equivalent to the mass m of the halo, and the concentration parameter c. Since simulations show that the concentration parameter depends strongly on the halo mass, it seems that the halo density profile is mainly determined by its mass. In spite of these results from numerical simulations, there is no analytical explanation for the existence of such a universal density profile. But there are successful models that find a mass dependence of the concentration parameter. Typically, they are of the form

$$c(m,z) = \frac{c_*}{1+z}\left(\frac{m}{m_*}\right)^{-\alpha}\,, \tag{4.118}$$

where $m_* = m_*(z=0)$ is the characteristic mass scale as introduced in section 4.3.1. For the NFW profile, Bullock et al. [13] find $c_* = 9$ and $\alpha = 0.13$ from numerical

simulations. A more recent analysis from Takada and Jain [84] finds $c_* = 10$ and $\alpha = 0.2$ instead, which we use for our halo model implementation. From Eq. (4.118), we can see that low mass halos tend to be more centrally concentrated on average than massive halos. Simulations show that the concentration parameters for halos of the same mass m scatter and that they follow a log-normal distribution

$$p(c)\mathrm{d}c = \frac{1}{\sqrt{2\pi\sigma^2_{\ln c}}} \exp\left[-\frac{(\ln c - \ln \bar{c})^2}{2\sigma^2_{\ln c}}\right] \mathrm{d}\ln c \tag{4.119}$$

according to Jing [38] with the average concentration parameter $\bar{c}$ following Eq. (4.118).

For our considerations it is important to transform the normalized profile to Fourier space. For a virialized halo of mass m we define

$$\tilde{u}(\boldsymbol{k}, m) \equiv \frac{\int_{r<r_{\mathrm{vir}}} \mathrm{d}^3x \rho(\boldsymbol{x}, m) \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}}{\int_{r<r_{\mathrm{vir}}} \mathrm{d}^3x \rho(\boldsymbol{x}, m)} \,. \tag{4.120}$$

In case of a spherically symmetric profile, this simplifies to

$$\tilde{u}(k, m) = \int_0^{r_{\mathrm{vir}}} \mathrm{d}r\, 4\pi r^2 \, \frac{\sin kr}{kr} \, \frac{\rho(r, m)}{m} \,, \tag{4.121}$$

where the integration is truncated at r_{vir}, as most of the mass is enclosed within this radius. Since we use in our implementation of the halo model mainly the NFW profile, we give an analytic expression for its Fourier transformation:

$$\begin{aligned}\tilde{u}(k, m) = \frac{4\pi\rho_{\mathrm{s}} r_{\mathrm{s}}^3}{m} \Big\{ & \sin(kr_{\mathrm{s}}) \left[\mathrm{Si}([1+c]kr_{\mathrm{s}}) - \mathrm{Si}(kr_{\mathrm{s}})\right] \\ & - \frac{\sin(ckr_{\mathrm{s}})}{(1+c)kr_{\mathrm{s}}} \cos(kr_{\mathrm{s}}) \left[\mathrm{Ci}([1+c]kr_{\mathrm{s}}) - \mathrm{Ci}(kr_{\mathrm{s}})\right] \Big\} \,,\end{aligned} \tag{4.122}$$

where we use the notation

$$\mathrm{Ci}(x) = -\int_x^\infty \frac{\cos t}{t} \, \mathrm{d}t \quad \text{and} \quad \mathrm{Si}(x) = \int_0^x \frac{\sin t}{t} \, \mathrm{d}t \,, \tag{4.123}$$

for the sine- and cosine integrals.

4.6 Building blocks for correlation functions

With the ingredients of the forgoing sections, it is possible to define a building block, which simplifies the notation for the correlation functions and spectra significantly:

$$M_{ij}(k_1,\ldots,k_j|m) \equiv \int \mathrm{d}m\, n(m) \left(\frac{m}{\bar{\rho}}\right)^j b_i(m)[\tilde{u}(k_1,m)\cdot\ldots\cdot\tilde{u}(k_j,m)]\,, \qquad (4.124)$$

where $n(m)$ is the halo mass function, $b_i(m)$ the halo bias parameter, $\tilde{u}(k,m)$ the normalized halo profile and $\bar{\rho}$ the comoving energy density of the Universe. For consistency reasons, we set $b_0 \equiv 1$. If we want to take into account the variation of the concentration parameter c for a fixed halo mass m as well, we have to perform an additional integration and the building block changes to

$$M_{ij}(k_1,\ldots,k_j|m,c) \equiv \int \mathrm{d}m\,\mathrm{d}c\, n(m)\, p(c) \left(\frac{m}{\bar{\rho}}\right)^j b_i(m)[\tilde{u}(k_1,m,c)\cdot\ldots\cdot\tilde{u}(k_j,m,c)]\,, \qquad (4.125)$$

where $p(c)$ is the probability distribution function for c as defined in Equation (4.119). Including additionally a redshift dependence, we would have to consider $b = b(m,z)$ and $n = n(m,z)$ instead.

Power spectrum

The power spectrum consists in the halo model approach of two terms

$$P(k) = P_{1\mathrm{h}}(k) + P_{2\mathrm{h}}(k) \qquad (4.126)$$

which are defined as

$$P_{1\mathrm{h}}(k) = M_{02}(k,k)\,, \qquad (4.127)$$
$$P_{2\mathrm{h}}(k) = [M_{11}(k)]^2 P_{\mathrm{lin}}(k) \qquad (4.128)$$

and denote correlations within one halo and between two halos.

Bispectrum

The bispectrum is defined as the connected third order moment in Fourier space and decomposes into three terms

$$B = B_{1\mathrm{h}} + B_{2\mathrm{h}} + B_{3\mathrm{h}}\,, \qquad (4.129)$$

where we omit the $(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3)$ dependency and define

$$B_{1\mathrm{h}} = M_{03}(k_1, k_2, k_3) \tag{4.130}$$

as the one-halo term,

$$\begin{aligned} B_{2\mathrm{h}} = {} & M_{11}(k_1) M_{12}(k_2, k_3) P_{\mathrm{lin}}(k_1) \\ & + M_{11}(k_3) M_{12}(k_1, k_2) P_{\mathrm{lin}}(k_3) \\ & + M_{11}(k_2) M_{12}(k_3, k_1) P_{\mathrm{lin}}(k_2) \end{aligned} \tag{4.131}$$

as the two-halo term,

$$\begin{aligned} B_{3\mathrm{h}} = {} & M_{11}(k_1) M_{11}(k_2) M_{11}(k_3) B_{\mathrm{pt}}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) \\ & + M_{11}(k_1) M_{11}(k_2) M_{21}(k_3) P_{\mathrm{lin}}(k_1) P_{\mathrm{lin}}(k_2) \\ & + M_{11}(k_3) M_{11}(k_1) M_{21}(k_2) P_{\mathrm{lin}}(k_3) P_{\mathrm{lin}}(k_1) \\ & + M_{11}(k_2) M_{11}(k_3) M_{21}(k_1) P_{\mathrm{lin}}(k_2) P_{\mathrm{lin}}(k_3) \end{aligned} \tag{4.132}$$

as the three-halo term, which describes correlations between three halos.

Trispectrum

The trispectrum is defined as the connected fourth order moment in Fourier space and splits in the halo model approach into four terms

$$T = T_{1\mathrm{h}} + T_{2\mathrm{h}} + T_{3\mathrm{h}} + T_{4\mathrm{h}}\,, \tag{4.133}$$

where we omit the $(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3, \boldsymbol{k}_4)$ dependency and define

$$T_{1\mathrm{h}} = M_{04}(k_1, k_2, k_3, k_4)\,, \tag{4.134}$$

as the one-halo term,

$$T_{2\mathrm{h}} = T_{2\mathrm{h}}^{31} + T_{2\mathrm{h}}^{22}\,, \tag{4.135}$$

as the two-halo term, which is further subdivided into T_{2h}^{31}, corresponding to correlations of three points in one halo and a fourth point in a second halo and T_{2h}^{22}, describing correlations of two points in one and another two points in the second halo:

$$\begin{aligned} T_{\mathrm{2h}}^{31} &= M_{13}(k_2,k_3,k_4)M_{11}(k_1)P_{\mathrm{lin}}(k_1)\,, \\ &\quad + M_{13}(k_1,k_2,k_3)M_{11}(k_4)P_{\mathrm{lin}}(k_4) \\ &\quad + M_{13}(k_4,k_1,k_2)M_{11}(k_3)P_{\mathrm{lin}}(k_3) \\ &\quad + M_{13}(k_3,k_4,k_1)M_{11}(k_2)P_{\mathrm{lin}}(k_2) \qquad (4.136) \\ T_{\mathrm{2h}}^{22} &= M_{12}(k_1,k_2)M_{12}(k_3,k_4)P_{\mathrm{lin}}(|\boldsymbol{k}_1+\boldsymbol{k}_2|) \\ &\quad + M_{12}(k_1,k_3)M_{12}(k_2,k_4)P_{\mathrm{lin}}(|\boldsymbol{k}_1+\boldsymbol{k}_3|) \\ &\quad + M_{12}(k_1,k_4)M_{12}(k_2,k_3)P_{\mathrm{lin}}(|\boldsymbol{k}_1+\boldsymbol{k}_4|) \qquad (4.137) \end{aligned}$$

(4.138)

the three-halo term is given by

$$\begin{aligned} T_{\mathrm{3h}} &= M_{12}(k_3,k_4)M_{11}(k_1)M_{11}(k_2)B_{\mathrm{pt}}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3+\boldsymbol{k}_4) + 5\,\mathrm{perms.} \\ &\quad + M_{22}(k_3,k_4)M_{11}(k_1)M_{11}(k_2)P_{\mathrm{lin}}(k_1)P_{\mathrm{lin}}(k_2) + 5\,\mathrm{perms.} \qquad (4.139) \end{aligned}$$

the four-halo term is

$$\begin{aligned} T_{\mathrm{4h}} &= M_{11}(k_1)M_{11}(k_2)M_{11}(k_3)M_{11}(k_4)T_{\mathrm{pt}}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\boldsymbol{k}_4) \\ &\quad + M_{11}(k_1)M_{11}(k_2)M_{11}(k_3)M_{21}(k_4)P_{\mathrm{lin}}(k_1)P_{\mathrm{lin}}(k_2)P_{\mathrm{lin}}(k_2) + 3\,\mathrm{perms.} \qquad (4.140) \end{aligned}$$

which describes correlations between points of four halos.

4.7 Summary

The basic assumption of the halo model is that all dark matter in the Universe is distributed in units of virialized halos. This suggests a scale-dependent consideration. On small scales, the correlation of dark matter is governed by the mass profile of the halo, whereas on large scales the halo clustering determines the nature of the correlation. As there are a multitude of models to describe the behavior on different scales and an even larger amount of parameters one has to set judiciously, there exists no such thing as a unique halo model. In order to have reproducible results, it is therefore necessary to specify ones choice of parameters. For this work, we will adopt the following parameters for the halo models. Deviations from this choice will be stated in the sections concerned:

1. The average mass of a halo is defined as the mass within a sphere of virial radius $r_{\rm vir}$ as $m \equiv (4\pi/3) r_{\rm vir}^3 \Delta_{\rm vir} \bar{\rho}$. In our implementation we use $\delta_{\rm sc}$ and $\Delta_{\rm vir}$ as calculated from the spherical collapse for a ΛCDM-Universe as given in Eq. (B.6) and (B.7).

2. N-body simulations suggest that the mass profile of a halo follows a universal function. As the NFW profile is in good agreement with numerical results and has an analytical Fourier transform, we use

$$\rho(r, m) = \frac{\rho_{\rm s}}{(r/r_{\rm s})^{\alpha}(1 + r/r_{\rm s})^{\beta}} \tag{4.141}$$

with $(\alpha, \beta) = (1, 2)$ as our standard profile.

3. Essential for the form of the halo profile is the concentration parameter, which is defined as $c \equiv r_{\rm vir}/r_{\rm s}$. From N-body simulations one can find that c is a function of the halo mass and follows

$$c(m, z) = \frac{c_*}{1+z} \left(\frac{m}{m_*} \right)^{-\alpha} , \tag{4.142}$$

where $m_* = m_*(z = 0)$ is the characteristic mass as defined within the Press-Schechter formalism. In the following we will use the values $c_* = 10$ and $\alpha = 0.2$ as proposed by Takada and Jain [84]. For the Millennium Run we will adopt the values $c_* = 11$ and $\alpha = 0.1$ as found recently by Neto et al. [57]. This implies that more massive halos are less centrally concentrated than less massive ones.

4. The abundance of halos depends on the mass m and is in our implementation equivalent to the Sheth and Tormen mass function

$$\nu f(\nu) = A(p) \left(1 + (q\nu^2)^{-p} \right) \sqrt{\frac{2q}{\pi}} \, \nu \exp\left(-q\nu^2/2 \right) , \tag{4.143}$$

with the parameters $p = 0.3$, $q = 0.707$ and $A(0.3) = 0.322$, which is a modification of the original Press-Schechter formulation.

5. As the spatial distribution of halos is biased, an important ingredient are the bias functions, which relate the dark matter density field to the halo density. It is derived on the basis of the Sheth-Tormen mass function above. For the linear bias one obtains then

$$b_1(\nu) = 1 + \frac{q\nu - 1}{\delta_{\rm sc}} + \frac{2p}{\delta_{\rm sc}[1 + (q\nu)^p]} , \tag{4.144}$$

where p and q match the values of the above mass function.

6. To obtain the final correlation function, one has to perform a mass integration, which goes formally from 0 to ∞. In practice, we use the mass limits $m_{\rm min} = 10^3 M_\odot$ and $m_{\rm max} = 10^{16} M_\odot$.

7. Due to the mass-cutoff the consistency relation

$$\int_{m_{\min}}^{m_{\max}} \mathrm{d}\nu \, n(\nu) b(\nu) = 1 \tag{4.145}$$

does not hold any longer. To cure this problem we introduce a normalization factor $b_{\mathrm{norm}}(z)$, which ensures that Eq. (4.145) is fulfilled. It has to be included for every appearance of the linear bias factor $b_1(\nu)$.

Chapter 5

Covariance of the power spectrum estimator

Weak gravitational lensing is an important tool to probe the large-scale structure of the Universe and to estimate cosmological parameters. The constraints it provides are independent of and complementary to those found by other measurements such as CMB anisotropies, SN type Ia or galaxy surveys. As the next generation of surveys will allow much more precise measurements of weak lensing effects, it is important to have a good understanding of the underlying physics and the expected errors. The quantity of main interest is the cosmic shear field. It quantifies the distortion of faint galaxy images by gravitational light deflection and depends directly on the projected mass density of the matter density. So far, mainly the convergence power spectrum has been measured and parameter constraints are based on the Gaussian approximation of the power spectrum covariance estimator. A more realistic analysis would imply the consideration of the non-Gaussian part of the covariance as well. In this chapter, we will investigate the non-Gaussianities in detail and determine the full covariance within the halo model approach. It continues and extends the work of Scoccimarro et al. [72] and Cooray and Hu [16].

Starting from a canonical estimator for the dark matter power spectrum, we derive in Sect. 5.1 its covariance for different k-modes and extend our consideration to the projected power spectrum estimator and the corresponding covariance in Sect. 5.2. As the evaluation of the covariance requires in both cases the calculation of trispectrum configurations where the wave-vectors form a parallelogram, we apply in Sect. 5.3 the halo model approach to find an analytic expression. In order to minimize the computational effort for the covariance, Sect. 5.3.2 addresses different approximations to the trispectrum and studies their validity for different configurations of the trispectrum wave-vectors. They are later applied to the actual covariance in Sect. 5.3.3 and used throughout this work. Sect. 5.4 covers the effect of a log-normal distribution for the concentration parameter in the power and trispectrum. To understand the origin of cross-correlations in the power spectrum covariance, Sect. 5.5 examines how the form and amplitude of the trispectrum depends on the angle between the trispectrum wave-vectors. Finally, in Sect. 5.6, we check the amount of non-Gaussianity on dif-

ferent scales for the case of the three-dimensional and the projected power spectrum covariance.

5.1 Covariance of the dark matter power spectrum estimator

We are interested in the statistical properties of the dark matter density field $\delta(\boldsymbol{x})$ and its Fourier modes

$$\delta(\boldsymbol{k}) = \int \mathrm{d}^3x\, \delta(\boldsymbol{x})\, \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}\,, \tag{5.1}$$

where $\boldsymbol{x} \in \mathbb{R}^3$. More precisely, we want to determine the power spectrum of $\delta(\boldsymbol{x})$ and the corresponding covariance for wave-vectors of different length k. As we will see later, this requires considering moments up to fourth order. Let us assume for our analysis that a survey (or simulation) of volume V is given from which $\delta(\boldsymbol{k})$ can be estimated. In Fourier space the volume can be divided into k-bins of shell width Δk. Following Scoccimarro et al. [72], we choose

$$\hat{P}(k_i) = \frac{1}{V} \int_{\mathrm{s},k_i} \frac{\mathrm{d}^3k}{V_\mathrm{s}(k_i)} \delta(\boldsymbol{k})\delta(-\boldsymbol{k}) \tag{5.2}$$

as an estimator for the power spectrum, where the integration is performed along the i-th shell of volume $V_\mathrm{s}(k_i)$ (the subscript 's' stands for *shell*). The estimator is constructed such that it is unbiased in the limit of infinitely small shells, which can be seen by calculating its expectation value

$$\begin{aligned}
\langle \hat{P}(k_i) \rangle &= \frac{1}{V} \int_{\mathrm{s},k_i} \frac{\mathrm{d}^3k}{V_\mathrm{s}(k_i)} \langle \delta(\boldsymbol{k})\delta(-\boldsymbol{k}) \rangle \\
&= \frac{1}{V} \int_{\mathrm{s},k_i} \frac{\mathrm{d}^3k}{V_\mathrm{s}(k_i)} (2\pi)^3 \delta_\mathrm{D}(\boldsymbol{0})\, P(k) \\
&\approx P(k) \int_{\mathrm{s},k_i} \frac{\mathrm{d}^3k}{V_\mathrm{s}(k_i)} \frac{(2\pi)^3}{V} \delta_\mathrm{D}(\boldsymbol{0}) = P(k)\,.
\end{aligned} \tag{5.3}$$

Note that we used the definition of the power spectrum in the second step (see Eq. 3.42) and the identity $\delta_\mathrm{D}(\boldsymbol{0}) = V/(2\pi)^3$ (use Eq. 3.27 and consider $\boldsymbol{k} = 0$). The last approximation is only valid for small bin-sizes, since the variation of $P(k)$ within $V_\mathrm{s}(k_i)$ is then negligible. We can also calculate the covariance of the power spectrum

estimator (5.2) for small shells:

$$
\begin{aligned}
C_{ij} &\equiv \mathrm{Cov}\left(\hat{P}(k_i), \hat{P}(k_j)\right) \\
&= \langle \hat{P}(k_i)\hat{P}(k_j)\rangle - \langle \hat{P}(k_i)\rangle\langle \hat{P}(k_j)\rangle \\
&= \frac{1}{V^2}\int_{\mathrm{s},k_i}\frac{\mathrm{d}^3k_1}{V_\mathrm{s}(k_i)}\int_{\mathrm{s},k_j}\frac{\mathrm{d}^3k_2}{V_\mathrm{s}(k_j)}\,\left[\langle\delta(\boldsymbol{k}_1)\delta(-\boldsymbol{k}_1)\delta(\boldsymbol{k}_2)\delta(-\boldsymbol{k}_2)\rangle\right. \\
&\quad \left. - \langle\delta(\boldsymbol{k}_1)\delta(-\boldsymbol{k}_1)\rangle\langle\delta(\boldsymbol{k}_2)\delta(-\boldsymbol{k}_2)\rangle\right] \\
&= \frac{1}{V^2}\int_{\mathrm{s},k_i}\frac{\mathrm{d}^3k_1}{V_\mathrm{s}(k_i)}\int_{\mathrm{s},k_j}\frac{\mathrm{d}^3k_2}{V_\mathrm{s}(k_j)}\,\left[\langle\delta(\boldsymbol{k}_1)\delta(-\boldsymbol{k}_1)\delta(\boldsymbol{k}_2)\delta(-\boldsymbol{k}_2)\rangle_c\right. \\
&\quad \left. + \langle\delta(\boldsymbol{k}_1)\delta(\boldsymbol{k}_2)\rangle\langle\delta(-\boldsymbol{k}_1)\delta(-\boldsymbol{k}_2)\rangle + \langle\delta(\boldsymbol{k}_1)\delta(-\boldsymbol{k}_2)\rangle\langle\delta(-\boldsymbol{k}_1)\delta(\boldsymbol{k}_2)\rangle\right] \\
&= \frac{(2\pi)^3}{V^2}\int_{\mathrm{s},k_i}\frac{\mathrm{d}^3k_1}{V_\mathrm{s}(k_i)}\int_{\mathrm{s},k_j}\frac{\mathrm{d}^3k_2}{V_\mathrm{s}(k_j)}\left\{\delta_\mathrm{D}(\boldsymbol{0})\,T(\boldsymbol{k}_1,-\boldsymbol{k}_1,\boldsymbol{k}_2,-\boldsymbol{k}_2)\right. \qquad (5.4)\\
&\quad \left. + (2\pi)^3P^2(k_1)\left[\delta_\mathrm{D}(\boldsymbol{k}_1+\boldsymbol{k}_2)\,\delta_\mathrm{D}(-\boldsymbol{k}_1-\boldsymbol{k}_2) + \delta_\mathrm{D}(\boldsymbol{k}_1-\boldsymbol{k}_2)\,\delta_\mathrm{D}(-\boldsymbol{k}_1+\boldsymbol{k}_2)\right]\right\} \\
&= \frac{(2\pi)^6}{V^2}\int_{\mathrm{s},k_i}\frac{\mathrm{d}^3k_1}{V_\mathrm{s}(k_i)V_\mathrm{s}(k_j)}\,2P^2(k_1)\delta_\mathrm{D}(\boldsymbol{0})\,\delta_{ij} + \frac{1}{V}\,\bar{T}(\boldsymbol{k}_1,-\boldsymbol{k}_1,\boldsymbol{k}_2,-\boldsymbol{k}_2) \\
&\approx \frac{1}{V}\left[\frac{(2\pi)^3}{V_\mathrm{s}(k_i)}2P^2(k_i)\delta_{ij} + \bar{T}_{ij}\right], \qquad (5.5)
\end{aligned}
$$

where $\bar{T}_{ij}$ is the bin-averaged three-dimensional matter trispectrum

$$
\bar{T}_{ij} \equiv \bar{T}(k_i,k_j) = \int_{\mathrm{s},k_i}\int_{\mathrm{s},k_j}\frac{\mathrm{d}^3k_1}{V_\mathrm{s}(k_i)}\frac{\mathrm{d}^3k_2}{V_\mathrm{s}(k_j)}\,T(\boldsymbol{k}_1,-\boldsymbol{k}_1,\boldsymbol{k}_2,-\boldsymbol{k}_2)\,. \qquad (5.6)
$$

For the derivation of Eq. (5.5), we made use of the definition of the trispectrum (see Eq. 3.42) and again of $\delta_\mathrm{D}(\boldsymbol{0}) = V/(2\pi)^3$. The approximation in the last step relies on the assumption that the power spectrum is almost constant over the shell integration. The resulting equation (5.5) consists of two terms: a Gaussian part which scales as the power spectrum squared and only gives a contribution to the diagonal of the covariance matrix, and a non-Gaussian part which introduces correlations between the wave-vectors of different shells. Both terms are inversely proportional to the survey volume V, but have a different behavior with respect to the shell volume $V_\mathrm{s}(k_i)$. While the Gaussian term decreases with the size of the shell, the non-Gaussian term is independent of the binning, since the volumes of the shells cancel out with the integration.

If we choose the coordinate system for the wave-vectors such that the integration shells are placed concentrically to its origin the problem becomes rotation-symmetric. This suggests to parametrize the three-dimensional trispectrum with three quantities, namely the length of the two wave-vectors which we denote with k_1 and k_2 and the

cosine of the angle between them $x \equiv (\boldsymbol{k}_1 \cdot \boldsymbol{k}_2)/k_1 k_2 = \cos\theta$. Accordingly, we use for the three-dimensional trispectrum the subsequent parametrization

$$T(k_1, k_2, x) \equiv T(\boldsymbol{k}_1, -\boldsymbol{k}_1, \boldsymbol{k}_2, -\boldsymbol{k}_2)\,. \tag{5.7}$$

The appropriate choice of the coordinate system allows us then to reduce the number of integrations in Eq. (5.6) from six to three, resulting in

$$\bar{T}_{ij} = \frac{8\pi^2}{V_\mathrm{s}(k_i)V_\mathrm{s}(k_j)} \int_{\mathrm{s},k_i} \mathrm{d}k_1\, k_1^2 \int_{\mathrm{s},k_j} \mathrm{d}k_2\, k_2^2 \int_{-1}^{1} \mathrm{d}x\, T(k_1, k_2, x)\,. \tag{5.8}$$

Note that for the calculation of the above equation we only have to consider parallelogram configurations of the trispectrum wave-vectors.

If we choose a linear binning for k_i and k_j, such that they are centered in the integration interval, the shells have a volume

$$V_\mathrm{s}(k_i) = 4\pi \int_{k_i-\frac{\Delta k}{2}}^{k_i+\frac{\Delta k}{2}} \mathrm{d}k\, k^2 = 4\pi \left(k_i^2 \Delta k + \frac{(\Delta k)^3}{12} \right)\,, \tag{5.9}$$

which is approximately $V_\mathrm{s}(k_i) \simeq 4\pi k_i^2 \Delta k$ for small bin-sizes Δk. Equation (5.8) has in this case the form

$$\bar{T}_{ij} = \frac{1}{2} \frac{1}{(k_i k_j \Delta k)^2} \int_{k_i-\frac{\Delta k}{2}}^{k_i+\frac{\Delta k}{2}} \mathrm{d}k_1\, k_1^2 \int_{k_j-\frac{\Delta k}{2}}^{k_j+\frac{\Delta k}{2}} \mathrm{d}k_2\, k_2^2 \int_{-1}^{1} \mathrm{d}x\, T(k_1, k_2, x)\,. \tag{5.10}$$

For a logarithmic binning, Eqs. (5.9) and (5.10) would have to be modified, because $\Delta k \propto k_i$.

5.2 Covariance of the lensing power spectrum estimator

In the same way as for the dark matter power spectrum one can find an expression for the two-dimensional case of the convergence power spectrum covariance (see Sect. 3.6). Instead of the density field $\delta(\boldsymbol{k})$, we now consider the statistics of the projected density field

$$\kappa(\boldsymbol{\theta}) = \int_0^{w_\mathrm{H}} \mathrm{d}w\, W(w)\, \delta[f_K(w)\boldsymbol{\theta}, w]\,, \tag{5.11}$$

where w_H denotes the comoving distance to the horizon and its Fourier counterpart

$$\kappa(\boldsymbol{l}) = \int \mathrm{d}^2\theta\, \mathrm{e}^{\mathrm{i}\boldsymbol{l}\cdot\boldsymbol{\theta}} \kappa(\boldsymbol{\theta})\,. \tag{5.12}$$

For our considerations, we will assume that all background sources are situated at a comoving distance w_s, such that the weight function has the form

$$W(w) = \frac{3}{2}\,\Omega_\mathrm{m}\,\frac{H_0^2}{c^2 a}\,\frac{f_K(w_\mathrm{s}-w)}{f_K(w_\mathrm{s})}\,. \tag{5.13}$$

Analogously to the three-dimensional case, we are interested in estimating the power spectrum and the corresponding covariance for wave-vectors of different length l_i. The power spectrum estimator for the convergence field is constructed as

$$\hat{P}_\kappa(l_i) = \frac{1}{A}\int_{\mathrm{r},l_i}\frac{\mathrm{d}^2 l}{A_\mathrm{r}(l_i)}\kappa(\boldsymbol{l})\kappa(-\boldsymbol{l})\,, \tag{5.14}$$

which is also unbiased in the limit of infinitely small shells, since $\langle\hat{P}_\kappa(l_i)\rangle = P_\kappa(l_i)$. Here $A = 4\pi f_\mathrm{sky}$ denotes the survey area or the solid angle in steradian[1], where f_sky is the fractional sky coverage and the integration is performed along a two-dimensional annulus of volume $A_\mathrm{r}(l_i)$ (the subscript 'r' stands for *ring*).

With the estimator from Eq. (5.14), the evaluation of the covariance results in an expression analogous to the three-dimensional case

$$\mathcal{C}_{ij} \equiv \mathrm{Cov}\left(\hat{P}_\kappa(l_i),\hat{P}_\kappa(l_i)\right) = \frac{1}{A}\left[\frac{(2\pi)^2}{A_\mathrm{r}(l_i)}2P_\kappa^2(l_i)\delta_{ij} + \bar{T}_\kappa(l_i,l_j)\right] \tag{5.15}$$

with the bin-averaged convergence trispectrum

$$\bar{T}_{ij}^\kappa \equiv \bar{T}_\kappa(l_i,l_j) = \int_{\mathrm{r},l_i}\int_{\mathrm{r},l_j}\frac{\mathrm{d}^2 l_1}{A_\mathrm{r}(l_i)}\frac{\mathrm{d}^2 l_2}{A_\mathrm{r}(l_j)}\,T_\kappa(\boldsymbol{l}_1,-\boldsymbol{l}_1,\boldsymbol{l}_2,-\boldsymbol{l}_2)\,. \tag{5.16}$$

Again, the covariance consists of two contributions: a Gaussian part proportional to the convergence power spectrum squared and a non-Gaussian part in which the trispectrum dependence enters. The convergence power spectrum $P_\kappa(l)$ and the convergence trispectrum $T_\kappa(\boldsymbol{l}_1,-\boldsymbol{l}_1,\boldsymbol{l}_2,-\boldsymbol{l}_2)$ are the projections of the corresponding three-dimensional spectra along the comoving distance convolved with the weight function $W(w)$ of Eq. (5.13) as defined in Sect. 3.6.

If we choose an appropriate coordinate system for the integration over the wave-vectors, the problem becomes rotation-symmetric and we can parametrize the convergence trispectrum by three quantities as well, namely the length of the wave-vectors l_1 and l_2 and the cosine of the angle between them $x \equiv (\boldsymbol{l}_1\cdot\boldsymbol{l}_2)/l_1 l_2 = \cos\varphi$. Analogously to the three-dimensional case we define

$$T_\kappa(l_1,l_2,x) \equiv T_\kappa(\boldsymbol{l}_1,-\boldsymbol{l}_1,\boldsymbol{l}_2,-\boldsymbol{l}_2)\,. \tag{5.17}$$

[1]Multiplication with $\left(\frac{180^\circ}{\pi}\right)^2$ converts it to square degrees.

Making use of the symmetry properties of this problem, one angular integration becomes trivial and the integration in Eq. (5.16) simplifies to

$$\bar{T}^{\kappa}_{ij} = \frac{2\pi}{A_{\rm r}(l_i)A_{\rm r}(l_j)} \int_{{\rm r},l_i} {\rm d}l_1\, l_1 \int_{{\rm r},l_j} {\rm d}l_2\, l_2 \int_0^{2\pi} {\rm d}\varphi\, T_\kappa(l_1, l_2, \cos\varphi)\,. \tag{5.18}$$

If the bins are spaced linearly and l_i is placed in the arithmetic mean of the integration interval the shell area is

$$A_{\rm r}(l) = 2\pi \int_{l_i-\Delta l/2}^{l_i+\Delta l/2} {\rm d}l_1\, l_1 {\rm d}l\, l = 2\pi l \Delta l\,. \tag{5.19}$$

In this case, the averaged trispectrum becomes

$$\bar{T}^{\kappa}_{ij} = \frac{1}{2\pi l_i l_j (\Delta l)^2} \int_{l_i-\Delta l/2}^{l_i+\Delta l/2} {\rm d}l_1\, l_1 \int_{l_j-\Delta l/2}^{l_j+\Delta l/2} {\rm d}l_2\, l_2 \int_0^{2\pi} {\rm d}\varphi\, T_\kappa(l_1, l_2, \cos\varphi)\,. \tag{5.20}$$

If the bin-width is sufficiently small, we can make use of the mean value theorem and approximate integral (5.20) in the following way:

$$\bar{T}^{\kappa}_{ij} \simeq \frac{1}{2\pi} \int_0^{2\pi} {\rm d}\varphi\, T_\kappa(l_i, l_j, \cos\varphi)\,. \tag{5.21}$$

Since the 1-halo term of the trispectrum (Eq. 4.134) is in fact independent of the angle φ, an approximation of the covariance can be calculated without having to perform an integration at all.

5.3 Calculating the covariance in the halo model approach

We have seen that the covariance of the power spectrum estimator consists of two terms: a Gaussian part, which is proportional to the power spectrum squared and a non-Gaussian part, which is the bin-averaged trispectrum. To model the non-linear power spectrum one can make use of different fitting formulae (see e.g. Peacock and Dodds [62], Smith et al. [79]), but for the trispectrum there exists currently no reliable prediction. For this reason, we use the halo model approach to describe the trispectrum. To be consistent, the same halo model is also applied to calculate the power spectrum. In the following, we will summarize the halo model terms in the form we need them for our covariance consideration and analyze which of them are required for a consistent description on different scales. Finally, we will show the numerical results of the power spectrum covariance as calculated with our halo model implementation.

5.3.1 Trispectrum

From the analytic results of the power spectrum covariance (Eq. 5.4) and the lensing power spectrum covariance (Eq. 5.15), we know that only parallelogram configurations of the trispectrum wave-vectors contribute to the non-Gaussian terms. Instead of four arbitrary wave-vectors $(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3, \boldsymbol{k}_4)$, we now only have to consider $(\boldsymbol{k}_1, -\boldsymbol{k}_1, \boldsymbol{k}_2, -\boldsymbol{k}_2)$ configurations, where the two opposing wave-vectors are of the same length. Restricting our calculations to these configurations, the four different halo terms which add to the total trispectrum (see also Sect. 4.6) simplify. In order to shorten the resulting expressions, we make use of the M_{ij} building blocks as defined in Sect. 4.6. Neglecting the $\boldsymbol{k}$-dependencies where no confusion arises, the trispectrum contributions in the halo model description take the following form:

The 1-halo term becomes

$$T_{\mathrm{1h}} = M_{04}(k_1, -k_1, k_2, -k_2)\,, \tag{5.22}$$

the 2-halo term is

$$T_{\mathrm{2h}} = T_{\mathrm{2h}}^{31} + T_{\mathrm{2h}}^{22}\,, \tag{5.23}$$

which consists of T_{2h}^{31}, corresponding to correlations of three points within one halo and a fourth point in a second halo and T_{2h}^{22}, describing correlations of two points in one and another two points in the second halo:

$$\begin{aligned} T_{\mathrm{2h}}^{31} &= 2M_{13}(k_1, k_2, k_1)M_{11}(k_1)P_{\mathrm{lin}}(k_1)\,, \\ &\quad + 2M_{13}(k_2, k_1, k_2)M_{11}(k_2)P_{\mathrm{lin}}(k_2) \end{aligned} \tag{5.24}$$

$$T_{\mathrm{2h}}^{22} = M_{12}^2(k_1, k_2)\left[P_{\mathrm{lin}}(|\boldsymbol{k}_1 + \boldsymbol{k}_2|) + P_{\mathrm{lin}}(|\boldsymbol{k}_1 - \boldsymbol{k}_2|)\right]\,. \tag{5.25}$$

The three-halo term is given by

$$\begin{aligned} T_{\mathrm{3h}} = {} & 2M_{12}(k_1, k_2)M_{11}(k_1)M_{11}(k_2) \\ & \cdot \left[B_{\mathrm{pt}}(\boldsymbol{k}_1, \boldsymbol{k}_2, -\boldsymbol{k}_1 - \boldsymbol{k}_2) + B_{\mathrm{pt}}(\boldsymbol{k}_1, -\boldsymbol{k}_2, -\boldsymbol{k}_1 + \boldsymbol{k}_2)\right] \\ & + M_{22}(k_1, k_1)M_{11}^2(k_2)P_{\mathrm{lin}}^2(k_2) \\ & + M_{22}(k_2, k_2)M_{11}^2(k_1)P_{\mathrm{lin}}^2(k_1) \\ & + 4M_{22}(k_1, k_2)M_{11}(k_1)M_{11}(k_2)P_{\mathrm{lin}}(k_1)P_{\mathrm{lin}}(k_2)\,, \end{aligned} \tag{5.26}$$

and the four-halo term is

$$\begin{aligned} T_{\mathrm{4h}} = {} & M_{11}^2(k_1)M_{11}^2(k_2)T_{\mathrm{pt}}(\boldsymbol{k}_1, -\boldsymbol{k}_1, \boldsymbol{k}_2, -\boldsymbol{k}_2) \\ & + 2M_{11}^2(k_1)M_{11}(k_2)M_{21}(k_2)P_{\mathrm{lin}}(k_1)P_{\mathrm{lin}}^2(k_2) \\ & + 2M_{11}^2(k_2)M_{11}(k_1)M_{21}(k_1)P_{\mathrm{lin}}(k_2)P_{\mathrm{lin}}^2(k_1)\,, \end{aligned} \tag{5.27}$$

which describes correlations of points between four different halos. B_{pt} and T_{pt} denote the lowest order, non-vanishing perturbation theory contribution to the bi- and

trispectrum as derived in Sect. 2. For parallelogram configurations of the trispectrum wave-vectors, we have to consider the following perturbation theory bispectrum dependencies

$$\begin{aligned} B_{\rm pt}(\boldsymbol{k}_1, \boldsymbol{k}_2, -\boldsymbol{k}_1 \pm \boldsymbol{k}_2) = {} & 2F_2(\boldsymbol{k}_1, \boldsymbol{k}_2) P_{\rm lin}(k_1) P_{\rm lin}(k_2) \\ & + 2F_2(\boldsymbol{k}_2, -\boldsymbol{k}_1 \pm \boldsymbol{k}_2) P_{\rm lin}(k_2) P_{\rm lin}(|-\boldsymbol{k}_1 \pm \boldsymbol{k}_2|) \\ & + 2F_2(-\boldsymbol{k}_1 \pm \boldsymbol{k}_2, \boldsymbol{k}_1) P_{\rm lin}(|-\boldsymbol{k}_1 \pm \boldsymbol{k}_2|) P_{\rm lin}(k_1)\,, \end{aligned} \tag{5.28}$$

where the $F_2(\cdot,\cdot)$ kernels denote the symmetrized second-order coupling function (see Sect. 2.4.3). The perturbation theory trispectrum simplifies in comparison to the one for arbitrary wave-vector configurations (see Eq. 3.59) to the following expression:

$$\begin{aligned} T_{\rm pt} = {} & 4P^2(k_1) \left\{ [F_2(\boldsymbol{k}_1, -\boldsymbol{k}_1 - \boldsymbol{k}_2)]^2 P(|\boldsymbol{k}_1 + \boldsymbol{k}_2|) \right. \\ & \left. + [F_2(\boldsymbol{k}_1, \boldsymbol{k}_2 - \boldsymbol{k}_1)]^2 P(|\boldsymbol{k}_1 - \boldsymbol{k}_2|) \right\} \\ & + 4P^2(k_2) \left\{ [F_2(\boldsymbol{k}_2, -\boldsymbol{k}_1 - \boldsymbol{k}_2)]^2 P(|\boldsymbol{k}_1 + \boldsymbol{k}_2|) \right. \\ & \left. + [F_2(\boldsymbol{k}_2, \boldsymbol{k}_1 - \boldsymbol{k}_2)]^2 P(|\boldsymbol{k}_1 - \boldsymbol{k}_2|) \right\} \\ & + 8P(k_1)P(k_2) \left[F_2(\boldsymbol{k}_1, -\boldsymbol{k}_1 - \boldsymbol{k}_2) F_2(\boldsymbol{k}_2, -\boldsymbol{k}_1 - \boldsymbol{k}_2) P(|\boldsymbol{k}_1 + \boldsymbol{k}_2|) \right. \\ & \left. + F_2(\boldsymbol{k}_1, \boldsymbol{k}_2 - \boldsymbol{k}_1) F_2(\boldsymbol{k}_2, \boldsymbol{k}_1 - \boldsymbol{k}_2) P(|\boldsymbol{k}_1 - \boldsymbol{k}_2|) \right] \\ & + 12\, P^2(k_1) P(k_2) F_3(\boldsymbol{k}_1, -\boldsymbol{k}_1, \boldsymbol{k}_2) + 12\, P^2(k_2) P(k_1) F_3(\boldsymbol{k}_1, \boldsymbol{k}_2, -\boldsymbol{k}_2)\,, \end{aligned} \tag{5.29}$$

where we dropped the $(\boldsymbol{k}_1, -\boldsymbol{k}_1, \boldsymbol{k}_2, -\boldsymbol{k}_2)$ dependency in $T_{\rm pt}$ and $F_3(\cdot,\cdot,\cdot)$ denotes the third-order coupling function as defined in Sect. 2.4.3. The full trispectrum is the sum of the four halo terms such that

$$T(k_1, k_2, x) = T_{\rm 1h} + T_{\rm 2h} + T_{\rm 3h} + T_{\rm 4h}\,. \tag{5.30}$$

A detailed derivation of equations (5.22) – (5.29) can be found in appendix A. To display the trispectrum graphically, it is convenient to parametrize an arbitrary parallelogram configuration in terms of the ratio of the wave-numbers k_1 and k_2, the angle between the corresponding vectors and the length of one wave-vector k_1. In this way, one can fix the values for the ratio and the angle and plot the trispectrum against a varying wave-length k_1. If we set $|\boldsymbol{k}_1| = a|\boldsymbol{k}_2|$, the dimensionless dark matter trispectrum becomes

$$\Delta^2_{\rm pc}(k_1, a, x) \equiv \frac{k_1^3}{2\pi^2}\, T^{1/3}(k_1, k_2, x)\,, \tag{5.31}$$

where the subscript 'pc' indicates that we consider parallelogram configurations. The same notation can be used for the projected case. If we have $|\boldsymbol{l}_1| = a|\boldsymbol{l}_2|$, the dimensionless lensing trispectrum is

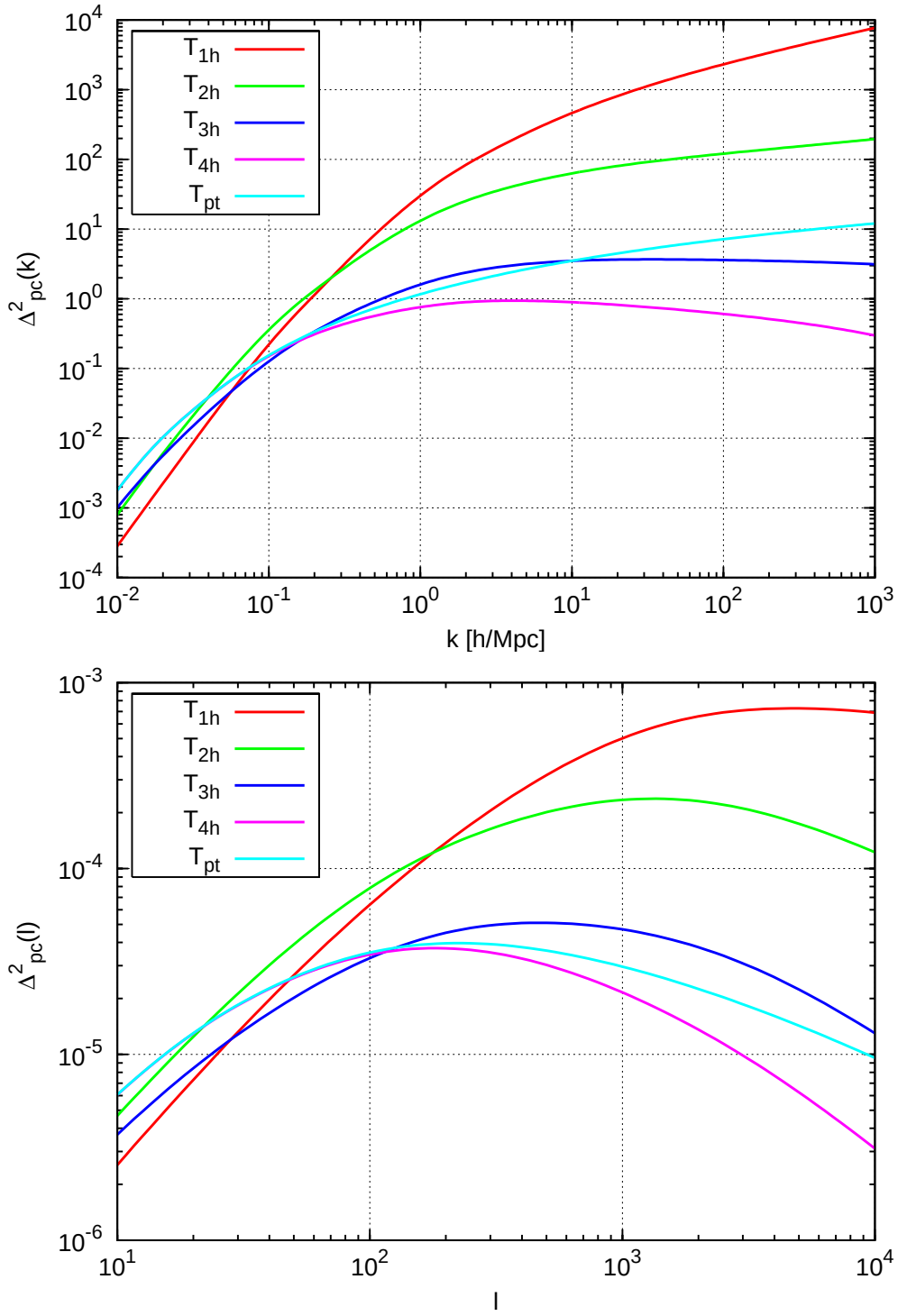

Figure 5.1: *Illustration of the different contributions to the dimensionless trispectrum against the wave-vector for a parallelogram configuration with* $a = 10$ *and* $x = \sqrt{2}/2$ *(see text for further explanations). The upper plot displays the three-dimensional case, whereas the lower plot shows the projected lensing trispectrum.*

Table 5.1: *Cosmological parameters used for calculating the trispectrum in the halo model approach. The parameters chosen match the cosmology of the Millennium Run simulation [82]. In order to set up the initial power spectrum the Eisenstein-Hu (EH) transfer function was used [22].*

Simulation	Ω_m	Ω_Λ	h	Ω_b	σ_8	n_s	Γ	z_s	$T(k)$
Millennium	0.25	0.75	0.73	0.045	0.9	1.0	0.14	1	EH

$$\Delta^2_\mathrm{pc}(l_1, a, x) \equiv \frac{l_1^2}{2\pi}\, T_\kappa^{1/3}(l_1, l_2, x)\,, \tag{5.32}$$

where $T_\kappa^{1/3}(l_1, l_2, x)$ denotes the projected version of the trispectrum. As an example, we calculate the different halo terms of a trispectrum configuration with $a = 10$ and $x = \sqrt{2}/2$. For our halo model implementation we use the same WMAP3-like cosmological parameters as for the Millennium Run simulation (see Tab. 5.1), which is applied throughout this chapter, if not stated otherwise. Fig. 5.1 illustrates the resulting trispectra for the three-dimensional (upper panel) and projected case (lower panel) against the corresponding wave-number k or l. As expected, in both cases the 1-halo term (red line) which accounts for dark matter correlations within one halo is dominant on small scales. Analogously, the 4-halo term (pink line) corresponding to correlations of dark matter between four different halos has the largest influence on large scales. For this configuration the 2-halo term (green line) has the largest contribution to the total trispectrum on intermediate scales and the 3-halo term (blue line) is of minor importance throughout the displayed scales. By construction the 4-halo term (turquoise line) equals the perturbation theory trispectrum on very large scales.

5.3.2 Contributions of the individual halo terms

As the calculation of the non-Gaussian covariance requires at least three integrations, it is useful to find approximations of the total trispectrum on different scales such that the computation-time is minimized. For this reason, we aim at quantifying the contributions of the individual halo terms to the overall trispectrum. From Fig. 5.1 we can rate the importance of the different halo terms. This suggests to examine the following approximations for different trispectrum configurations more closely:

$$\mathcal{T}_1 = T_\mathrm{1h}\,, \tag{5.33}$$

$$\mathcal{T}_2 = T_\mathrm{1h} + T_\mathrm{2h}\,, \tag{5.34}$$

$$\mathcal{T}_3 = T_\mathrm{1h} + T_\mathrm{2h} + T_\mathrm{3h}\,, \tag{5.35}$$

$$\mathcal{T}_4 = T_\mathrm{1h} + T_\mathrm{2h} + T_\mathrm{pt}\,. \tag{5.36}$$

where we additionally set $T_\mathrm{4h} \approx T_\mathrm{pt}$, since the perturbation theory trispectrum equals by construction the 4-halo term on very large scales and requires less computational

effort. Subsequently, we investigate these approximations to the total trispectrum T in two ways: first by calculating the ratio $\mathcal{T}_i/T$ and plotting it against the corresponding wave-number and second by analyzing the trispectrum configuration dependence of these approximations in order to generalize the results found. Note that the total trispectrum corresponds either to the three-dimensional trispectrum of Eq. (4.133) or to the corresponding projected version depending on the case considered. Furthermore, we choose a parametrization (k, a, x) or (l, a, x) for the trispectra as described in the previous section.

Quality of the different approximations

Fig. 5.2 illustrates the different approximations $\mathcal{T}_i$ compared with the overall trispectrum T for a parallelogram configuration with $a = 10$ and $x = \sqrt{2}/2$. More precisely, we calculate the ratio $\mathcal{T}_i/T$ for each of the approximations defined (see Eqs. 5.33-5.36) and plot it against the corresponding wave-number. In this way, values close or equal to 1 mean good or perfect agreement with the total trispectrum, whereas low or high values indicate a poor resemblance. In the following, we say that an approximation is in good agreement with the total trispectrum if the error $|\mathcal{T}_i/T - 1| \leq 10\%$. As we are interested in the three-dimensional as well as the projected trispectrum, the different configurations are plotted and analyzed for both cases.

The upper panel of Fig. 5.2 illustrates the quality of the four different approximations for the three-dimensional trispectrum in a range $10^{-2}\,h\,\mathrm{Mpc}^{-1} < k < 10^{2}\,h\,\mathrm{Mpc}^{-1}$. The $\mathcal{T}_1$ approximation (pink line) which consists only of the 1-halo term of the trispectrum provides the worst approximation. Only for wave-numbers larger than $k \gtrsim 30\,h\,\mathrm{Mpc}^{-1}$ the approximation deviates less than 10% from the total trispectrum. $\mathcal{T}_2$ (blue line) and $\mathcal{T}_3$ (green line) already provide a better approximation and can be applied safely on scales smaller than $k \simeq 0.2\,h\,\mathrm{Mpc}^{-1}$. Since both approximations are of similar quality, including the 3-halo term in the approximation is of minor importance for this trispectrum configuration. The best agreement with the total trispectrum is achieved for the $\mathcal{T}_4$-approximation (red line). It should be used if a good resemblance on very large scales is required.

Analogously, the lower panel of Fig. 5.2 compares the same approximations to the total convergence trispectrum. For our consideration we chose a range $10 < l < 10^4$. As one can see, the qualitative behavior of the approximations is identical to the three-dimensional case. The $\mathcal{T}_1$-approximation (pink line) is only applicable on scales smaller than $l \simeq 3000$, whereas the other approximations can already be used on scales smaller than $l \simeq 200$. Again, $\mathcal{T}_3$ (green line) does not improve the estimate of $\mathcal{T}_2$ (blue line) by much. If a good agreement on large scales is necessary, the $\mathcal{T}_4$-approximation (red line) should be preferred over the others.

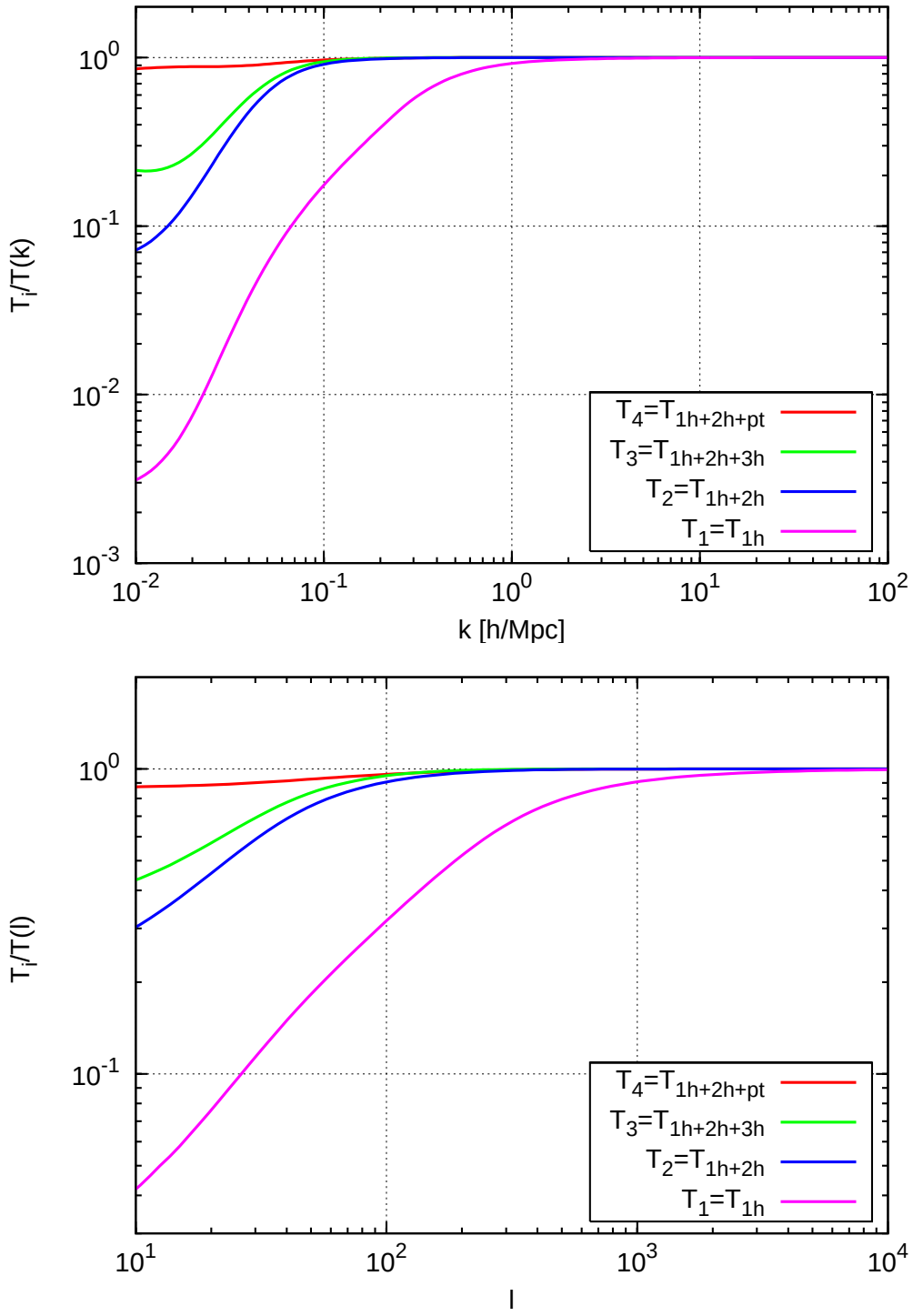

Figure 5.2: *Illustration of the different approximations of the trispectrum divided by the total trispectrum $\mathcal{T}_i/T(k)$ against the corresponding wave-number. The upper panel displays the three-dimensional case, while the lower panel shows the projected case for a parallelogram configuration, with ratio $a = 10$ and $x = \sqrt{2}/2$. $\mathcal{T}_4$ provides in both cases the best approximation to the total trispectrum, whereas $\mathcal{T}_1$ is only applicable on small scales. A good compromise between the two approximations is given by $\mathcal{T}_2$ (see text for a detailed discussion).*

Configuration dependence of the approximations

As the calculation of the power spectrum requires to average over all possible configurations of the trispectrum, we are interested to see if the results of the previous section hold as well for arbitrary configurations of the trispectrum. For this purpose, we investigate the validity of the approximations for four different configurations of the trispectrum. As before, we parametrize the parallelogram configurations by the length of one wave-vector $\boldsymbol{k}_1$ or $\boldsymbol{l}_1$, the ratio a between both wave-vectors and the cosine of the angle between them x. We focus on three configurations with fixed ratio a specified by

$$a = 1, \quad x = 0\,, \tag{5.37}$$

$$a = 10, \quad x = \sqrt{2}/2\,, \tag{5.38}$$

$$a = 100, \quad x = 0\,, \tag{5.39}$$

where in the three-dimensional case one wave-vector is varied in the range $10^{-2}\,h\,\mathrm{Mpc}^{-1} < k_1 < 10^{2}\,h\,\mathrm{Mpc}^{-1}$ and in the projected case one wave-vector is varied in the range $10 < l_1 < 10^4$. Additionally, we consider one configuration where the ratio a varies specified by

$$1 < a < 1000, \quad x = \sqrt{2}/2\,. \tag{5.40}$$

In this case, one wave-vector has a fixed value. For the three-dimensional consideration it is set to $k_2 = 0.1\,h\,\mathrm{Mpc}^{-1}$, for the projected case we set $l_2 = 100$. The first parallelogram configuration (5.37) corresponds geometrically to the special case of a square configuration. As the wave-vectors have the same length, this configuration only contributes to the diagonal part of the covariance matrix. The next two configurations (5.38, 5.39) are parallelogram configurations with different ratio and angle between the wave-vectors and correspond to off-diagonal entries of the covariance matrix. The larger the ratio a, the larger the separation from the diagonal of the covariance. Finally, the last configuration with varying ratio a probes the border region of the covariance matrix where one wave-length is fixed to a small value. In the following, we consider the validity of the $\mathcal{T}_1$ and $\mathcal{T}_2$ approximations for the four parallelogram configurations specified. Fig. 5.3 displays the results for the three-dimensional trispectrum, while Fig. 5.4 shows the results for the convergence trispectrum.

Results

The upper panel of Fig. 5.3 summarizes the results for the $\mathcal{T}_1$ approximation of the trispectrum in the three-dimensional case. For all configurations except (5.40), $\mathcal{T}_1$

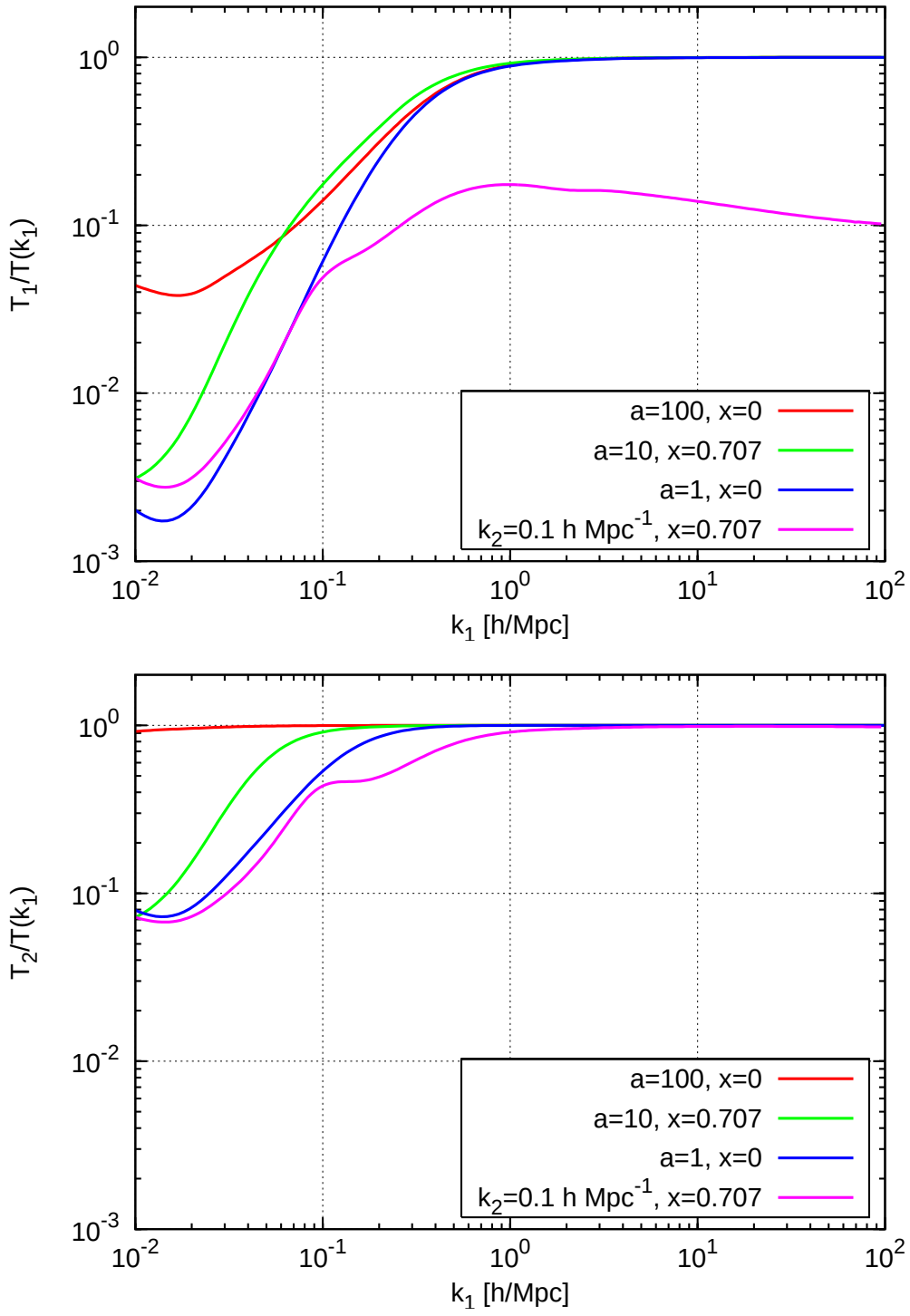

Figure 5.3: *Illustration of the different approximations of the three-dimensional trispectrum $\mathcal{T}_i$ divided by the total one against the wave-number k for four different parallelogram configurations. The upper plot displays the $\mathcal{T}_1$-approximation which considers only the 1-halo term contribution. The lower plot shows the results of the $\mathcal{T}_2$-approximations, which consists of 1-halo and 2-halo term. The trispectrum configurations are specified by the ratio a of the wave-length k_1 and k_2 and the cosine of the angle x between the wave-vectors $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$.*

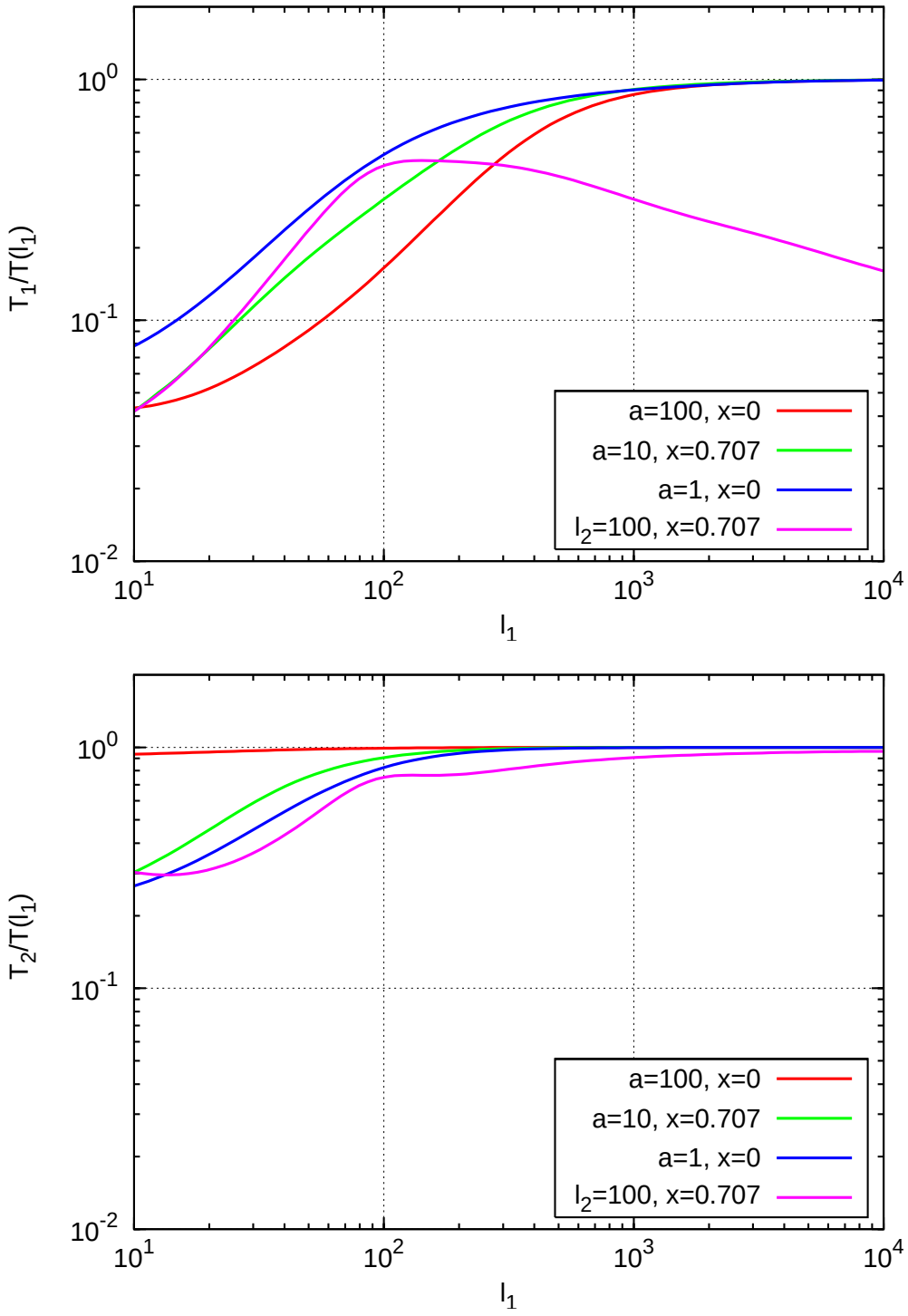

Figure 5.4: *Illustration of the different approximations of the projected trispectrum $\mathcal{T}_i$ divided by the total one against the wave-number l for four different parallelogram configurations. The upper plot displays the $\mathcal{T}_1$-approximation which considers only the 1-halo term contribution. The lower plot shows the results of the $\mathcal{T}_2$-approximations, which consists of 1-halo and 2-halo term. The trispectrum configurations are specified by the ratio a of the wave-length l_1 and l_2 and the cosine of the angle x between the wave-vectors $\boldsymbol{l}_1$ and $\boldsymbol{l}_2$.*

approximates the total trispectrum well for wave-length larger than $k \simeq 20\,h\,\mathrm{Mpc}^{-1}$. Configuration (5.40) was chosen to probe the border region of the covariance matrix. As we can see from Fig. 5.3, $\mathcal{T}_1$ fails to approximate the total trispectrum. It deviates over the considered range by more than 80%. Considering the four configurations in combination with the $\mathcal{T}_2$ approximation shows a significant improvement. On scales smaller than $k \simeq 0.3\,h\,\mathrm{Mpc}^{-1}$, $\mathcal{T}_2$ resembles the total trispectrum well for all configurations except (5.40). In case of the off-diagonal configuration (5.39) the $\mathcal{T}_2$ approximation even provides an accurate approximation of the whole range considered. Hence in regions further away from the diagonal of the covariance the 3-halo and 4-halo term are of minor importance and can be safely neglected.

Fig. 5.4 shows the same type of plots for the projected trispectrum configurations. Again, the qualitative behavior of the configurations is similar to the three-dimensional case. The $\mathcal{T}_1$ approximation deviates 10% or less from the total trispectrum T on scales smaller than $l \simeq 3000$ for all configurations considered except (5.40). The deviation in case of configuration (5.40) is less severe than in the three-dimensional case but is still larger than 50% on the scales considered. The lower panel of Fig. 5.4 shows the situation of the four parallelogram configurations for the $\mathcal{T}_2$ approximation of the total trispectrum. As for the three-dimensional case, there is a significant improvement in the quality of the approximations. All parallelogram configurations considered are approximated accurately by $\mathcal{T}_2$ on scales smaller than $l \simeq 3000$. If one is only interested in approximating configurations which are further away from the covariance diagonal, $\mathcal{T}_2$ already provides an accurate approximation for wave-length larger than $l \gtrsim 200$. For a square configuration, $\mathcal{T}_2$ resembles the total trispectrum very well on all scales from $200 < l < 10^4$.

From this analysis, we can draw the following conclusions: the 1-halo term ($\mathcal{T}_1$ approximation) is sufficient to describe the total trispectrum on small scales as long as the configurations considered deviate not too much from a square configuration. To describe intermediate scales accurately, one has to consider the 2-halo contribution as well ($\mathcal{T}_2$ approximation). For degenerate parallelogram configurations where one wave-number is very small, one has to consider the 3- or 4-halo term as well. The projection of the three-dimensional trispectrum along the redshift-axis has no significant effect on the qualitative behavior of the approximations considered. The results from this section valid for the large majority of trispectrum configurations are summarized in Tab. 5.2 and apply for a standard ΛCDM cosmology as given in Tab. 5.1. The three-dimensional trispectrum is considered at $z = 0$. For the convergence trispectrum we applied a single source redshift plane at $z_\mathrm{s} = 1$.

5.3.3 Lensing power spectrum covariance

With the results of the previous sections, we can finally calculate the covariance of the lensing power spectrum $\mathcal{C}_{ij}$ in the halo model description. As shown in Sect. 5.2 it consists of a Gaussian part which depends on the square of the convergence power

Table 5.2: *Validity range of the $\mathcal{T}_1$ and $\mathcal{T}_2$ approximations for the three-dimensional (3D) and projected (2D) trispectrum. For larger wave-numbers than the ones stated the corresponding approximation deviates less than 10% from the total trispectrum. The results apply for the majority of parallelogram configurations of the trispectrum wave-vectors in a standard ΛCDM cosmology (see Tab. 5.1).*

Approximation	3D trispectrum $k[\,h\,\mathrm{Mpc}^{-1}]$	2D trispectrum $l\ [\mathrm{rad}^{-1}]$
$\mathcal{T}_1$	$\gtrsim 3$	$\gtrsim 3000$
$\mathcal{T}_2$	$\gtrsim 0.3$	$\gtrsim 300$

spectrum and a non-Gaussian part consisting of the bin-averaged convergence trispectrum. The term bin-averaged here describes the averaging over all possible trispectrum configurations which can be found for wave-vectors within two annuli. In contrast to calculating the 'pure' trispectrum, for the covariance we have to specify the survey size and the bin-width that define the width of the annuli to be averaged over. We focus here on the projected case of the covariance as it has more relevance for observations in weak lensing. After considering the covariance for a specific example, we will apply the results of the different trispectrum approximations from Sect. 5.3.2 to the lensing power spectrum covariance and explore their relevance for upcoming weak lensing surveys.

Covariance for a standard ΛCDM cosmology

As illustrative example, we consider the covariance of the projected power spectrum estimator with cosmological parameters matching a standard ΛCDM cosmology (see Tab. 5.1). We adopt the same survey size $A = 25$ deg^2 and the same linear binning scheme with $\Delta l = 72$ starting from $l_0 = 144$ up to $l_{40} = 3024$ as used for the ray-tracing through the Millennium Run, which we will consider in the following chapter. The resulting covariance matrix is presented in Fig. 5.5 on a logarithmic scale. Typically, the covariance matrix has a maximum amplitude on large scales, i.e. for small wave-numbers l (red region) and decreases towards larger wave-numbers (dark blue region). The largest values can be found along the diagonal, corresponding to the variance of the power spectrum estimator. Under the Gaussian assumption, only a fraction of this diagonal is used for the error analysis. The non-Gaussian contribution to the diagonal and the non-Gaussian off-diagonals of the covariance matrix are neglected. In a subsequent section, we will analyze the amount of non-Gaussianity in comparison to the Gaussian contribution for the diagonal of the covariance in more detail.

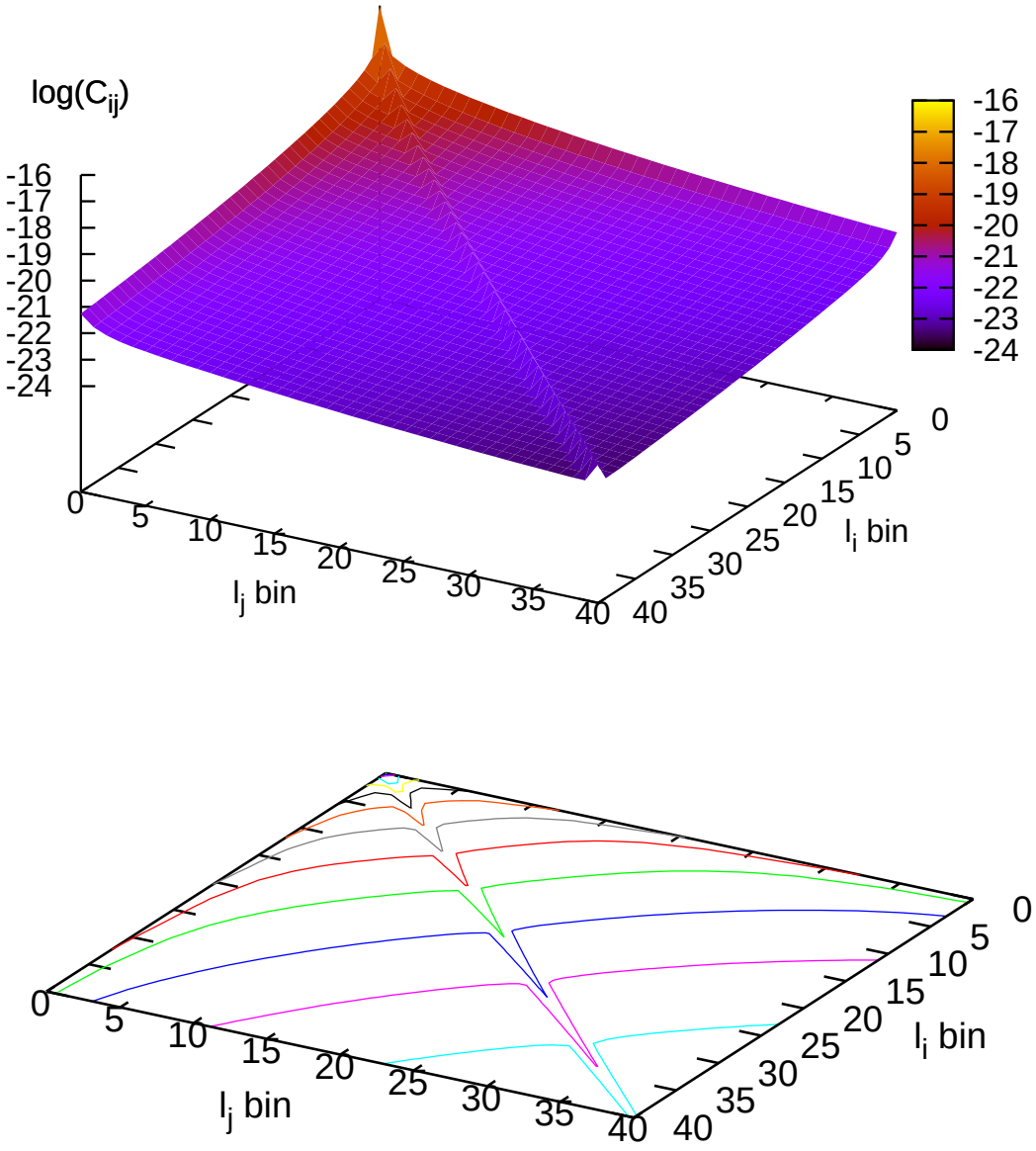

Figure 5.5: *Logarithmic convergence power spectrum covariance* $\log(C(l_i, l_j))$ *against wavenumbers* (l_i, l_j) *as calculated with the halo model. Survey area and cosmological parameters match the Millennium Run simulation. The binning scheme is linear with a width* $\Delta l = 72$ *ranging from* $l_0 = 144$ *up to* $l_{40} = 3024$*. The upper plot shows a three-dimensional visualization of the lensing covariance, whereas the lower plot depicts the corresponding contour representation. As can be seen from the contours the diagonal marks the highest amplitude of the total covariance.*

Approximations for the covariance

From the results of the trispectrum analysis in the previous section, we expect the 1-halo and the 2-halo term of the bin-averaged trispectrum to be the major contributions to the overall covariance. In order to analyze this issue further, we define analogous covariance approximations as for the trispectrum. The quality of these approximations is then visualized in contour plots.

For the considerations, it is convenient to decompose the covariance of the lensing power spectrum estimator into the Gaussian part and the four halo terms of the bin-averaged trispectrum

$$\begin{aligned}\mathcal{C}_{ij} &= \mathcal{C}_{ij}^{\mathrm{G}} + \mathcal{C}_{ij}^{\mathrm{NG}} \\ &= \mathcal{C}_{ij}^{\mathrm{G}} + \mathcal{C}_{ij}^{\mathrm{1h}} + \mathcal{C}_{ij}^{\mathrm{2h}} + \mathcal{C}_{ij}^{\mathrm{3h}} + \mathcal{C}_{ij}^{\mathrm{4h}}\,,\end{aligned} \tag{5.41}$$

where the components are defined as

$$\mathcal{C}_{ij}^{\mathrm{G}} = \frac{1}{A}\,\frac{(2\pi)^2}{A_{\mathrm{r}}(l_i)}\,2P_\kappa^2(l_i)\,\delta_{ij}\,, \tag{5.42}$$

$$\mathcal{C}_{ij}^{n\mathrm{h}} = \frac{1}{A}\int_{\mathrm{r},l_i}\frac{\mathrm{d}^2 l_1}{A_{\mathrm{r}}(l_i)}\int_{\mathrm{r},l_j}\frac{\mathrm{d}^2 l_2}{A_{\mathrm{r}}(l_j)}\,T_\kappa^{n\mathrm{h}}(\boldsymbol{l}_1,\boldsymbol{l}_2)\,,\quad n=1,2,3,4\,. \tag{5.43}$$

As already discussed in Sect. 5.2, these terms generally depend on the survey size A and the binning scheme since $A_{\mathrm{r}}(l) \propto \Delta l\, l$. Since the Gaussian part is well described by applying the standard fitting formulae for the power spectrum and its computation can be performed quickly, we focus in the following on the four halo terms of the bin-averaged trispectrum. In analogy to the analysis of the trispectrum in Sect. 5.3.2, the following four approximations are investigated further:

$$\mathcal{C}_1 = \mathcal{C}^{\mathrm{1h}}\,, \tag{5.44}$$

$$\mathcal{C}_2 = \mathcal{C}^{\mathrm{1h}} + \mathcal{C}^{\mathrm{2h}}\,, \tag{5.45}$$

$$\mathcal{C}_3 = \mathcal{C}^{\mathrm{1h}} + \mathcal{C}^{\mathrm{2h}} + \mathcal{C}^{\mathrm{3h}}\,, \tag{5.46}$$

$$\mathcal{C}_4 = \mathcal{C}^{\mathrm{1h}} + \mathcal{C}^{\mathrm{2h}} + \mathcal{C}^{\mathrm{4h}}\,, \tag{5.47}$$

where we omit the subscripts i and j for notational convenience. The different approximations $\mathcal{C}_i$ are visualized in four contour plots in Fig. 5.6. Each plot shows a different approximation divided by the total lensing covariance $\mathcal{C}$ versus wave-numbers (l_i, l_j). As one can see, the $\mathcal{C}_1$-approximation works well on small scales. For wave-numbers larger than $l \gtrsim 2500$, the deviation from the total trispectrum is 6% or less except along the diagonal. Here and close to the left and lower borders of the covariance

matrix where one wave-number is fixed to a small value the error is 12% or less. Using the $\mathcal{C}_2$-approximation the deviation to the total covariance along the borders and the diagonal decreases significantly. For wave-numbers larger than $l \simeq 1000$ the error is 3% or less. If one desires a smaller deviation on large scales more halo-terms must be considered, since for $l \lesssim 500$ the discrepancy to the total covariance is 13% or more. If a high precision is required, one should choose the $\mathcal{C}_3$-approximation. For wave-numbers $l \gtrsim 300$, the deviation is 6% or less. In contrast to $\mathcal{C}_2$ the $\mathcal{C}_4$-approximation improves the boundary region of the covariance matrix but is less accurate than the $\mathcal{C}_3$-approximation. A summary of this analysis can be found in Tab. 5.3, which is valid for a standard ΛCDM cosmology as given in Tab. 5.1. As we consider the ratio of the approximations to the total covariance, the results found are independent of the survey size. For small bin-width, they are also independent of the binning scheme and can be applied to all weak lensing surveys with source redshift planes situated at $z_\mathrm{s} = 1$.

Table 5.3: *Validity range of the approximations for the covariance of the lensing power spectrum as defined in Eqs. (5.44)-(5.47). For values larger than the wave-numbers stated the error is less than* 12% *or less than* 6% *depending on the column considered. The results apply to a standard* Λ*CDM cosmology (see Tab. 5.1) and source redshift planes situated at* $z_\mathrm{s} = 1$.

Approximation	Error $\lesssim 12\%$ l [rad^{-1}]	Error $\lesssim 6\%$ l [rad^{-1}]
$\mathcal{C}_1$	$\gtrsim 1200$	$\gtrsim 2500$
$\mathcal{C}_2$	$\gtrsim 200$	$\gtrsim 500$
$\mathcal{C}_3$	$\gtrsim 100$	$\gtrsim 200$
$\mathcal{C}_4$	$\gtrsim 100$	$\gtrsim 200$

5.4 Effect of a concentration parameter distribution

In Sect. 4.5 we saw that the density profile of a halo can be parametrized by its virial radius r_vir and the concentration parameter c. Up to now, we assumed that the concentration parameter in our halo model implementation is described by a simple mass-redshift dependency of the form

$$c(m,z) = \frac{c_*}{1+z}\left(\frac{m}{m_*}\right)^{-\alpha} \tag{5.48}$$

with $m_* = m_*(z=0)$. Recent numerical results [13, 38] indicate that this relation holds only for the average halo concentration, i.e. there is a scatter in the concentration parameter for halos of the same mass. Furthermore, Jing [38] proposes that the underlying concentration distribution is better described by a log-normal distribution

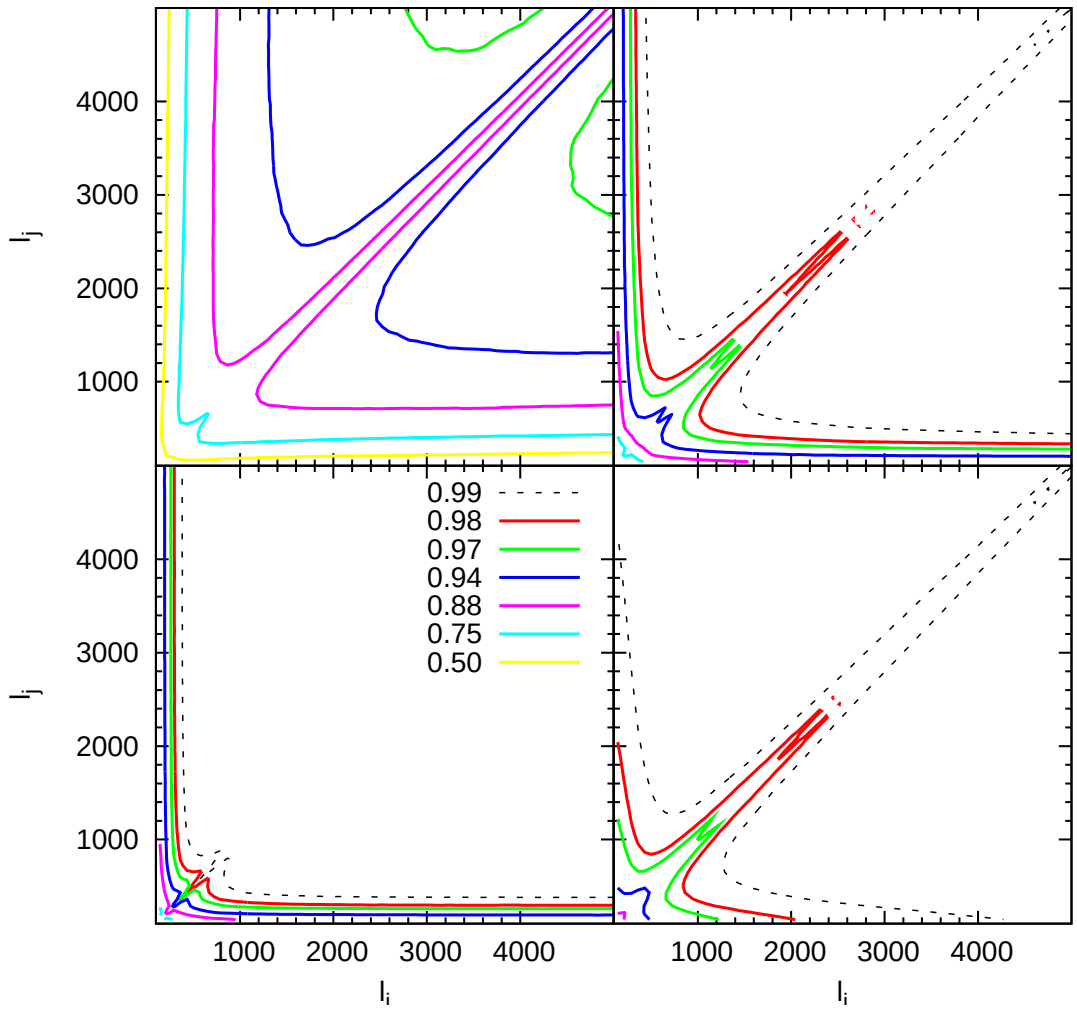

Figure 5.6: *Contour plots of the approximated to total lensing covariance $\mathcal{C}_k/\mathcal{C}$ against the wave-numbers (l_i, l_j) as calculated with the halo model. Each panel depicts a different approximation. In reading order these are: $\mathcal{C}_1$, $\mathcal{C}_2$, $\mathcal{C}_3$ and $\mathcal{C}_4$ as defined in Eqs. (5.44)-(5.47). To have an error smaller than 6% for wavenumbers larger than 600 one requires at least the 1-halo and 2-halo terms of the bin-averaged trispectrum.*

$$p(c)\mathrm{d}c = \frac{1}{\sqrt{2\pi\sigma^2_{\ln c}}} \exp\left[-\frac{(\ln c - \ln \bar{c})^2}{2\sigma^2_{\ln c}}\right] \mathrm{d}\ln c\,, \tag{5.49}$$

with the mean $\bar{c}$ following the original relation (5.48). Typical values for the concentration dispersion range from $\sigma_{\ln c} = 0.18$ to $\sigma_{\ln c} = 0.32$ [38, 86]. This variation of the halo concentration can be explained by the different merger histories of the halos (see Wechsler et al. [86]). Studying the concentration of halos with different formation redshifts z_f, they found a fitting formula

$$c(z_\mathrm{f}) = c_0(1 + z_\mathrm{f})\,, \tag{5.50}$$

where $c_0 = 4.1$ corresponds to the typical concentration of halos forming at the present time. This implies that, for a fixed mass m, halos forming at earlier epochs have a larger concentration than halos forming at later epochs which reflects the higher density of the Universe at earlier times.

In the following, we analyze the impact of a scatter in the halo concentration on the 1- and 2-halo terms of the power and trispectrum. To avoid confusion, we refer to the original concentration-mass relation in Eq. (5.48) as the *deterministic* concentration relation, since the probability distribution can be described by a Dirac delta distribution as

$$p(c)\mathrm{d}c = \delta_\mathrm{D}\left(c - \bar{c}\right) c\,\mathrm{d}\ln c \tag{5.51}$$

and denote the log-normal concentration distribution of Eq. (5.49) as *probabilistic* concentration relation.

5.4.1 Three-dimensional spectra

Since the concentration parameter characterizes the halo profile (see Sect. 4.5), we expect the change from a deterministic to a probabilistic concentration relation to have the strongest effect on small scales. Thus in our halo model description of structure, the 1-halo term and possibly the 2-halo term of the spectrum considered should be most sensitive to a scatter in the concentration. Since our main interest in this work is the power spectrum covariance, we analyze in the following the impact of a probabilistic concentration on the 1- and 2-halo term of the power and trispectrum for the three-dimensional and for the projected case.

Fig. 5.7 illustrates the results for the 1-halo terms of the three-dimensional power spectrum (upper panel) and the trispectrum (lower panel) for different dispersions $\sigma_{\ln c} = \{0.1, 0.2, 0.3, 0.4\}$ and the deterministic concentration denoted by $\sigma_{\ln c} = 0$. In order to show the effect more clearly we display in the plots the ratio of the 1-halo

term contribution as a function of the concentration dispersion to the 1-halo term contribution from the deterministic concentration relation $\Delta^2(k;\sigma_{\ln c})/\Delta^2(k;\sigma_{\ln c}=0)$ against k. Additionally, we consider for the trispectrum only square configurations, where $k_1 = k_2 = k$ and $x = 0$. One can clearly see that a probabilistic concentration raises the amplitude of the 1-halo terms significantly on small scales with respect to the original deterministic concentration. The larger the dispersion $\sigma_{\ln c}$, the stronger the deviation from the original deterministic concentration relation as given in Eq. (5.48). Note that the impact of the probabilistic concentration distribution is stronger on the trispectrum than the power spectrum. This is not surprising as the evaluation of the trispectrum requires to average over the fourth power of the halo profile, whereas to derive the power spectrum one averages merely over the squared halo profile.

The upper panel of Fig. 5.8 displays the results for the 2-halo term of the three-dimensional trispectrum for different dispersions $\sigma_{\ln c} = \{0.1, 0.2, 0.3, 0.4\}$ and the deterministic concentration denoted by $\sigma_{\ln c} = 0$ in the same way as for the 1-halo term. In the following, we do not consider the influence of a probabilistic concentration on the 2-halo term of the power spectrum, as it is vanishingly small (less than 1% deviation from the zero dispersion result). Similar to the 1-halo term consideration, the plot indicates that a larger concentration dispersion for the 2-halo term results in a stronger deviation from the 2-halo term with a deterministic concentration. In comparison to the 1-halo term of the trispectrum, the influence of the concentration distribution on the 2-halo term is less pronounced, but still noticeable. This is in agreement with the weaker dependence of the 2-halo term on the halo profile. For wave-numbers around $k \simeq 10^3\,h\,\mathrm{Mpc}^{-1}$ the deviation from the deterministic concentration varies between 10% (for $\sigma_{\ln c} = 0.1$) and 50% (for $\sigma_{\ln c} = 0.4$).

5.4.2 Projected spectra

As the probabilistic concentration has a strong effect on small scales of the three-dimensional spectra, we are also interested its impact on the projected spectra. For this reason, we consider in Fig. 5.9 the influence of a concentration dispersion on the 1-halo term of the convergence power and trispectrum. Additionally, the lower panel of Fig. 5.8 depicts the same dispersions in the 2-halo term of the projected trispectrum. In analogy to the three-dimensional case we plot the ratio of the corresponding halo term contribution as a function of the concentration dispersion to the corresponding halo term contribution from the deterministic concentration relation $\Delta^2_\kappa(l;\sigma_{\ln c})/\Delta^2_\kappa(l;\sigma_{\ln c}=0)$ against l. Again, we consider for the trispectrum only square configurations, where $l_1 = l_2 = l$ and $x = 0$. The plotted results show the same behavior as in the three-dimensional case. But in contrast to the three-dimensional counterparts, the effect of a concentration dispersion is diminished in the projected spectra. This is a result of the projection along the redshift axis. In case of the 1-halo term of the power spectrum with $\sigma_{\ln c} = 0.3$ (upper panel, blue line) this means a deviation from a deterministic concentration relation of about $1\% - 2\%$ for wave-numbers

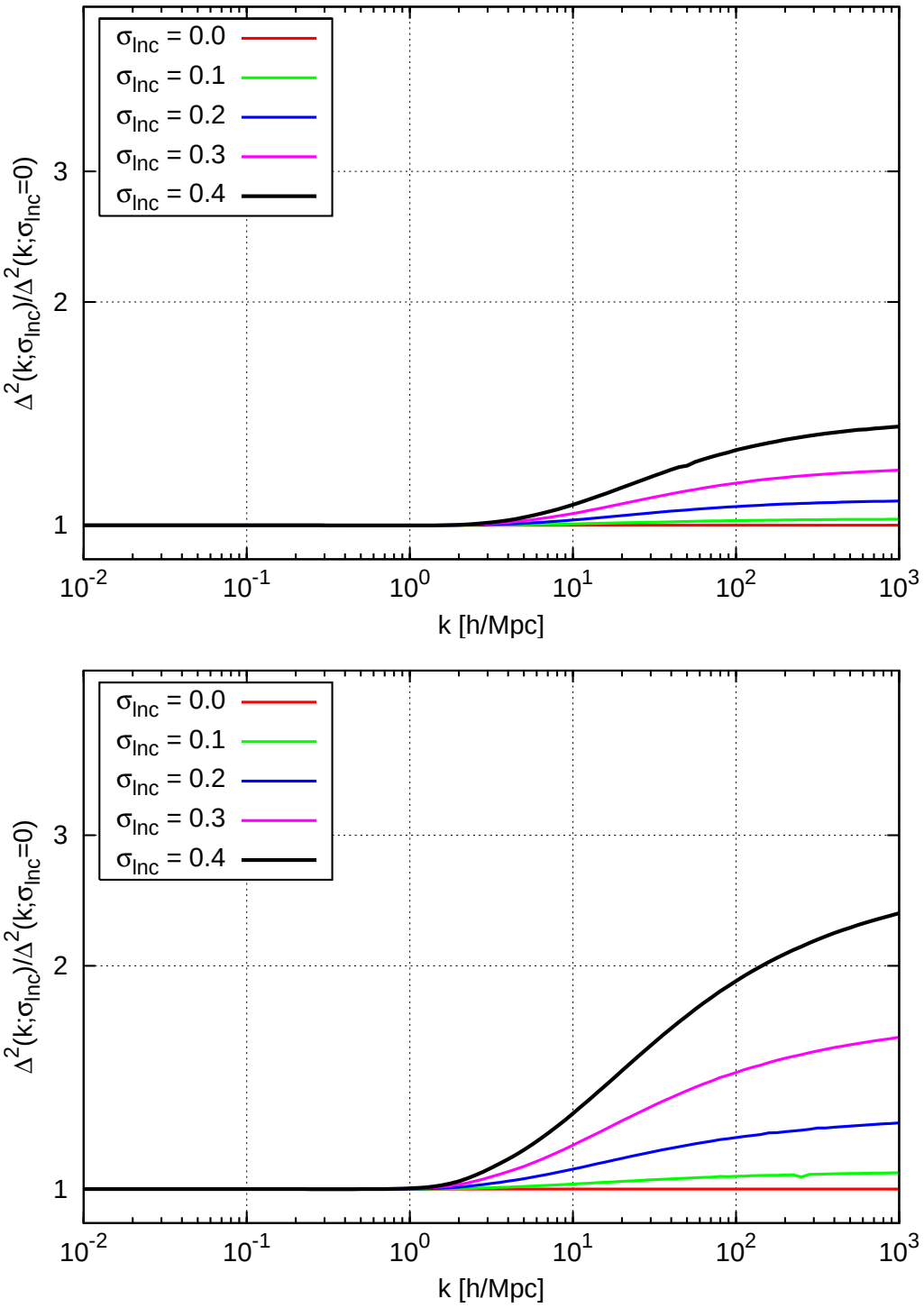

Figure 5.7: *Ratio of the 1-halo term contribution as a function of the concentration dispersion to the 1-halo term contribution from the deterministic concentration relation of different three-dimensional spectra against wave-number k. The curves show the ratio for four different dispersions of the halo concentration $\sigma_{\ln c} = \{0.1, 0.2, 0.3, 0.4\}$ and the deterministic concentration denoted by $\sigma_{\ln c} = 0$. The upper plot displays different concentration dispersions of the power spectrum, whereas the lower panel shows the same dispersions for the trispectrum. As one can see the dispersion has a strong effect on small scales. For the trispectrum the effect is more pronounced than for the power spectrum.*

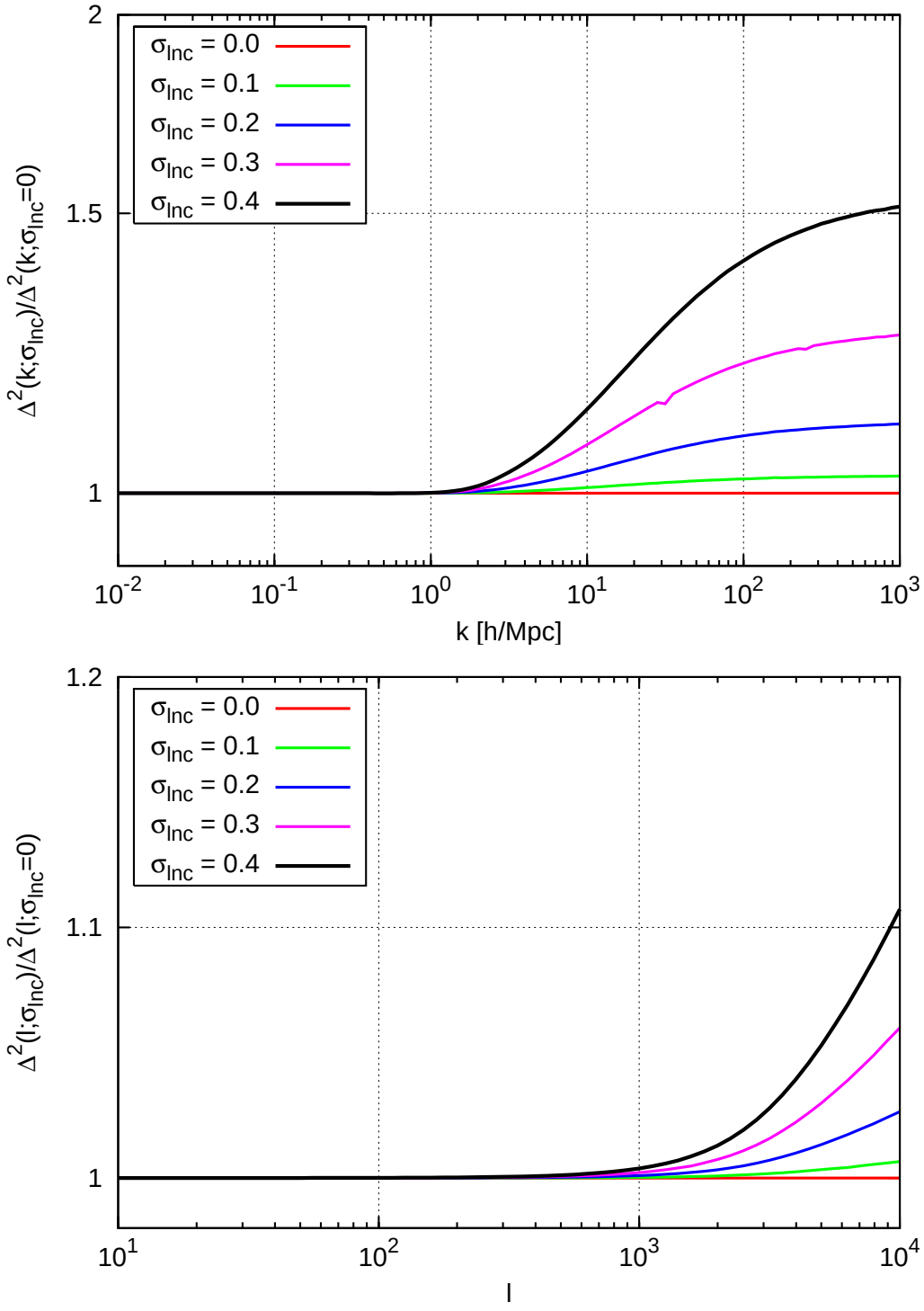

Figure 5.8: *Ratio of the 2-halo term contribution as a function of the concentration dispersion to the 2-halo term contribution from the deterministic concentration relation of the three-dimensional (upper panel) and the convergence trispectrum (lower panel) against wavenumber l. The curves show the ratio for four different dispersions of the halo concentration $\sigma_{\ln c} = \{0.1, 0.2, 0.3, 0.4\}$ and the deterministic concentration denoted by $\sigma_{\ln c} = 0$. The effect of the concentration dispersion on the 2-halo term is less pronounced than for the 1-halo term, but still important to consider for large wave-numbers. Around $l \simeq 10^4$ the deviation from a deterministic concentration relation can amount up to 10%.*

larger than $l \sim 10^3$ and for the 1-halo term of the trispectrum (lower panel, blue line) a deviation in the order of $10\% - 15\%$ in the same l-range. Thus, when considering the diagonal part of the covariance, one should take into account the concentration dispersion in the 1-halo term of the trispectrum but can safely neglect it for the power spectrum. It might be useful to apply the dispersion as well on the 2-halo term of the trispectrum. For a large dispersion and on small scales it can raise the amplitude of the 2-halo term up to 10% in comparison to the same 2-halo contribution with a deterministic concentration.

5.4.3 Projected power spectrum covariance

From the results above, we expect the strongest effect of a concentration distribution on the non-Gaussian part of the covariance since it depends on the bin-averaged trispectrum. To understand the influence of a concentration distribution on the covariance better, we calculate the 1-halo contribution of the non-Gaussian covariance $\mathcal{C}^{1h}$ (see Eq. (5.43)) for four different concentration dispersions $\sigma_{\ln c}$ with our halo model implementation and divide it by non-Gaussian covariance $\mathcal{C}^{1h}$ with a deterministic concentration relation. The cosmological parameters and survey characteristics are chosen such that they match the Millennium Run simulation. The considered wave-number range is from $l_0 = 144$ to $l_{49} = 3672$. Fig. 5.10 displays the results for $\sigma_{\ln c} = \{0.1, 0.2, 0.3, 0.4\}$ in four contour plots. In consistency with the previous results, the largest impact of the concentration dispersion occurs for wave-numbers larger than $l_i \simeq l_j \simeq 2000$. For $\sigma_{\ln c} = 0.1$ we have only a small deviation of the non-Gaussian covariance $\mathcal{C}^{1h}$ from the original deterministic concentration relation of about $1\% - 2\%$. The effect is already stronger for $\sigma_{\ln c} = 0.2$, where $\mathcal{C}^{1h}$ deviates on small scales ($l_i \simeq l_j \gtrsim 3300$) around 6% from a deterministic concentration and is non-negligible for $\sigma_{\ln c} = 0.3$, where we find a deviation from 3% (for $l_i \simeq l_j \gtrsim 1000$) to 12% (for $l_i \simeq l_j \gtrsim 3300$). For $\sigma_{\ln c} = 0.4$ the largest deviation is about 25% for $l_i = l_j \gtrsim 4000$. N-body simulations suggest that concentration dispersions up to $\sigma_{\ln c} = 0.3$ can be expected whereas the last scenario with $\sigma_{\ln c} = 0.4$ is rather unlikely.

5.5 Mode coupling in the power spectrum covariance

In the previous sections we found that the non-Gaussian part of the power spectrum covariance can be described as the bin-averaged trispectrum. In other words, to obtain the entry $C^{\mathrm{NG}}(l_i, l_j)$ we average for a parallelogram configuration of the trispectrum with fixed wave-numbers (l_i, l_j) over all possible angles between the wave-vectors $\boldsymbol{l}_1$ and $\boldsymbol{l}_2$ (see Eq. 5.18). In order to understand the origin of the cross-correlations or *mode coupling* in the power spectrum covariance, we analyze parallelogram configurations of the three-dimensional trispectrum with a fixed ratio of the wave-vectors such that

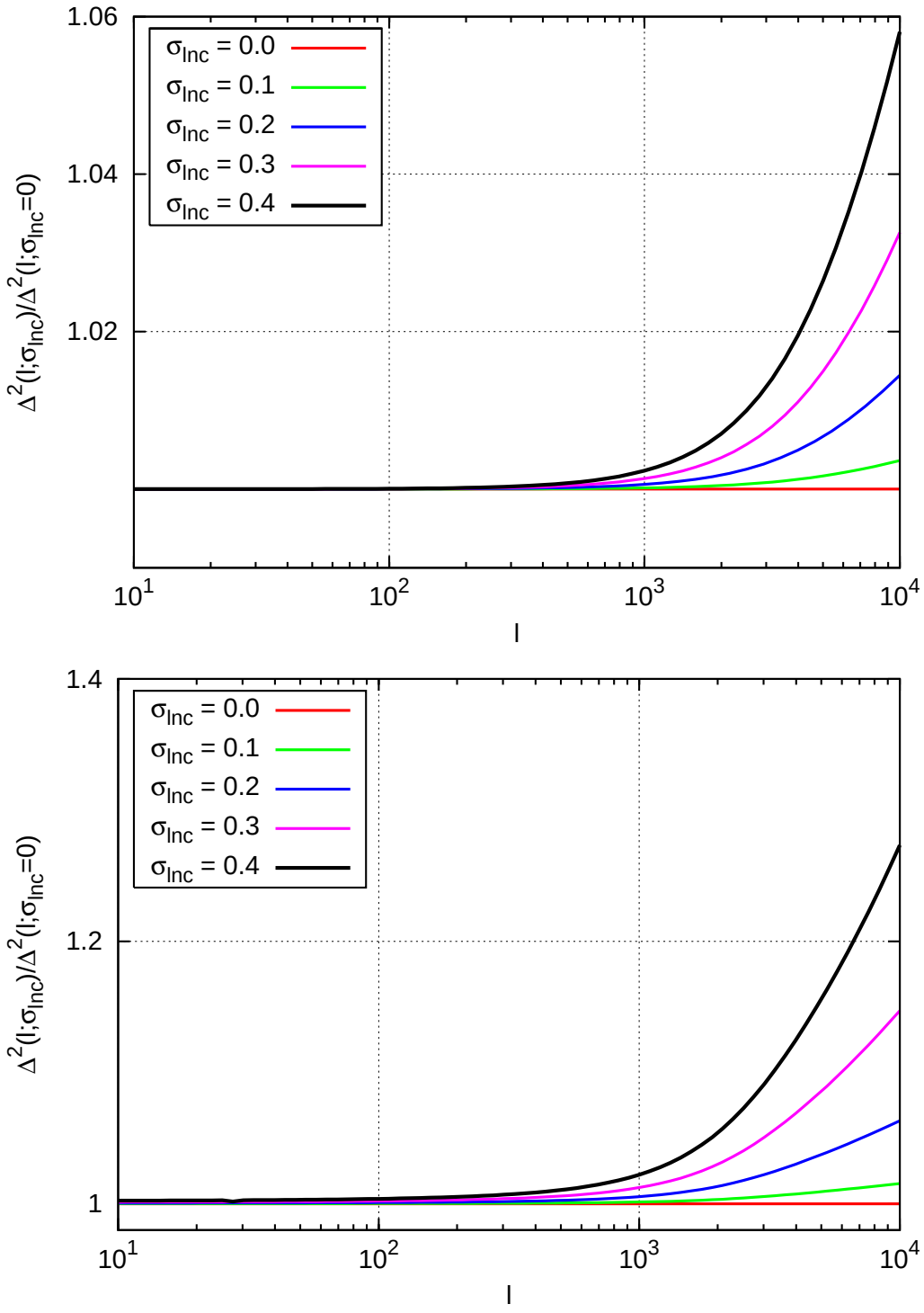

Figure 5.9: *Ratio of the 1-halo term contribution as a function of the concentration dispersion to the 1-halo term contribution from the deterministic concentration relation of different convergence spectra against wave-number l. The curves show the ratio for four different dispersions of the halo concentration $\sigma_{\ln c} = \{0.1, 0.2, 0.3, 0.4\}$ and the deterministic concentration denoted by $\sigma_{\ln c} = 0$. The upper plot displays different concentration dispersions for the projected power spectrum, whereas the lower panel shows the same dispersions for the projected trispectrum. The projection along the redshift-axis diminishes the effect of the concentration dispersion in both spectra. The strongest deviation from a deterministic concentration relation occurs for large wave-numbers. For the projected power spectrum, one can expect a deviation up to 6%, while for the projected trispectrum it differs up to 30%.*

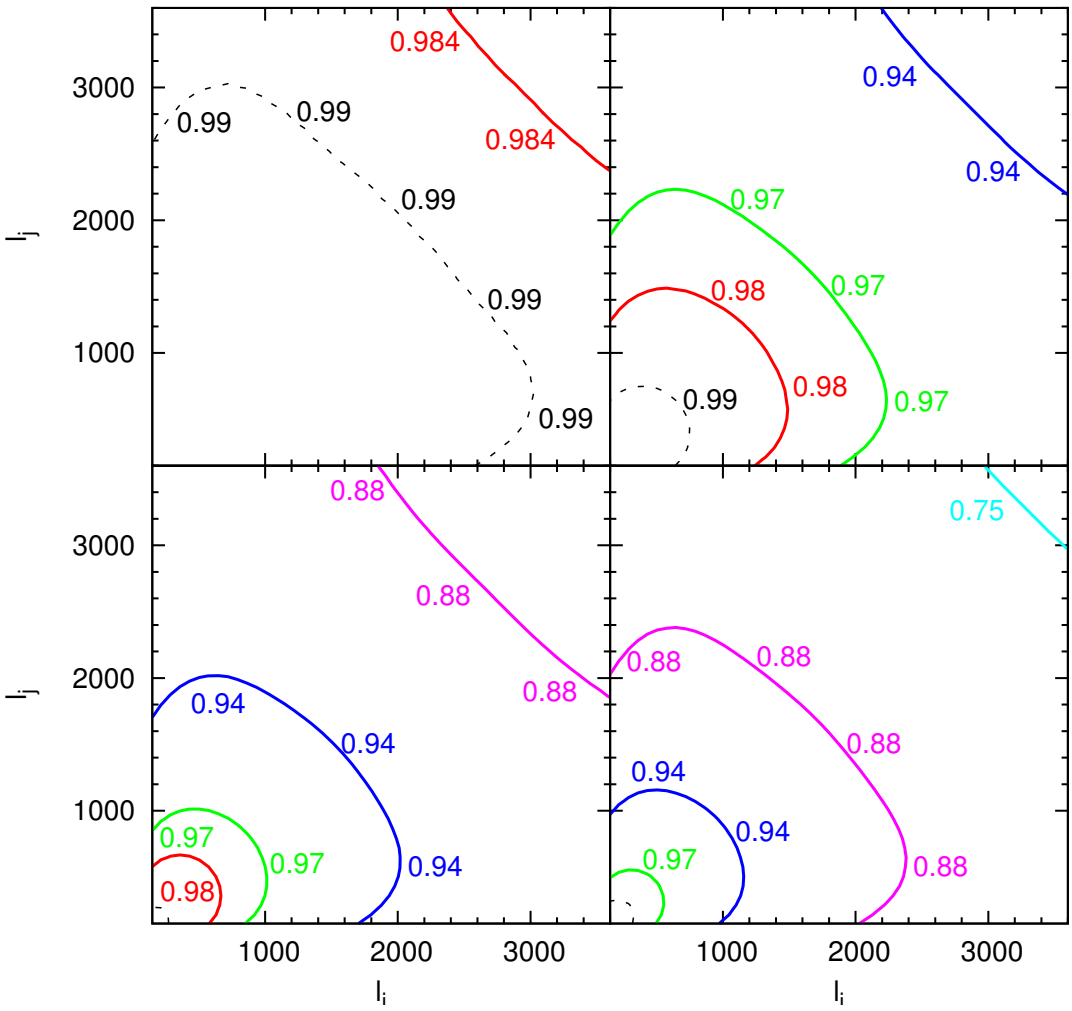

Figure 5.10: *Contour plots of the ratio of the covariance 1-halo term contribution with different concentration dispersions to that with a deterministic concentration against wavenumbers* (l_i, l_j)*. Each panel displays the ratio for a different concentration dispersion: for the upper left panel we use* $\sigma_{\ln c} = 0.1$*, for the upper right panel* $\sigma_{\ln c} = 0.2$*, for the lower left panel* $\sigma_{\ln c} = 0.3$*, while the lower right panel considers* $\sigma_{\ln c} = 0.4$*. The binning scheme is linear from* $l_0 = 144$ *to* $l_{49} = 3672$*.*

$a = k_1/k_2$ for different angles $x \equiv \cos\varphi$. A convenient way to illustrate this dependence is by normalizing the three-dimensional trispectrum according to

$$\beta(k_1, k_2, x) \equiv \frac{T(k_1, k_2, x)}{[P(k_1)P(k_2)]^{3/2}} ,$$

where $P(k_i)$ denotes the three-dimensional power spectrum, which is calculated with the halo model as well. In this way we expect to eliminate the $T(k, k, x) \propto P^3(k)$ scaling on large scales, which is predicted by perturbation theory (see Sect. 3.5.3). Accordingly, we normalize the convergence trispectrum as

$$\beta_\kappa(l_1, l_2, x) \equiv \frac{T_\kappa(l_1, l_2, x)}{[P_\kappa(l_1)P_\kappa(l_2)]^{3/2}} , \tag{5.52}$$

where $P_\kappa(l_i)$ denotes the convergence power spectrum. Since the 1-halo term depends only on the length of the wave-vectors (see Eq. 5.22) and the angular dependence of the 2-halo term is expected to be small as it only enters in the linear power spectrum (see Eq. 5.23), we concentrate on large scales, i.e. small wave-numbers k. In this regime the angular dependent 4-halo term dominates and we expect to see a strong mode coupling. To verify this behavior, we decompose in the upper plot of Fig. 5.11 the trispectrum for a configuration with $k_1 = 0.1\,h\,\mathrm{Mpc}^{-1}$ and $k_2 = 0.2\,h\,\mathrm{Mpc}^{-1}$ into the individual halo terms and consider their angular dependence. As expected the largest correlations arise from the perturbation theory trispectrum or 4-halo term, whereas the 1-halo and 2-halo term stay roughly constant. The maximum correlation is obtained for colinear configurations ($\varphi = 0$) and the smallest contribution comes from perpendicular modes ($\varphi = \pi/2$).

In the lower plot of Fig. 5.11 we display the angular dependence of the full normalized trispectrum for five different configurations. It shows that the amount of correlation at a specific angle depends on the actual lengths of the modes and not only on the wave-number ratio. As we will see later, the behavior of the normalized convergence trispectrum is similar.

5.6 Testing the amount of non-Gaussianity

One reason for considering the full covariance of the power spectrum is to take into account the non-Gaussian terms in form of the trispectrum, which are expected to have a large contribution on small scales according to our understanding of structure formation. For this reason, we determine and analyze the ratio of non-Gaussian to Gaussian contributions to the covariance. Since the off-diagonal entries of the covariance matrix are of purely non-Gaussian origin, we determine the ratio only for the diagonal entries.

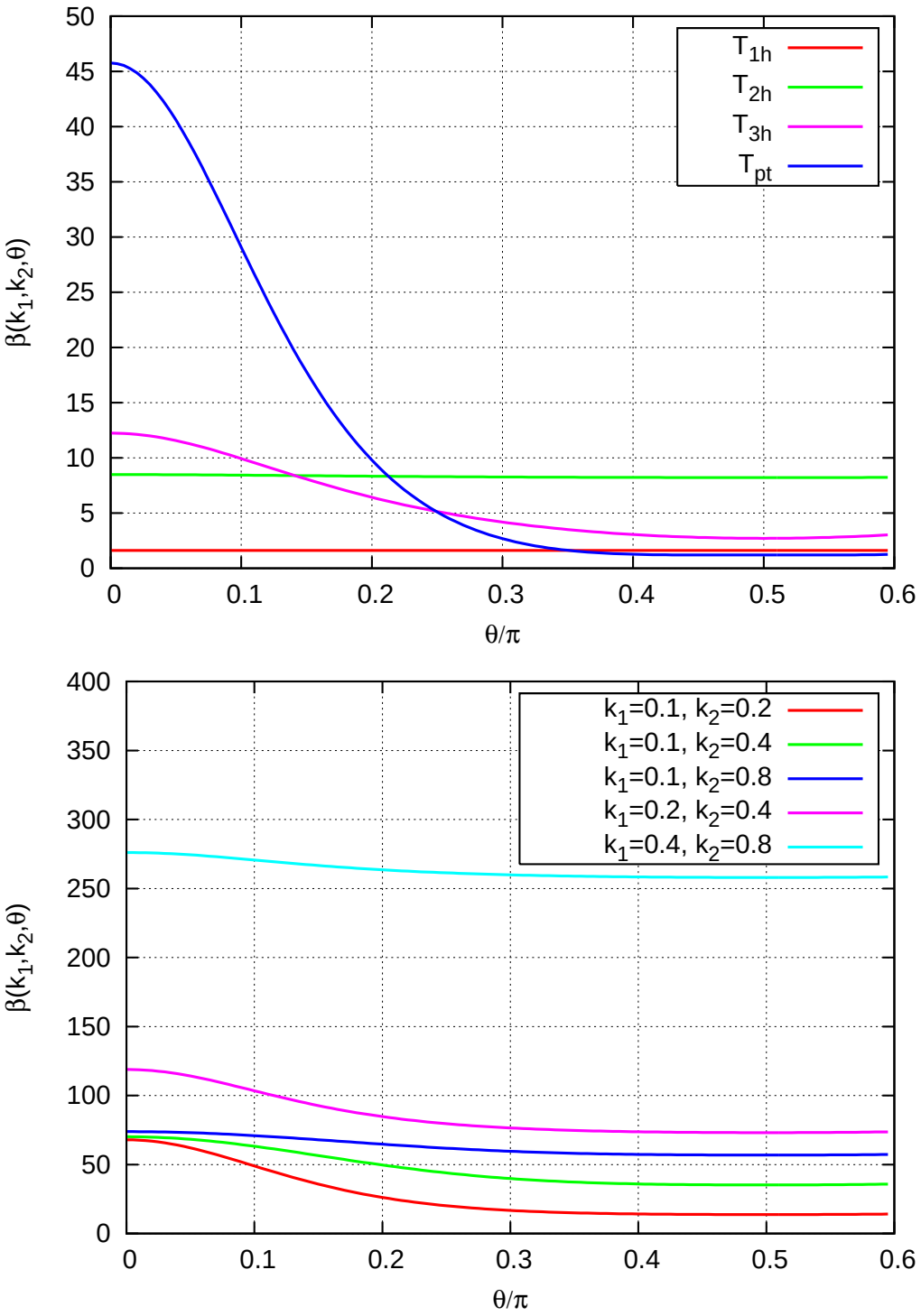

Figure 5.11: *Normalized trispectrum $\beta_{12}(k_1, k_2, \theta)$ against angle θ in terms of π (see text for the exact definition). The upper panel depicts a normalized trispectrum with $k_1 = 0.1\,h\,\mathrm{Mpc}^{-1}$ and $k_2 = 0.2\,h\,\mathrm{Mpc}^{-1}$ decomposed into the different halo terms, whereas the lower panel shows the full normalized trispectrum for different configurations as denoted in the key.*

5.6.1 Three-dimensional case

For the three-dimensional power spectrum covariance of dark matter, we define the ratio of non-Gaussian to Gaussian contribution

$$R_k \equiv \frac{C^{\mathrm{NG}}}{C^{\mathrm{G}}} = \frac{\bar{T}(k,k)}{2P^2(k)} \frac{V_{\mathrm{s}}(k)}{(2\pi)^3}\,, \tag{5.53}$$

where

$$C^{\mathrm{G}} = \frac{1}{V}\,\frac{(2\pi)^3}{V_{\mathrm{s}}(k)}\,2P^2(k) \quad \text{and} \quad C^{\mathrm{NG}} = \frac{1}{V}\,\bar{T}(k,k)\,. \tag{5.54}$$

Note that the ratio is independent of the actual survey size V, whereas it is proportional to the bin-width Δk leading to

$$R_k \approx \frac{\bar{T}(k,k)}{4\pi^2 P^2(k)} \Delta k \cdot k^2\,, \tag{5.55}$$

where we neglect higher order terms in Δk. In order to analyze the 'amount of non-Gaussianity', one has to choose a fixed binning before comparing e.g. results from different simulations. The upper panel of Fig. 5.12 displays the ratio R_k for a linear binning scheme and a bin-size $\Delta k = 2\pi/10\,h\,\mathrm{Mpc}^{-1}$. As expected, the non-Gaussian contribution increases towards smaller scales. On very large scales, the ratio deviates from the expected behavior: perturbation theory predicts the ratio to behave as $R_k \propto P(k)k^2$ for small wave-numbers k, since $\bar{T}(k,k) \propto k^3$. Since $P(k) \propto k$ on large scales, one would expect the ratio to increase with $R_k \propto k^3$. For the two smaller bin-width (green and blue lines) the ratio instead decreases rapidly. This effect has to be analyzed further in future work.

5.6.2 Projected case

In case of the projected power spectrum covariance (see Eq. 5.15), the ratio is

$$R_l \equiv \frac{\mathcal{C}^{\mathrm{NG}}}{\mathcal{C}^{\mathrm{G}}} = \frac{\bar{T}_\kappa(l,l)}{2P_\kappa^2(l)} \frac{A_{\mathrm{r}}(l)}{(2\pi)^2}\,, \tag{5.56}$$

where

$$\mathcal{C}^{\mathrm{G}} = \frac{1}{A}\,\frac{(2\pi)^2}{A_{\mathrm{r}}(l)}\,2P_\kappa^2(l) \quad \text{and} \quad \mathcal{C}^{\mathrm{NG}} = \frac{1}{A}\,\bar{T}_\kappa(l,l)\,. \tag{5.57}$$

Again the ratio is independent of the actual survey size A, but proportional to the chosen bin width Δl, leading to

$$R_l = \frac{\bar{T}_\kappa(l,l)}{4\pi P_\kappa^2(l)} \Delta l \cdot l \tag{5.58}$$

for a linear binning scheme. The lower panel of Fig. 5.12 displays the ratio R_l for a bin-size $\Delta l = 72$. For small wave-numbers the non-Gaussian contribution becomes less important. Contrary to our expectations the non-Gaussian contribution drops also towards larger l with a maximal value around $l = 800$.

In order to understand the shape of the curves, we construct an approximation for the non-Gaussian to Gaussian ratio, based on two simplifying assumptions: we estimate the projected power spectra as defined in Eq. (3.67) by

$$P_\kappa(l) \simeq \Delta w\, G^2(w_*)\, P(l/f_K(w_*), w_*)\,, \tag{5.59}$$

and the bin-averaged, projected trispectrum by

$$\bar{T}_\kappa(\boldsymbol{l}, -\boldsymbol{l}) \simeq \Delta w\, \frac{G^4(w_*)}{f_K^2(w_*)}\, T(\boldsymbol{l}/f_K(w_*), -\boldsymbol{l}/f_K(w_*), w_*)\,, \tag{5.60}$$

where we made use of Eq. (5.16) and T denotes the three-dimensional trispectrum as defined in Eq. (4.133). This corresponds to assuming that the lenses are mainly distributed in the interval $w_* \pm \Delta w$, where w_* is comoving distance to the peak of the lensing window function $G(w)$. Additionally, we neglect the configuration dependence of the trispectrum such that the non-Gaussian contribution is $C^{\mathrm{NG}} \propto T$. Combining these, Eq. (5.56) is approximated by

$$R_l = \frac{1}{\Delta w\, f_K^2(w_*)}\, \frac{T_\kappa(l/f_K(w_*))}{P^2(l/f_K(w_*))}\, \Delta l \cdot l\,. \tag{5.61}$$

One can interpret this as a substitution of the bin-averaged, projected spectra by their three-dimensional counterparts divided by an effective volume $\Delta w\, f_K^2(w_*)$. For large scales perturbation theory then predicts $R_l \propto l^3$. Applying hierarchical clustering on small scales, we expect $R_l \propto l^{-1}$. This corresponds indeed roughly to the behavior of the curves in the lower plot in Fig. 5.12.

5.7 A fitting formula for the lensing power spectrum covariance

A suitable quantity for a fitting formula to the lensing power spectrum covariance as predicted in the halo model approach is the bin-averaged version of the normalized trispectrum $\beta_\kappa(l_1, l_2)$ that we defined in Sect. 5.5. As one can see in Fig. 5.13, it steeply decreases towards large scales, but is well behaved on scales smaller than $l \simeq 1000$. Additionally this quantity is independent of the binning scheme chosen for the observation. Since the Gaussian contribution is bin-dependent instead and only

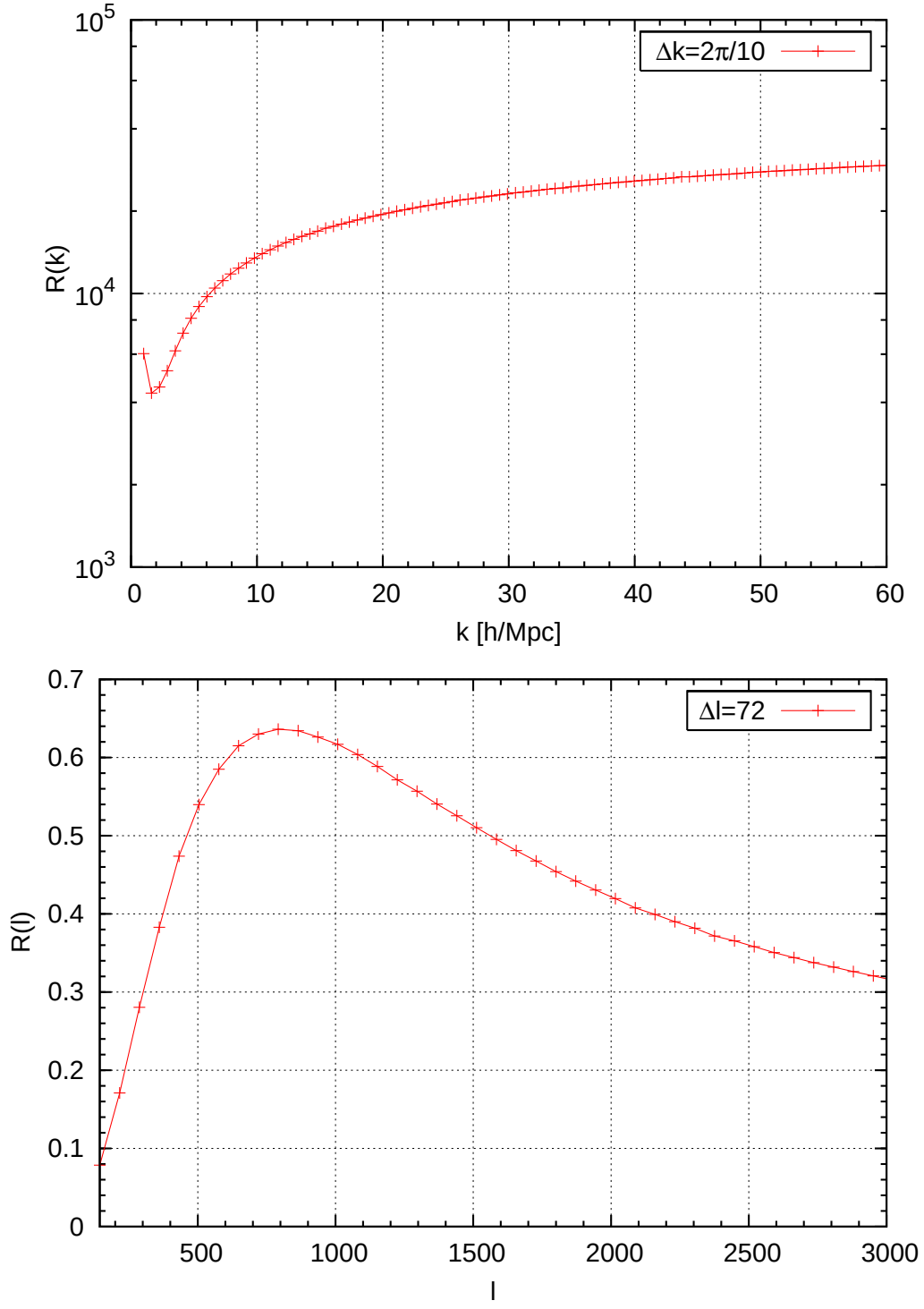

Figure 5.12: *Ratio R of non-Gaussian to Gaussian contribution of the power spectrum covariance against wave-number for different bin-sizes and a linear binning scheme. In the three-dimensional case (upper plot) the non-Gaussian term increases as expected, whereas in the projected case (lower plot) the ratio has a maximum around $l \simeq 800$.*

Table 5.4: *Fiducial cosmological model as used for the fitting formula for the bin-averaged, normalized trispectrum $\bar{\beta}(l_1, l_2)$. The parameter $\sigma_{\ln c}$ denotes the dispersion in the halo concentration distribution as defined in Sect. 5.4. In order to set up the initial power spectrum the Eisenstein-Hu (EH) transfer function was used [22].*

Ω_m	Ω_Λ	h	Ω_b	σ_8	n_s	$\sigma_{\ln c}$	z_s	$T(k)$
0.28	0.72	0.73	0.045	0.82	1.0	0.3	1	EH

contributes to the diagonal part of the covariance, we neglect it in our fitting formula. The contribution can be easily added on top of the fitting results as it depends only on the squared convergence power spectrum. As fitting function we choose a second order polynom in the dimensionless power spectrum, such that

$$\frac{\bar{\beta}(l_1, l_2)}{2\pi^2} = a_0 + a_1\Delta^2_\mathrm{min} + a_2\Delta^2_\mathrm{max} + a_3\Delta^4_\mathrm{min} + a_4\Delta^2_\mathrm{min}\Delta^2_\mathrm{max} + a_5\Delta^4_\mathrm{max}\,, \tag{5.62}$$

where $\Delta^2_\mathrm{max} \equiv \max(\Delta^2(l_1), \Delta^2(l_2))$ and $\Delta^2_\mathrm{min} \equiv \min(\Delta^2(l_1), \Delta^2(l_2))$ denote the dimensionless convergence power spectra. This construction allows us to model the inner 'square shape' (see lower panel of Fig. 5.13) of the normalized trispectrum accurately in the range $1000 \leq l \leq 5000$. This choice of a second order polynom with six free parameters for a fixed cosmological model is a compromise between expressive power and the danger of overfitting.

In order to provide a fitting formula which is valid for different cosmologies, we treat the six fitting parameters as cosmology dependent on Ω_m and σ_8. These are the parameters of special interest for applications in weak lensing. To achieve this with a minimal set of new fitting parameter, we Taylor expand a_k/a_k^fid to first order around a fiducial model, such that

$$a_k(\Omega_\mathrm{m}, \sigma_8) = a_k^\mathrm{fid}\left(1 + b_k(\Omega_\mathrm{m} - \Omega_\mathrm{m}^\mathrm{fid}) + c_k(\sigma_8 - \sigma_8^\mathrm{fid})\right)\,, \quad \text{for} \quad k = 0, \ldots, 5 \tag{5.63}$$

where b_k and c_k take into account the cosmology dependence in Ω_m and σ_8, respectively. As fiducial cosmological model we choose a standard flat ΛCDM model with values that are in agreement with the latest WMAP results [47]. The total amount of fitting parameter thus becomes 18 and consists of $a_0^\mathrm{fid}, \ldots, a_5^\mathrm{fid}, b_0, \ldots, b_5, c_0, \ldots, c_k$. Fig. 5.15 shows that this linear approximation is justified.

The fitting procedure is performed in the following way. We first apply a least square fit to the polynom in Eq. (5.62) for the fiducial model and determine the six parameters a_k^fid. With a second least square fit to equation (5.63), we find the best fit parameters b_k and c_k for different cosmologies varying in the $(\Omega_\mathrm{m}, \sigma_8)$ parameter space. To model the bin-averaged and normalized trispectrum $\bar{\beta}(l_1, l_2)$, we choose a halo model with parameters as described in Sect. 4.7. It takes into account the corresponding 1-halo

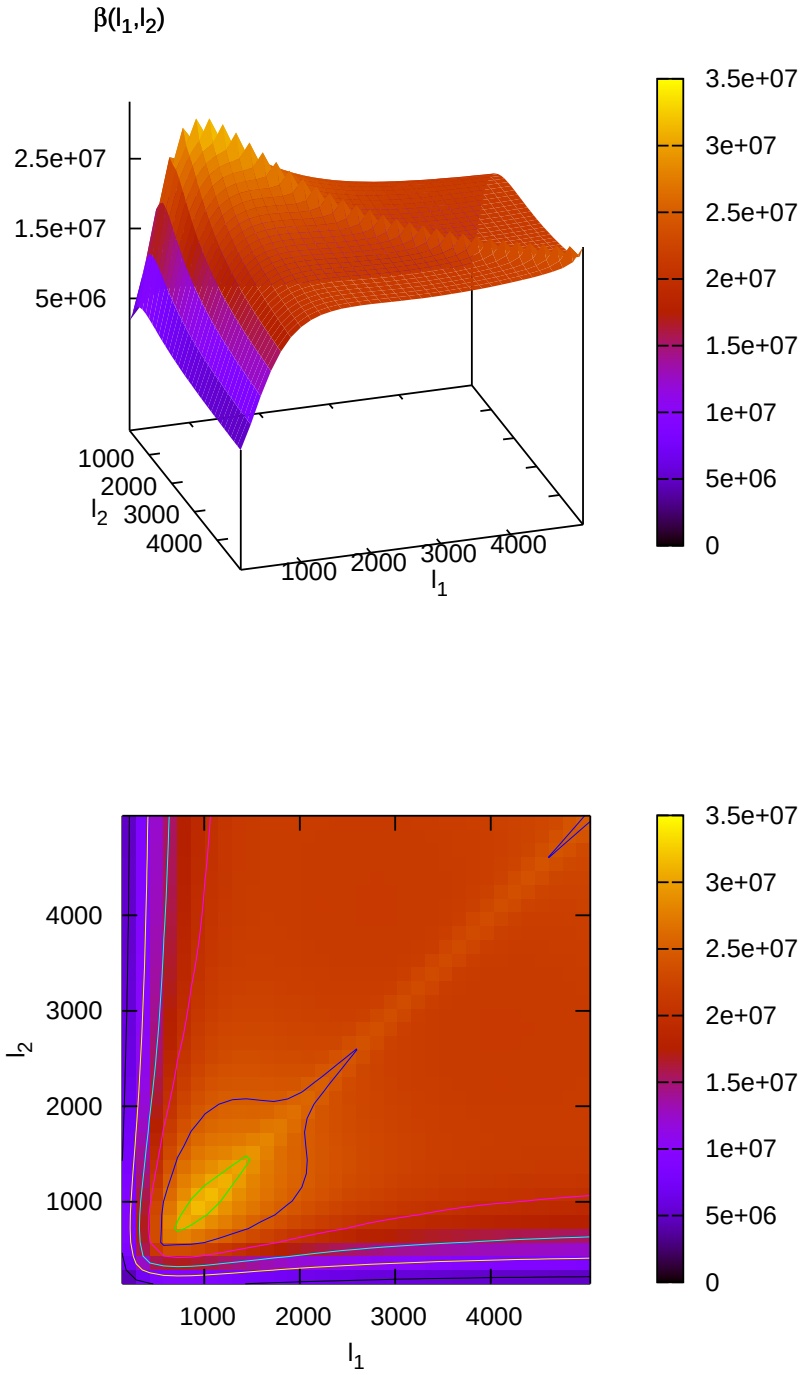

Figure 5.13: *Bin-averaged, normalized convergence trispectrum* $\bar{\beta}(l_1, l_2)$ *against wave-numbers* (l_1, l_2) *as calculated with the halo model for our fiducial* ΛCDM *cosmology with parameters as summarized in Tab. 5.4. The wave-number bins range from* $l_0 = 144$ *up to* $l_{34} = 5040$*. The upper plot shows a three-dimensional visualization of the bin-averaged trispectrum, whereas the lower plot depicts the corresponding contour representation. As can be seen,* $\bar{\beta}$ *decreases steeply towards large scales, but is well behaved on scales smaller than* $l \simeq 1000$*.*

Table 5.5: *Best-fit parameters for a_k^{fid}, b_k, c_k as obtained when fitting the bin-averaged, normalized convergence trispectrum calculated in the halo model approach to a second order polynom in the dimensionless convergence power spectrum.*

k	0	1	2	3	4	5
a_k^{fid}	$4.96 \cdot 10^7$	$-5.54 \cdot 10^{11}$	$-5.59 \cdot 10^{11}$	$2.50 \cdot 10^{13}$	$0.996 \cdot 10^{13}$	$-1.02 \cdot 10^{13}$
b_k	-10.2	-16.2	-24.8	-21.3	-25.6	-36.6
c_k	-1.0	-3.8	18.3	-6.0	-5.4	-102.0

and 2-halo terms and includes a stochastic concentration-mass relation with $\sigma_{\ln c} = 0.3$ for the 1-halo term of the trispectrum. In order to obtain reasonable values for the fitting formula, we calculate the bin-averaged, normalized trispectrum for 25 different cosmological models where $\Omega_{\mathrm{m}} \in [0.26, 0.30]$ and $\sigma_8 \in [0.80, 0.84]$. This covers more than a 3σ deviation from the standard WMAP results. The other parameters are fixed to the fiducial cosmology as summarized in Tab. 5.4. The best-fit parameters as found with this method are listed in Tab. 5.5.

In order to check the performance of the fitting formula, we calculate the relative deviation between fitting formula and halo model results

$$\Delta\bar{\beta}_{12} \equiv \frac{\bar{\beta}_{\mathrm{fit}}(l_1, l_2)}{\bar{\beta}_{\mathrm{halo}}(l_1, l_2)} - 1 \tag{5.64}$$

for every wave-number pair (l_1, l_2). Fig. 5.14 shows the deviation for two different cosmologies. The upper panel depicts the results for our fiducial cosmological model, whereas the lower panel displays the results for a cosmology with $\Omega_{\mathrm{m}} = 0.26$ and $\sigma_8 = 0.84$. Both models are in good agreement with the halo model results. In the fitting region between $1000 < l < 5000$ the deviation amounts less than 10%. Only the diagonal differs on small scales up to 15% from the halo model prediction.

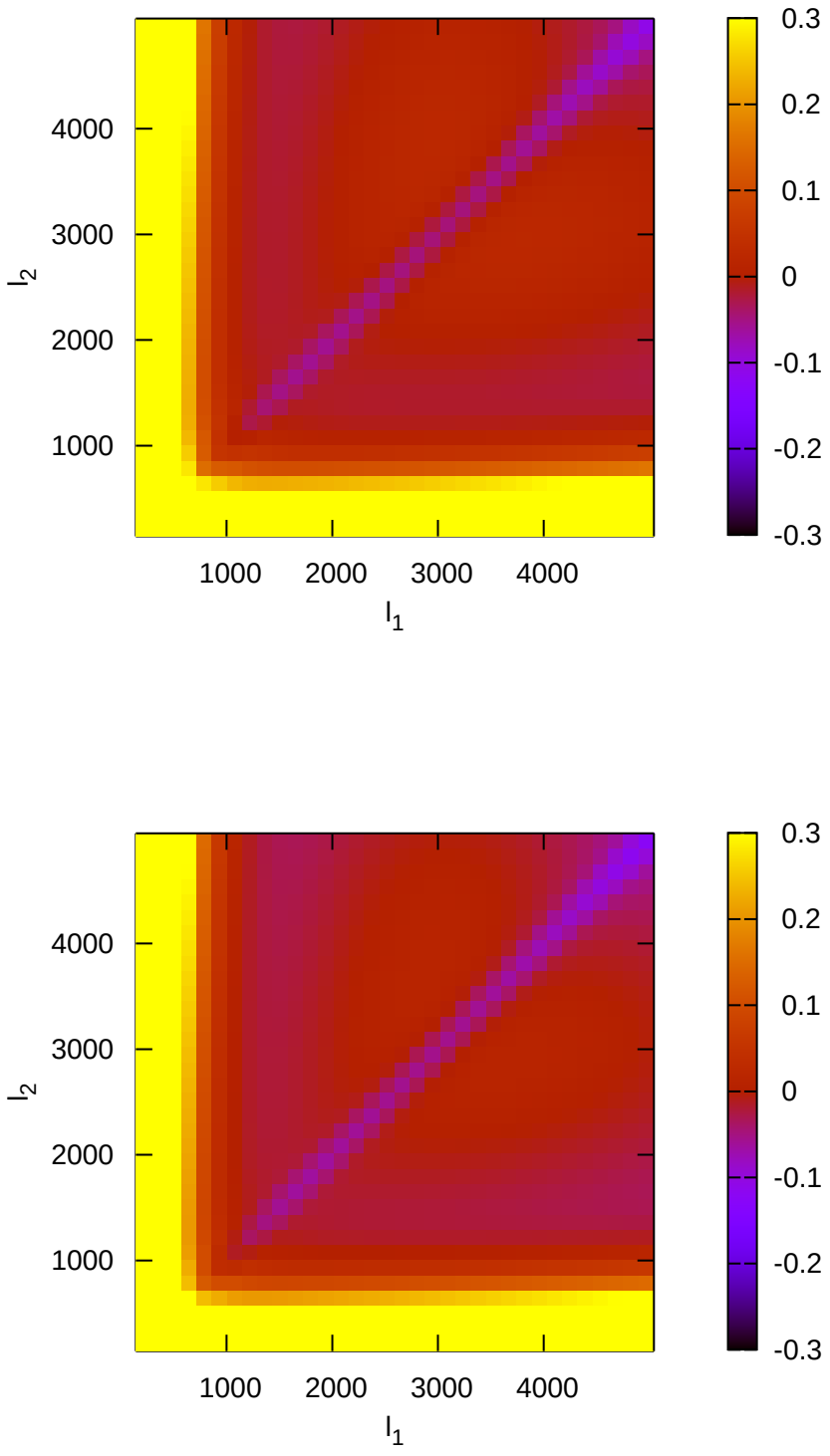

Figure 5.14: *Relative deviation $\Delta\bar{\beta}$ between fitting formula and halo model prediction against wave-numbers (l_i, l_j). The upper panel considers the deviation of the fiducial cosmological model to the halo model; the lower panel is compared to a cosmology which deviates from the fiducial model about 10% in Ω_m and σ_8.*

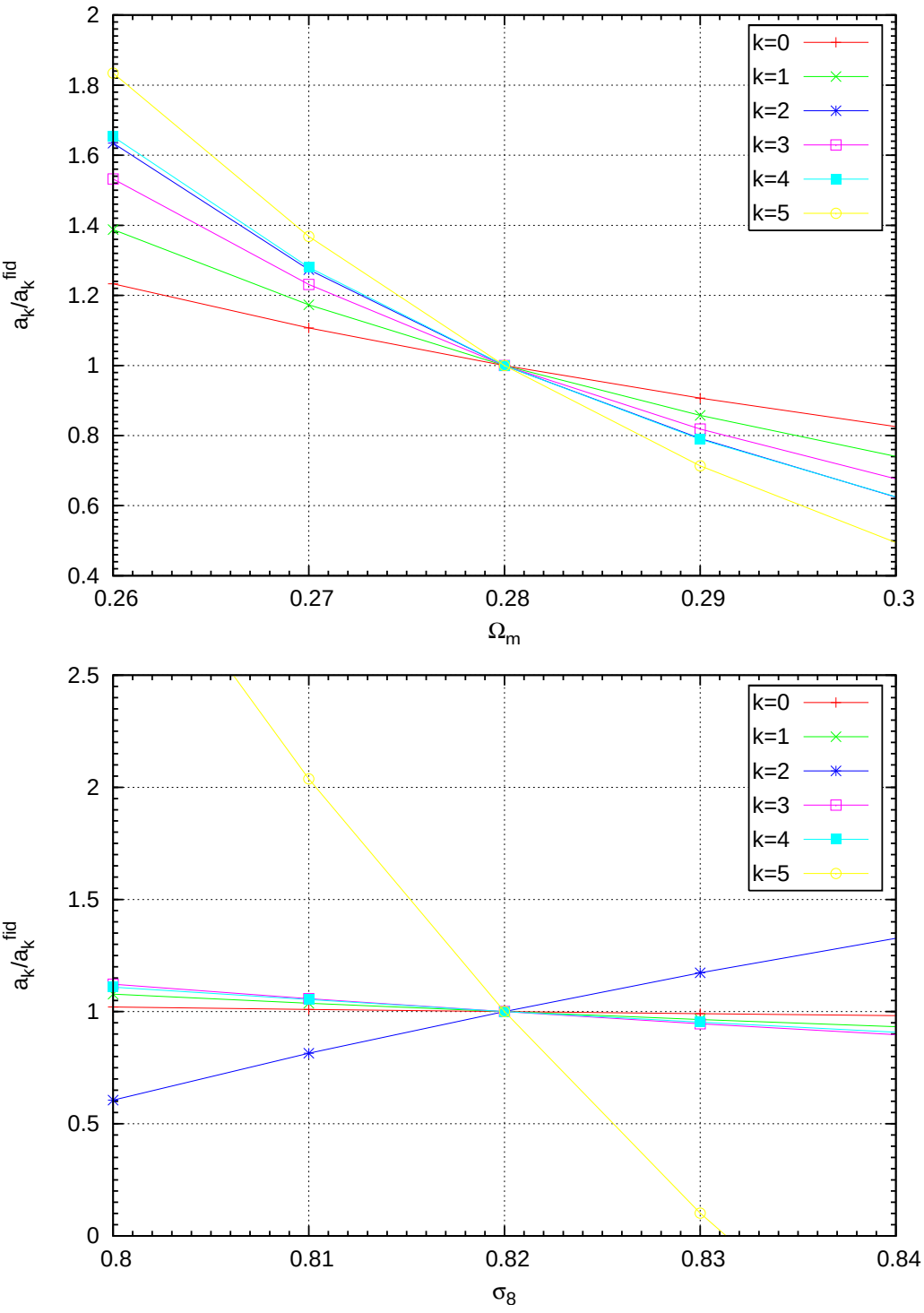

Figure 5.15: *Cosmology dependence of the fitting parameters a_k. The upper panel shows the ratio a_k/a_k^{fid} against Ω_m, whereas the lower panel shows the same ratio against σ_8. The bins correspond to the results obtained from the 25 different cosmologies and justify a linear approximation of the cosmology dependence in the considered parameters.*

Chapter 6

Comparison with N-body simulations

Since the perturbation theory description of structure formation breaks down as soon as the density contrast becomes comparable to 1 (see Sect. 2), numerical N-body simulations have become the theoretical tool of choice to analyze the growth of structure well into the non-linear regime. The underlying idea of these numerical simulations is to discretize the dark matter distribution by considering a sample of $N_{\rm par}$ particles and follow their dynamics in an expanding Universe. Having started with small N-body simulations of only few hundred particles in the 60s and 70s to understand cluster formation [1, 64], current state-of-the-art simulations master the equations of motions of about 10^{10} particles [82]. However, although there has been dramatic improvements in computational efficiency and speed over the last decades, numerical simulations are still very time-consuming and suffer from discreteness effects due to limited mass resolution of the particles used and finite box size. Capturing the evolution of structure formation in an accurate analytic model would allow much more flexibility in investigating observations from large-scale galaxy clustering or weak gravitational lensing. As we have seen in the previous chapters, a promising candidate for such a theoretical description of large-scale structure is provided by the halo model. It offers a good alternative to the available fitting formulae for the power spectrum and is currently tested against higher-order correlation functions from observations and simulations [16, 24, 70, 84].

This chapter aims at comparing the results for the non-Gaussian power spectrum covariance as calculated using the halo model description with that estimated from different types of numerical N-body simulations. In Sect. 6.1, we discuss the basics of N-body simulations and consider their limitations due to discreteness effects. Furthermore, this section summarizes the characteristic parameters of simulations and gives an overview of the different simulations used for this analysis. The actual analysis starts in Sect. 6.2 with a comparison of the convergence power spectrum as predicted from the halo model against the simulations. Sect. 6.3 extends the comparison to the lensing power spectrum covariance. Finally, in Sect. 6.4, we investigate and compare the non-Gaussian-to-Gaussian ratio of the theoretical power spectrum variance with the simulations.

6.1 Basics of N-body simulations

6.1.1 Dynamics of collisionless CDM particles

The most important ingredient for understanding structure formation in an expanding Universe is dark matter. Although the nature of dark matter has not yet been fully understood, observations indicate that they are CDM particles with extremely light masses in comparison to the mass scale of typical galaxies [46]. As we have already seen in Sect. 2, the standard approach to describe the interaction between these particles is by considering them as a pressureless and non-relativistic species. In the analytical approach of perturbation theory, one solves the corresponding set of equations – consisting of Continuity, Euler and Poisson equation Eq. (2.1)-(2.3) – directly for the density and velocity fields of a dark matter fluid. In contrast to this, numerical simulations consider the phase-space distribution function

$$f \equiv f(\boldsymbol{x}, \boldsymbol{p}, t) \tag{6.1}$$

of dark matter, where $\boldsymbol{x}$ is the comoving position and $\boldsymbol{p} = ma^2\dot{x}$ is the momentum and solve the combination of collisionless Boltzmann and Poisson equation instead. The mass density

$$\rho(\boldsymbol{x}, t) = \int f(\boldsymbol{x}, \boldsymbol{p}, t)\, \mathrm{d}^3 p \tag{6.2}$$

can be inferred by integrating the distribution function over the momentum. Since the set of equations poses a high-dimensional problem, N-body simulations solve them by discretizing phase-space with a finite number of N_{par} tracer particles. Starting from initial conditions obtained from linear perturbation theory, numerical simulations follow the trajectory of each particle in phase-space. The most time-consuming part of N-body simulations is the calculation of the gravitational force acting on each particle as it depends on the positions of all particles. The technique applied determines the actual speed and accuracy of a simulation. The simplest and most accurate method is a pairwise force summation over all particles to find the acceleration on one particle (*particle-particle (PP) method*). Modern simulations use more effective ways of solving the N-body problem such as *tree algorithms*, *particle-mesh (PM) methods* or a combination of both.

The outcome of numerical simulations are three-dimensional distributions of N_{par} particles in cubic boxes over a range of redshift values. In order to compare the results e.g. with weak lensing observations, one makes use of the *multiple-lens-plane ray-tracing* algorithm (see e.g. [30, 35]). The basic idea is to introduce a series of lens planes perpendicular to the central line-of-sight of the observer's backward light cone. The matter distribution within the light cone is sliced by this and can be projected

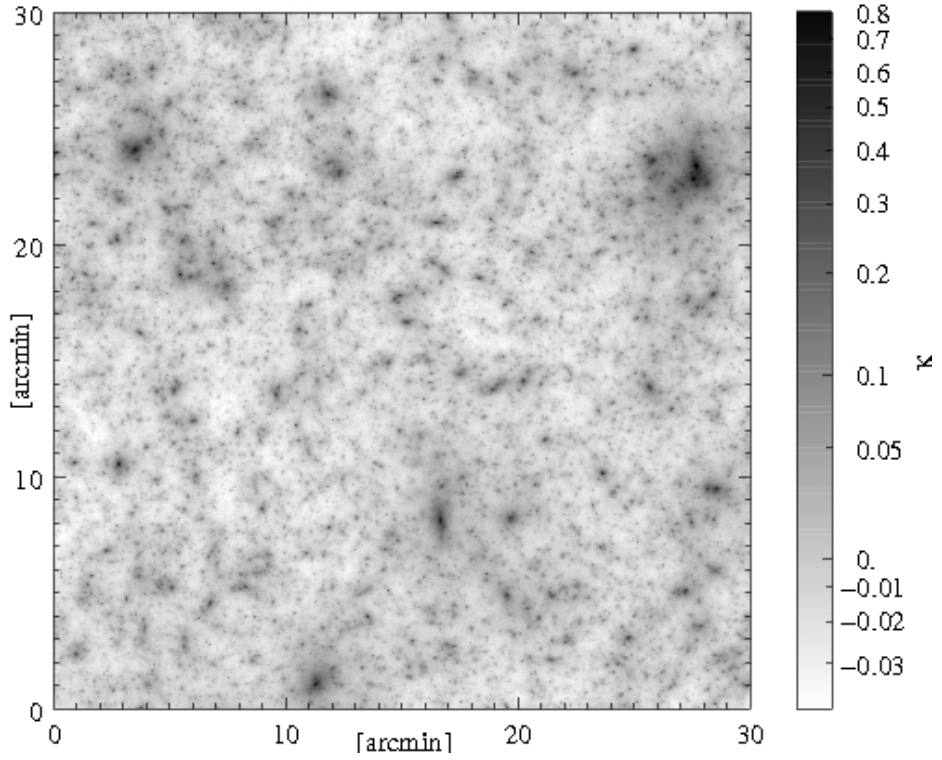

Figure 6.1: *Example of an effective convergence map as obtained from ray-tracing through the Millennium Run simulation (Source: S. Hilbert, J. Hartlap). Dark areas correspond to overdense regions, whereas light areas correspond to slightly underdense regions.*

on the corresponding lens plane. By computing the deflection of light rays and its derivatives at each lens-plane, one simulates the photon trajectory from the observer to the source and also keeps track of the distortions of ray bundles. In this way, the continuous deflection of light rays experienced while propagating through the matter inhomogeneities in the light cone is approximated. As a result, one obtains the Jacobian matrix for the lens mapping from source to observer and can construct so-called *effective convergence maps* (κ-maps) as illustrated in Fig. 6.1. Comprehensive reviews on numerical simulations and the most common techniques to calculate particle accelerations efficiently are e.g. provided by Bertschinger [7], Klypin [45] or Dolag et al. [19]. A detailed description of the ray-tracing technique can be found in Hartlap [30], Jain et al. [35].

6.1.2 Summary of important parameters

Apart from the actual technique used to calculate the gravitational force acting on each dark matter particle (see e.g. Hockney and Eastwood [33]), the outcome of each N-body simulation depends on a number of parameters. We distinguish between three types of quantities:

- **Cosmological parameters:** They consist of the usual set of parameters that determine a cosmological model: $\Omega_{\rm m}$, Ω_Λ, $\Omega_{\rm b}$, h, σ_8, $n_{\rm s}$ and Γ. Together with the transfer function $T(k)$ they determine the initial power spectrum used to set up the initial Gaussian distribution of dark matter particles in the simulations (see Chapter 1 for a discussion of these parameters).

- **Simulation parameters:** The *number of dark matter particles* $N_{\rm par}$ and the *side length* $L_{\rm box}$ of the volume modeled determine the basic setup of an N-body simulation. In order to consider a representative sample of the Universe the side length of the cube should be at least $100 - 200\ h^{-1}$Mpc. Above these scales no structure has been observed which could interact with structure on smaller length-scales, i.e. in a usual ΛCDM Universe we expect no significant power on scales larger than this. Typically, the choice for the size of the box simulated and the number of particles are limited by the computer's memory and the computation time available.

- **Ray-tracing parameters:** They are important to obtain the effective convergence maps (κ-maps), which we use to estimate the projected power spectrum. First of all we need to know how the sources are distributed along the line-of-sight to the observer to perform the actual projection along the redshift-axis. This is specified by a weight function $W(w_{\rm s})$ and the redshift $z_{\rm s}$ where the majority of sources is situated (see Sect. 3.6). In order to process the ray-tracing data one introduces a grid with $N_{\rm bin}^2$ points. This discretization restricts wave-numbers to integer multiples of the *Nyquist frequency*

$$l_{\rm Ny} = \frac{\pi}{\Delta x}\,, \tag{6.3}$$

where $\Delta x = L_{\rm map}/N_{\rm bin}$ is the mesh spacing and $L_{\rm map}$ the *length of the convergence map*. Wave-numbers smaller than the Nyquist frequency cannot be resolved properly by the grid under consideration. Additionally, one expects numerical smoothing effects on scales of the order of twice the mesh spacing. This affects the power spectrum on scales smaller than $l \simeq 10^4$, as was shown by Jain et al. [35]. If the ray-tracing is done along different light cones, one obtains $N_{\rm maps}$ different effective convergence maps. These can be considered as quasi-independent from each other, since each light cone considers usually a fractional part of the simulation volume. Averaging over the estimates obtained from the different κ-maps allows one to find a better estimate for the power spectrum with a reduced variance. For a detailed description of the ray-tracing technique see e.g. Hartlap [30].

6.1.3 Limitations

N-body simulations model the evolution of the collisionless dark matter fluid in the Universe with a set of discrete point particles in a periodic box. Obviously, this is only a coarse approximation of reality and results in a number of discreteness effects one has to take into account when comparing the results with observations and interpreting them. The most important limitations are listed in the following.

Mass resolution

The comoving side length $L_{\rm box}$ of the simulation box and the number of particles $N_{\rm par}$ determine the actual mass of each particle, i.e. the *mass resolution* of the simulation. It is then fixed to

$$m_{\rm par} = \frac{\bar{\rho} L_{\rm box}^3}{N_{\rm par}} \,, \tag{6.4}$$

where $\bar{\rho} \equiv \Omega_{\rm m} \rho_{\rm crit}$ is the comoving average mass density of the Universe. Typically, the mass resolution varies from $10^5\,h^{-1}\,M_\odot$ up to $10^{10}\,h^{-1}\,M_\odot$ depending on the size of objects one wants to consider. Furthermore, this sets the mass limit when objects can be considered as sufficiently resolved. Assuming that 100 particles are necessary for resolving a halo accurately, only objects larger than $M_{\rm min} \approx 100\ m_{\rm par}$ are considered to have a sufficiently high resolution.

Sampling variance

The side length of the cubical simulation box is equivalent to the maximum wavelength a simulation can consider. As a consequence, modes with wavelengths comparable to the box side length $L_{\rm box}$ are only poorly represented. This results in a large *sampling variance* on the corresponding length scales. The sampling variance corresponds to the *cosmic variance* considered for surveys. In Fourier space this limitation sets the smallest wave-number one can consider. The resulting *fundamental mode* is then given by

$$k_{\rm min} = \frac{2\pi}{L_{\rm box}} \,, \tag{6.5}$$

which corresponds to the Fourier counterpart of $L_{\rm box}$.

Shot noise

The discreteness of the mass particles leads to *shot noise* $P_{\rm shot}$, which is especially severe when only few particles describe a region of interest. When estimating the power

spectrum this effect can be mitigated by subtracting a constant term $P_{\rm shot} = 1/\bar{n}$, where $\bar{n}$ denotes the mean number density of particles in the box [30].

Force softening

N-body simulations represent the dark matter fluid as a discrete number of very massive particles. As a result Newton's law predicts strong collisions between particles more frequently than for less massive particles which produces undesirable artifacts in the simulations. Increasing the number of particles would decrease the abundance of strong collisions, but is usually not possible to the extent needed due to limited computer capacity. Thus, in order to correct for the artifacts produced by strong collisions, one modifies Newton's force equation below a certain length-scale, which is called the *softening length* $l_{\rm soft}$. If the separation between the particles corresponds to the softening length Newton's force goes to 0. On length-scales below $l_{\rm soft}$, this modification results in a deviation from the actual $\propto 1/r^2$-behavior of the gravitational force. Structure formation below this scale is inhibited and results from simulations are not reliable below $2-3l_{\rm soft}$. The actual choice for the softening length depends on other simulation parameters, mainly the number of particles $N_{\rm par}$ and the time-step for the integration over the equations-of-motion [33]. In general, a larger number of particles requires a smaller softening length.

6.1.4 Overview of the simulations

In order to have representative results for our comparison of the halo model with N-body simulations and to understand the effect of the simulation and ray-tracing parameters, we cover a wide spectrum of simulations with different box sizes, number of particles and number of available convergence maps. All simulations consider cosmologies close to the standard ΛCDM model consistent with the WMAP observation [80] and assume single redshift sources, unless stated otherwise.

In the following, we briefly summarize the properties of the simulations used for the subsequent analysis. The relevant parameters for each simulation are listed in Tab. 6.1 and 6.2.

- **Virgo (1997):** This simulation was carried out by the Virgo-Consortium for a ΛCDM cosmology with $N_{\rm par} = 256^3$ particles in a periodic box of side length $L_{\rm box} = 141.3\,h^{-1}\,$Mpc. It uses the PP-/PM-code HYDRA, which places subgrids of higher resolution in highly clustered regions. Structures on scales larger than $2l_{\rm soft} \approx 40h^{-1}$kpc can be considered as well resolved. Details about the simulation and results are published in Jenkins et al. [36].
- **VLS (2001):** The Very Large N-body simulation (VLS) was carried out by the Virgo-Consortium and performed using a parallel P^{3}M-code [33]. It employs 512^3

dark matter particles in a box of side length $L_{\rm box} = 479\ h^{-1}$Mpc. Structure on scales larger than $2l_{\rm soft} \approx 30h^{-1}$kpc are well resolved. Details on the ray-tracing techniques used to produce the convergence maps can be found in Ménard et al. [50].

- **Borgani (2004):** This hydrodynamical simulation of a standard ΛCDM cosmology uses the GADGET-2 code. The side length of the box is $L_{\rm box} = 192h^{-1}$Mpc and the simulation is performed with 480^3 dark matter particles and the same number of gas particles. The gravitational softening is set to $l_{\rm soft} = 7.5\,h^{-1}$ kpc. Details about the simulation and results are published in Borgani et al. [10] and Pace et al. [59].

- **Millennium Run (2005):** The Millennium Run simulation is the largest N-body simulation performed up to the present day with $N_{\rm par} = 2160^3$ particles and a box side length $L_{\rm box} = 500\,h^{-1}$ Mpc (Springel et al. [82]). It uses a specially customized version of the GAGDET-2 code (Springel [81]) and applies the TreePM method to evaluate gravitational forces. The calculation was performed on 512 processors of an IBM p690 parallel computer and produced in 28 days an amount of 20 TB data. The particle resolution is $m_{\rm par} = 8.6 \times 10^8 h^{-1} M_\odot$ and structures down to $2l_{\rm soft} \approx 10h^{-1}$kpc are well resolved in high density regions.

- **Grossi (2007):** This simulation was set up to investigate the effect of deviations from non-Gaussian initial conditions in a ΛCDM cosmology (Grossi et al. [27]). It uses the GADGET-2 code in a cubic box of side length $L_{\rm box} = 500\ h^{-1}$Mpc with 800^3 dark matter particles. The softening length is $2l_{\rm soft} \approx 40h^{-1}$kpc.

- **Gems (2007):** The setup of this simulation is a cubic volume of side length $L_{\rm box} = 150h^{-1}$Mpc with 256^3 particles. The cosmology chosen reflects the WMAP5 results [47] and thus has a slightly smaller value for σ_8 than the other simulations. It uses the GADGET-2 code and has a softening length of $2l_{\rm soft} \approx 30h^{-1}$kpc.

All simulations were provided in form of convergence power spectra. For the Borgani and Millennium Run simulations, we additionally obtained κ-maps.

6.2 Projected power spectrum

As a first test of how well the halo model approach describes the non-linear evolution of dark matter, we compare the projected power spectra predicted by the halo approach to the ones found from numerical N-body simulations and to the fitting formulae by Peacock and Dodds [62] and Smith et al. [79].

Table 6.1: *Parameters used for generating the simulations and producing the resulting convergence maps. A description of these parameters can be found in Sect. 6.1.2. More details on the simulations used for this work are given in Sect. 6.1.4.*

Simulation	$L_{\rm box}/\,h^{-1}\,{\rm Mpc}$	$N_{\rm par}$	$m_{\rm par}/\,h^{-1}\,M_\odot$	$N_{\rm map}$	$A_{\rm map}/({\rm deg})^2$
Virgo	141.3	256^3	1.4×10^{10}	200	0.25
Gems	150.0	256^3	1.4×10^{10}	220	16.00
Borgani	192	480^3	6.6×10^9	60	13.10
VLS	479.0	512^3	6.9×10^{10}	36	11.62
Grossi	500.0	800^3	2.033×10^{10}	60	34.57
Millennium	500.0	2160^3	8.6×10^8	20	25.00

Table 6.2: *Cosmological parameters used to set up the initial power spectrum, which determines how the simulation particles are distributed initially. The simulations employ either the BBKS or the Eisenstein-Hu (EH) transfer function. For the Borgani et al. simulation the CMBFAST programme was applied to find a transfer function. See Sect. 2.7 for a description of the different transfer functions.*

Simulation	$\Omega_{\rm m}$	Ω_Λ	h	$\Omega_{\rm b}$	σ_8	n_s	Γ	$z_{\rm s}$	$T(k)$
Virgo	0.3	0.7	0.7	0.0	0.9	1.0	0.21	1 (2)	BBKS
Gems	0.25	0.75	0.7	0.04	0.78	1.0	0.14	1 (2)	EH
Borgani	0.3	0.7	0.7	0.04	0.8	1.0	0.17	1.5	BBKS
VLS	0.3	0.7	0.7	0.04	0.9	1.0	0.17	0.97	CMBFAST
Grossi	0.3	0.7	0.7	0.04	0.9	1.0	0.17	4	BBKS
Millennium	0.25	0.75	0.73	0.045	0.9	1.0	0.14	1 (2)	EH

6.2.1 Estimating the power spectrum from simulations

The output of N-body simulations are cubical volumes of discrete matter particles distributed in a way that reflects the dark matter distribution of the Universe at a certain redshift z. With the help of ray-tracing techniques it is possible to project the matter between observer and distant sources along the line of sight. As a result, one obtains a two-dimensional discrete convergence map of side length L_{map} on a $N_{\mathrm{bin}} \times N_{\mathrm{bin}}$ grid as illustrated in Fig. 6.1. To estimate the convergence power spectrum from this map, we transform the continuous estimator from Eq. (5.14) into its discrete version applying the following procedure:

1. Perform a two-dimensional discrete Fourier decomposition of the convergence map such that
$$\kappa(\boldsymbol{\theta}) = \frac{1}{A_{\mathrm{map}}} \sum_{\boldsymbol{l}} \tilde{\kappa}_{\boldsymbol{l}} \, \mathrm{e}^{\mathrm{i}\boldsymbol{l}\cdot\boldsymbol{\theta}} \,, \tag{6.6}$$
where the summation runs over the combination of integers (n_x, n_y) for $\boldsymbol{l} = l_{\mathrm{min}}(n_x, n_y)$ with $l_{\mathrm{min}} = 2\pi/L_{\mathrm{map}}$ corresponding to the fundamental mode.[1]
2. Calculate the squared absolute value of $\tilde{\kappa}_{\boldsymbol{l}}$ at each coordinate (n_x, n_y).
3. Sort the squared $\tilde{\kappa}$ according to their wave-number length $|\boldsymbol{l}|$ given as
$$l = \sqrt{n_x^2 + n_y^2} \cdot l_{\mathrm{min}} \,. \tag{6.7}$$
All squared Fourier modes with a length $l - \Delta l/2 \leq |\boldsymbol{l}| \leq l + \Delta l/2$ are combined to form one wave-number band l_b of width Δl. In this way, one obtains a reasonable number of Fourier modes to average over for the estimate at each wave-number l. Geometrically, this is equivalent to partitioning the Fourier transformed convergence maps into (quarter) annuli of bin-width Δl.
4. Average over all Fourier modes belonging to one wave-number band l_b such that
$$\hat{P}_\kappa(l) = \frac{1}{A_{\mathrm{map}} N_p(l)} \sum_{\boldsymbol{l}; l \in l_b} \tilde{\kappa}_{\boldsymbol{l}} \tilde{\kappa}_{-\boldsymbol{l}} \tag{6.8}$$
where $N_p(l)$ denotes the number of modes in each band l_b to be averaged over. The result is the desired estimate of the power spectrum from a given convergence map.

Usually a certain number of convergence maps N_{maps} are available from one simulation and the described procedure is applied to each of them. In order to obtain a less noisy estimate for the power spectrum, one averages afterwards for each bin over all N_{maps}. Since the κ-maps are usually non-periodic, the estimator is on small scales additionally affected by *aliasing* [33, 39].

[1] The fastest way to perform the transformation to Fourier space is by using a *Fast Fourier Transformation* (FFT), which we apply for our calculation.

6.2.2 Results

In Figures 6.2-6.6 we illustrate the comparison of the lensing power spectra as predicted by the halo model approach (black lines) with the ones found from N-body simulations (red lines with errorbars). Additionally, we compare both results to the standard fitting formulae of Peacock and Dodds [62] and Smith et al. [79] (green and blue lines). The plots show the results for the different simulations as described in Sect. 6.1.4 with sources situated at single redshift planes between $z_s = 1 - 4$. In order to see the deviations from the simulations more clearly, we use for our comparison the dimensionless convergence power spectrum as defined in Eq. (3.70). The error bars correspond to the dispersion due to averaging over the $N_{\rm map}$ power spectrum estimates from the convergence maps and increase on large scales due to the *sampling variance* (see Sect. 6.1.3). Since numerical artifacts from smoothing on the lens planes, aliasing and shot noise become the dominant effect on scales smaller than $l \simeq 10^4 - 10^5$ [30, 35], we compare the halo model prediction with the numerical results only for wave-numbers up to $l_{\rm max} = 10^4$. Above this wave-number, the ray-tracing simulations at hand cannot be considered as reliable [35].

Comparison with the Halo Model

Altogether, the plots illustrate a good agreement between the halo model and N-body simulation convergence power spectra. Comparing the agreement of the same simulation for different source redshifts z_s with the halo prediction indicates a better correspondence for low redshifts (see e.g. Fig. 6.6). Deviations of the halo model approach to the simulations can either be found on large scales for the Borgani et al. and Grossi et al. simulations (see Fig. 6.4), where the *sampling variance* of the simulations is very large or, for most simulations, on small scales, where numerical effects as shot noise or force softening become important (see Sect. 6.1.3). The deviation on small scales are strongest for the Borgani and Grossi simulations. Since the drop in power occurs rather abruptly at a wave-number around $l \simeq 7 \cdot 10^3$, it is likely that for these simulations the numerical effects become important earlier. The overall good correspondence of the results is not too surprising since ingredients of the halo model as the mass function and the halo profile were adopted from N-body simulations and tuned to fit the simulation power spectrum.

Concentration-mass relation

Since recent numerical simulations suggest a scatter of the concentration parameter for halos of the same mass [38, 86], we additionally consider a dispersion in the concentration parameter ($\sigma_{\ln c} = 0.3$) for the 1-halo term of the power spectrum and compare it with the simulations (see also Sect. 5.4). The corresponding halo model prediction for the convergence power spectrum is denoted by HM2 in Figs. 6.2-6.6. The plots

show that the prediction of the halo model with a probabilistic halo concentration deviates from our fiducial halo model (HM), as expected, only on small scales. In general the deviation is rather small and occurs on scales where numerical effects become important. So there is currently no reason to prefer, in case of the convergence power spectrum, a probabilistic concentration relation over a deterministic one. Nevertheless, we expect that this effect is important for the consideration of the non-Gaussian part of the power spectrum covariance, since the largest contribution to this term comes from the 1-halo term of the trispectrum.

For the deterministic concentration-mass relation, we have applied so far a parametrization as suggested by Takada and Jain [84] with $c_* = 10$ and $\alpha = 0.2$ in Eq. (4.118). Recent studies of the Millennium Run simulations confirm the monotonic decline of the concentration parameter with mass but suggest a different choice of parameters where $c_* = 11$ and $\alpha = 0.1$ (see Neto et al. [57]). For this reason, we compare for the Millennium Run simulation (Fig. 6.6) the Takada and Jain choice of parameters (denoted by HM) with the choice of Neto et al. (denoted by HM3). Obviously, the new choice of parameters for the concentration-mass relation describes the numerical simulation much better. In the following we therefore adopt the choice of parameters as suggested by Neto et al., i.e. $c_* = 11$ and $\alpha = 0.1$.

Comparison with fitting formulae

The comparison of the simulation with the fitting formulae clearly favor the Smith et al. predictions (Smith, green line) over the one from Peacock and Dodds (PD, dark blue line). In most cases, the Smith et al. estimate is tantamount to the halo model prediction, whereas the PD estimate has – especially on smaller scales – often too little power. In case of the Gems and the Grossi simulation, the Smith et al. prediction is better on intermediate scales which hints that the halo model has a problem on these scales. Indeed the strongest deviation between the Smith et al. fitting formula and the halo model occurs around the scale where the 1-halo and 2-halo term of the power spectrum are of the same size. So the halo model may provide an inaccurate description of the intermediate scales, at the transition between small and large-scale behavior.

6.3 Projected power spectrum covariance

The good resemblance of the simulation and halo model power spectrum was to some extent expected. The real challenge for the halo model is a good description of higher-order correlations. It proposes the actual test of how well the halo model describes the underlying physics of structure formation. Since we are interested in estimating the error of the power spectrum and using it for parameter estimates, we focus in this section on the covariance of the projected power spectrum.

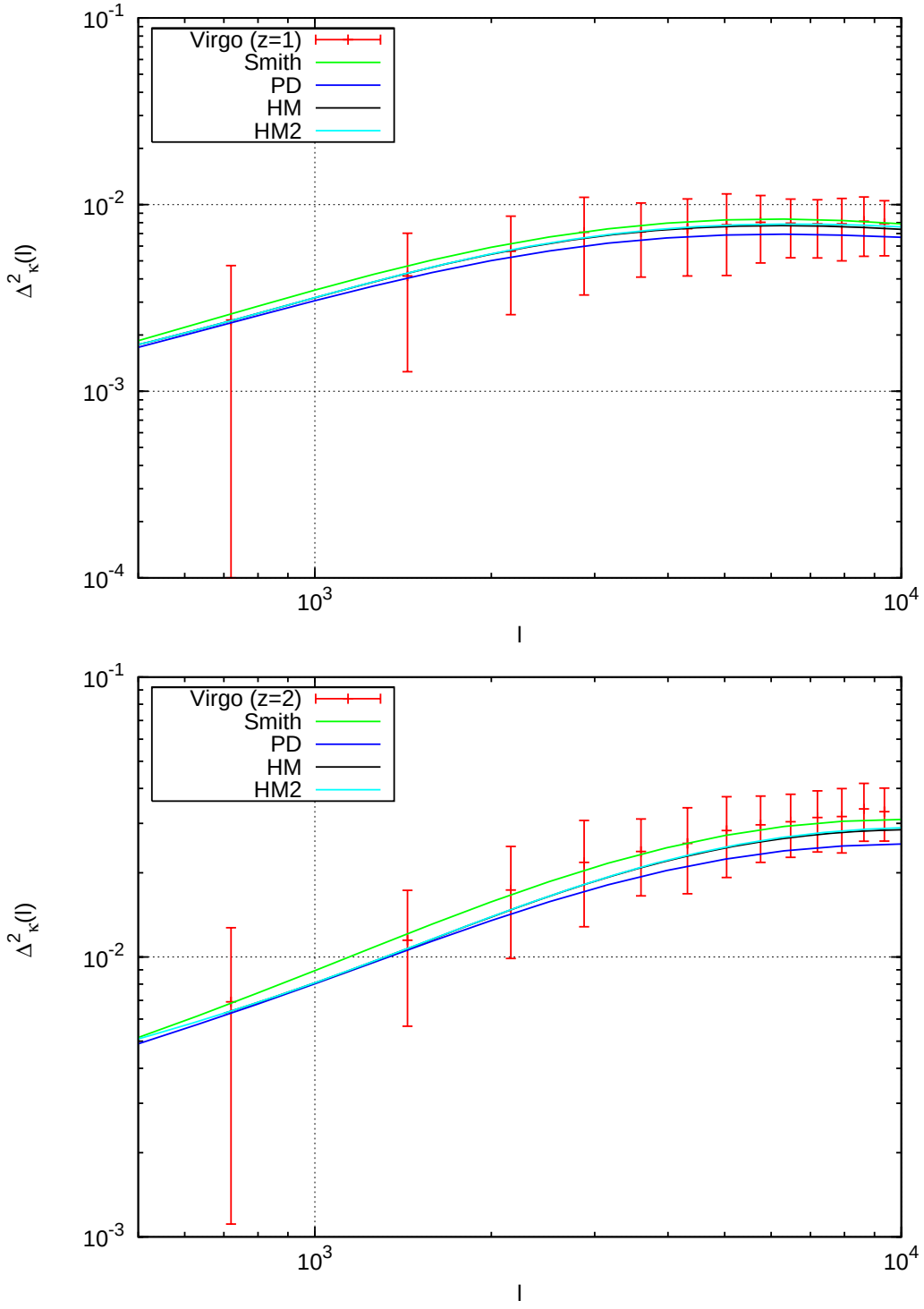

Figure 6.2: *Dimensionless convergence power spectrum against wave-number l. Both plots display the estimates obtained for the Virgo simulation. For the upper panel we considered the results for a single source redshift plane at $z_s = 1$, whereas the lower panel illustrates the results for $z_s = 2$. In both plots the red lines with errorbars correspond to the results as obtained from the numerical simulations. They are compared with the corresponding fitting formulae from Smith et al. (Smith, green line) and Peacock-Dodds (PD, blue line) and the halo model predictions for a deterministic halo concentration (HM, black line) and a concentration dispersion with $\sigma_{\ln c} = 0.3$ (HM2, turquoise line).*

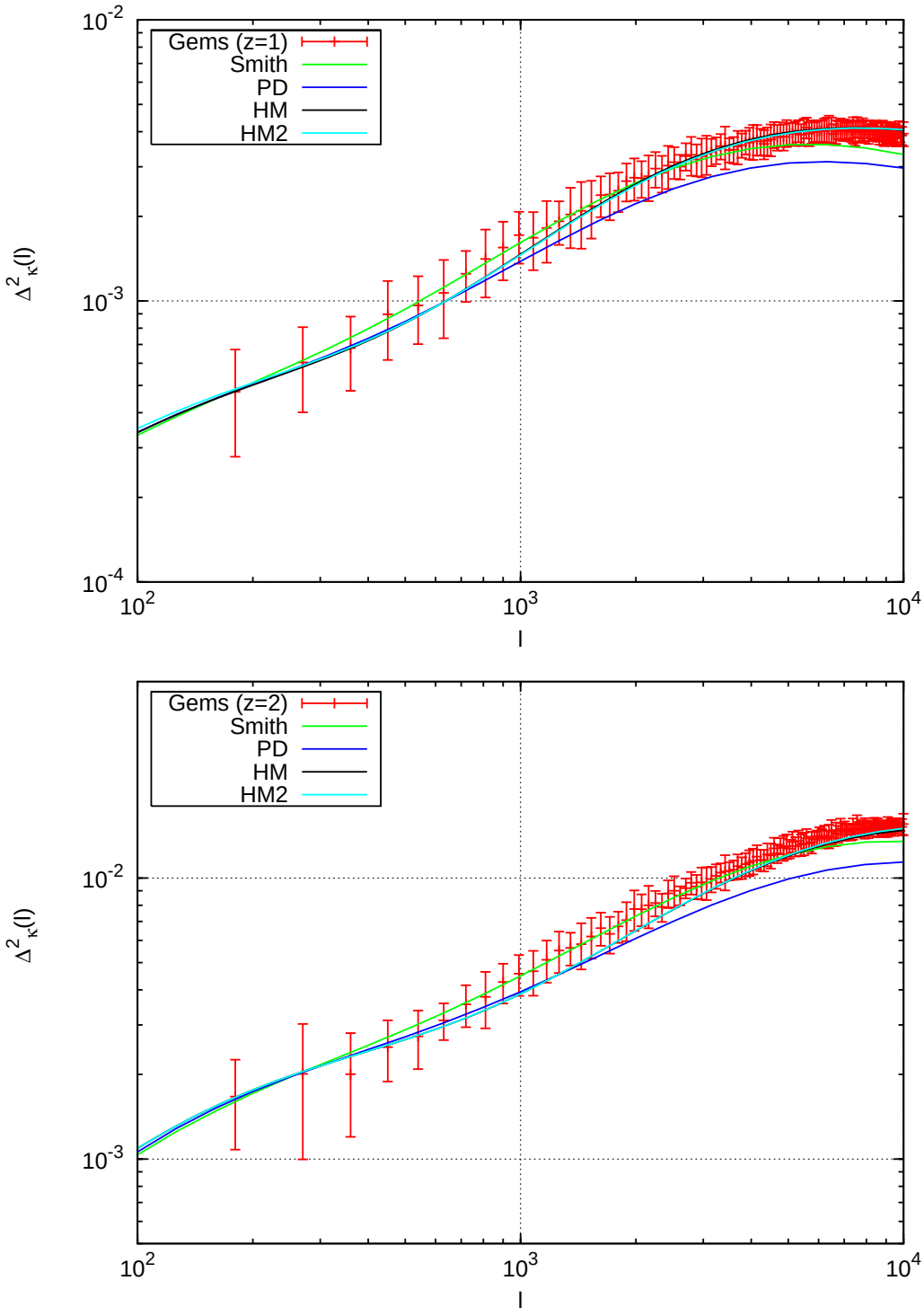

Figure 6.3: *Dimensionless convergence power spectrum against wave-number l. Both plots display the estimates obtained for the Gems simulation. For the upper panel we considered the results for a single source redshift plane at $z_s = 1$, whereas the lower panel illustrates the results for $z_s = 2$. In both plots the red lines with errorbars correspond to the results as obtained from the numerical simulations. They are compared with the corresponding fitting formulae from Smith et al. (Smith, green line) and Peacock-Dodds (PD, blue line) and the halo model predictions for a deterministic halo concentration (HM, black line) and a concentration dispersion with $\sigma_{\ln c} = 0.3$ (HM2, turquoise line).*

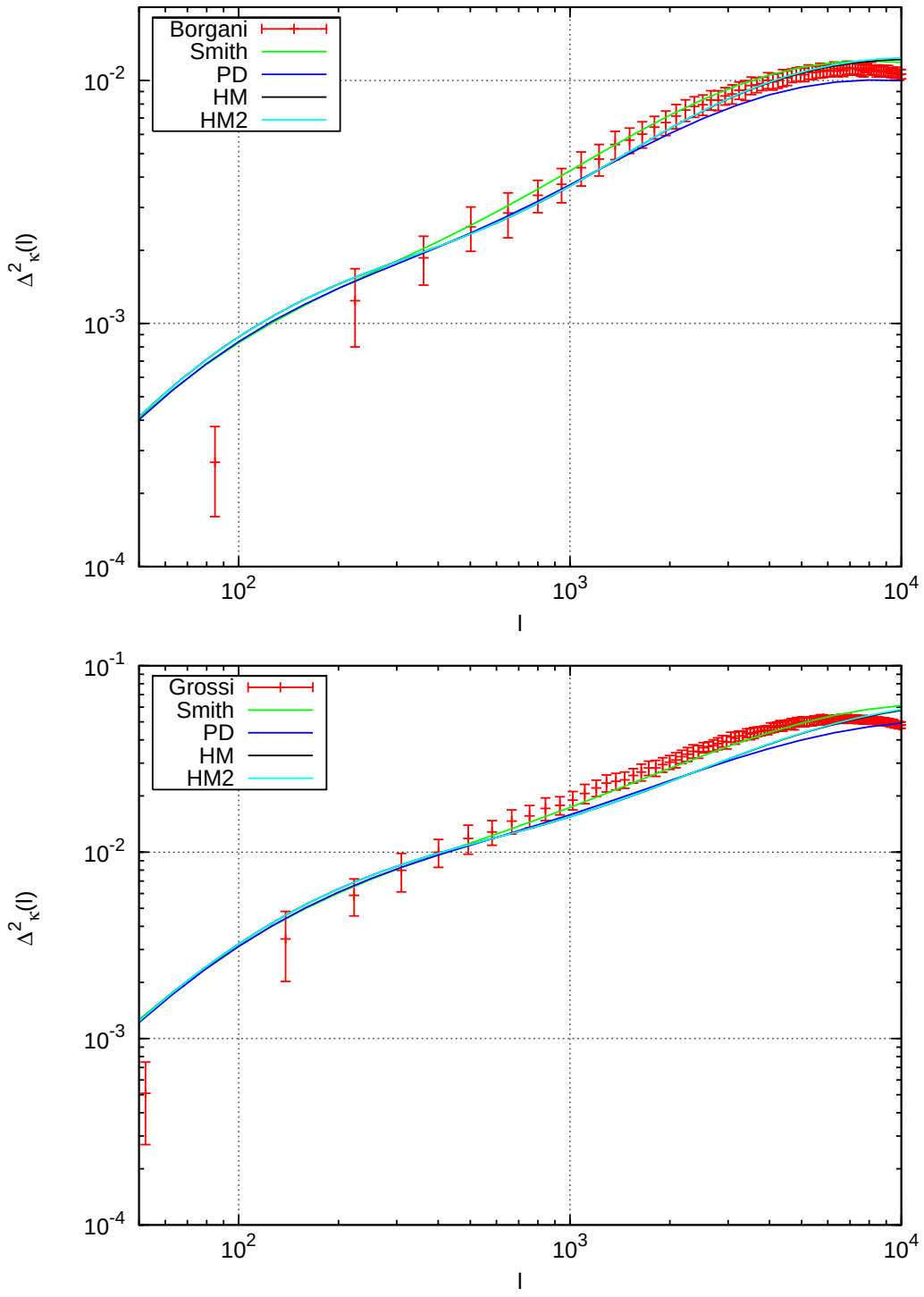

Figure 6.4: *Dimensionless convergence power spectrum against wave-number l. The upper panel displays the results as obtained from the analysis of the Borgani et al. simulation for a single source redshift plane $z_s = 1.5$, whereas the lower panel illustrates the convergence power spectrum estimate of the Grossi et al. simulation for $z_s = 4$. In both plots the red lines with errorbars correspond to the results as obtained from the numerical simulations. They are compared with the corresponding fitting formulae from Smith et al. (Smith, green line) and Peacock-Dodds (PD, blue line) and the halo model predictions for a deterministic halo concentration (HM, black line) and a concentration dispersion with $\sigma_{\ln c} = 0.3$ (HM2, turquoise line).*

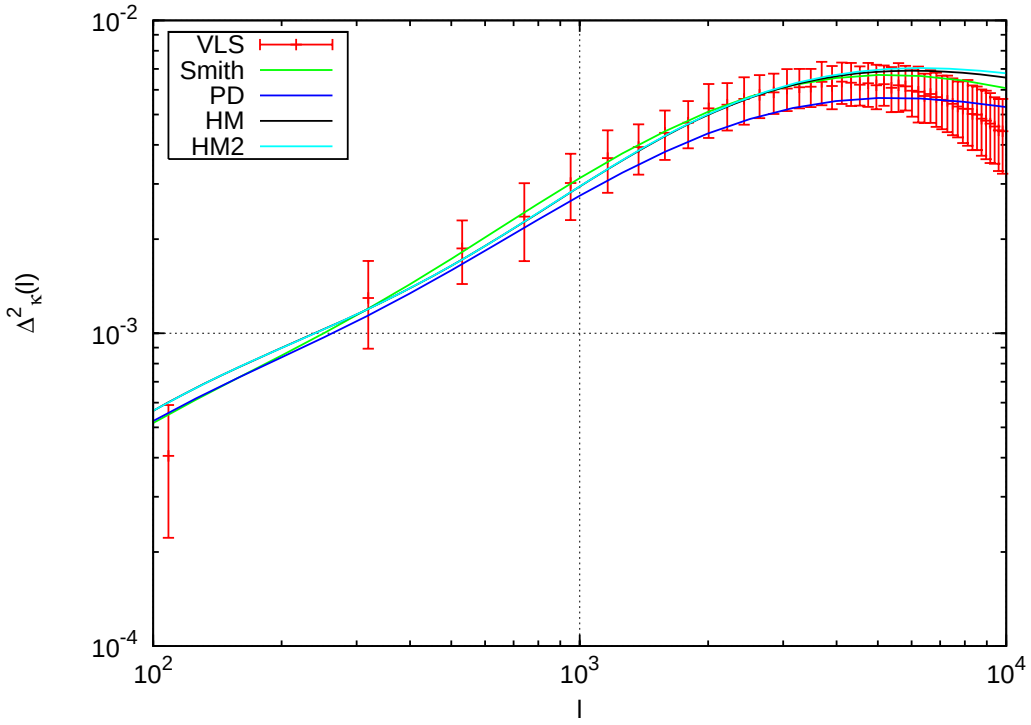

Figure 6.5: *Dimensionless convergence power spectrum against wave-number l. The plot illustrates the resulting estimate for the VLS simulation for sources situated at $z_s = 1$ (red lines with errorbars). It is compared with the corresponding fitting formulae from Smith et al. (Smith, green line) and Peacock-Dodds (PD, blue line) and the halo model predictions for a deterministic halo concentration (HM, black line) and a concentration dispersion with $\sigma_{\ln c} = 0.3$ (HM2, turquoise line).*

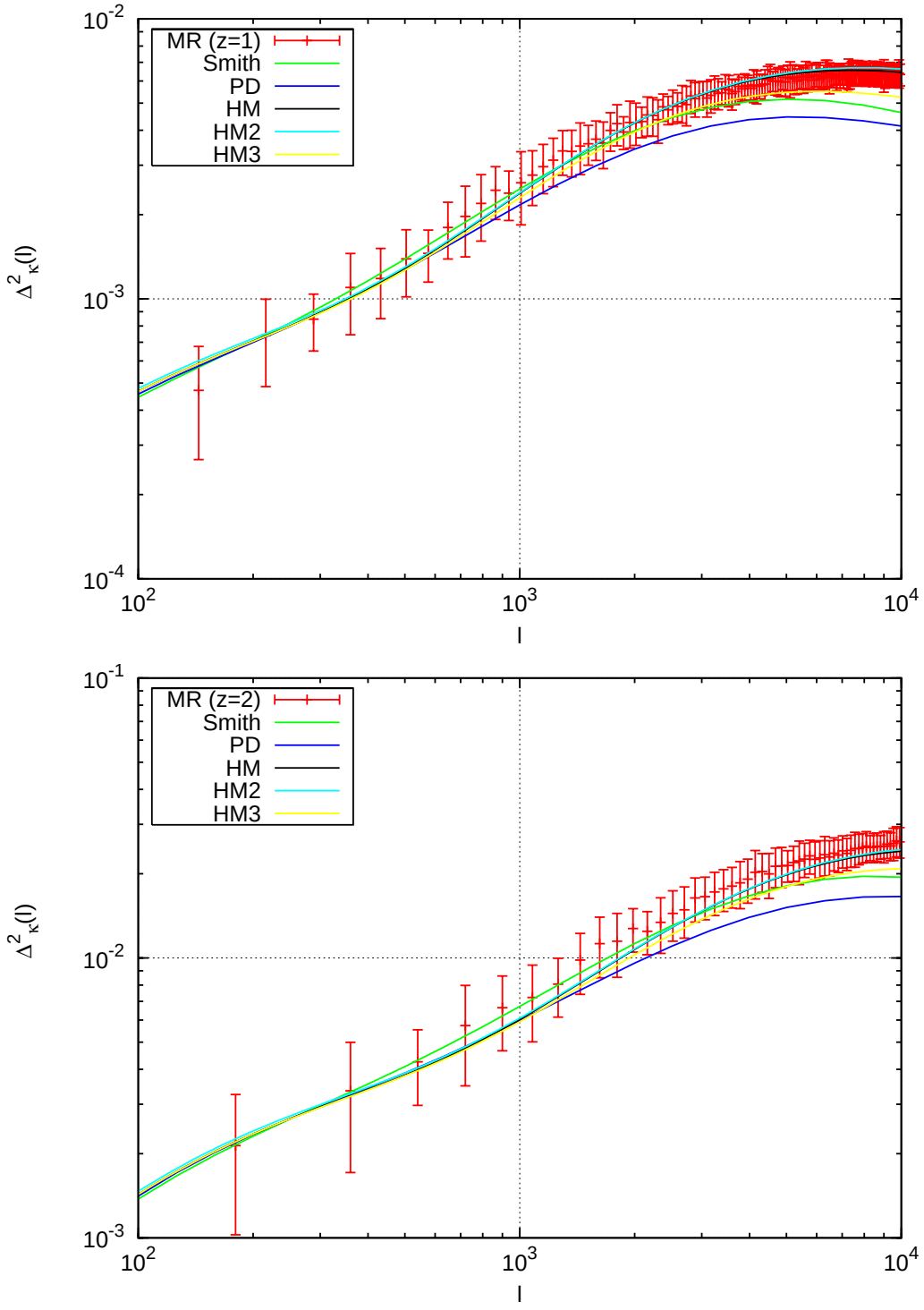

Figure 6.6: *Dimensionless convergence power spectrum against wave-number l. Both plots display the obtained estimates for the Millennium Run simulation. For the upper panel we considered the results for a single source redshift plane at $z_s = 1$, whereas the lower panel illustrates the results for $z_s = 2$. In both plots the red lines with errorbars correspond to the results as obtained from the numerical simulations. They are compared with the corresponding fitting formulae from Smith et al. (Smith, green line) and Peacock-Dodds (PD, blue line) and the three different halo model predictions. The black line (HM) shows the halo model prediction for a deterministic halo concentration with $c_* = 11$ and $\alpha = 0.2$, whereas the turquoise line (HM2) displays a probabilistic concentration-mass relation with $\sigma_{\ln c} = 0.3$. The halo model prediction for the Takada and Jain concentration relation is illustrated by the yellow line (HM3).*

6.3.1 Estimating the covariance from simulations

For the comparison with the halo model prediction, we need an appropriate estimator to find the power spectrum covariance for each simulation. In the last section we saw how to find an estimate for the power spectrum from a κ-map of the simulation. As each simulation provides N_{map} different κ-maps and thus realizations for the power spectrum, we can apply the unbiased sample covariance estimator as defined in Eq. (3.79). For our purpose it has the form:

$$\widehat{\mathrm{Cov}}(l_i, l_j) = \frac{1}{N-1}\left(\sum_{k=1}^{N} \hat{P}^{(k)}(l_i)\hat{P}^{(k)}(l_j) - \frac{1}{N}\sum_{k=1}^{N} \hat{P}^{(k)}(l_i)\sum_{k=1}^{N} \hat{P}^{(k)}(l_j)\right), \qquad (6.9)$$

where $\hat{P}^{(k)}(l_i)$ is the projected power spectrum estimate of the k-th effective convergence map at a wave-number l_i. The resulting quantity is then a measure of the coupling between the power at different wave-numbers. If the underlying density field were purely Gaussian, the off-diagonal entries would vanish as we saw in Sect. 3.4. The covariance matrix obtained in this way is then compared to the halo model prediction of the covariance estimate (see Eq. 5.15). For each simulation we derive a theoretical prediction of the covariance with the corresponding cosmological and simulation parameters as listed in Tabs. 6.1 and 6.2. The ingredients used for the halo model are as stated in Sect. 4.7.

To get a qualitative impression of how the results from N-body simulations match the theoretical halo model prediction we consider the covariances found for two different types of simulations in Fig. 6.7. Since the covariance estimator depends directly on the simulation parameters as the considered area and the binning scheme, we chose simulations that differ strongly in the two quantities. The upper panels illustrate covariances for the Borgani et al. simulation which uses a standard ΛCDM cosmology. They chose a linear binning scheme with a bin-width $\Delta l = 140$ from $l_0 = 84$ to $l_{70} = 9744$. In the lower panels, we consider covariances for the Virgo simulation with a similar cosmology. The binning scheme is also linear but coarser with a bin width $\Delta l = 720$ going from $l_0 = 720$ to $l_{20} = 4400$. Both simulations have a similar box size, but differ in their mass resolution and number of available κ-maps (see Tab. 6.1). Fig. 6.7 shows that the covariance matrix typically exhibits the largest values along the diagonal which is due to the contributions from a Gaussian and a non-Gaussian term (see Eq. 5.15). Towards smaller wave-numbers the covariance values become smaller. The left plots in Fig. 6.7 correspond to the theoretical predictions, whereas the right plots are the covariance estimates from the simulations. As one can see, the amplitude and overall shape of the covariances are in good agreement but the simulation covariances are noisier. This effect is much stronger for the Borgani simulation than for the Virgo simulation, which has two reasons: the different bin width and the number of available κ-maps to average over. Both arguments favor the

Virgo simulation as it uses a coarser binning scheme and offers a larger number of κ-maps. As a result, this simulation has much more estimates for each bin to average over which makes the resulting covariance apparently smoother.

6.3.2 Stability of the covariance

Relative deviation

Fig. 6.7 suggests that we should not only consider the absolute value but also the relative error of the covariance estimator. We study this effect for the Gems simulation for which the largest number of convergence maps is available ($N_{\rm map} = 220$). For this purpose, we split the convergence maps into ten subsets of equal number and find a covariance estimate for each subset. Four of the ten covariances found are displayed in Fig. 6.8 and show how strong the estimate can vary. To quantify the amount of variation we average over the ten covariance matrices found and calculate the standard variance $\mathrm{Var}(\mathcal{C})$ for each pair of bins (l_i, l_j). The upper plot of Fig. 6.9 illustrates the resulting relative deviation matrix for the ten Gems covariance estimates, which we define for this purpose as

$$\Delta\mathcal{C}_{ij} \equiv \sqrt{\mathrm{Var}(\mathcal{C}_{ij})}/\bar{\mathcal{C}}_{ij}\,. \tag{6.10}$$

Around the borders where one wave-number l is small the deviation is largest and can vary by more than 100%. The inner part of the covariance matrix still has a relative deviation of 65% and even the diagonal entries deviate by approximately 40%. This means that we should consider the results from simulations carefully, especially if only few tens of κ-maps are available to average over.

Bootstrap

Splitting the available effective convergence maps into smaller subsamples increases the variance of the estimator and is therefore likely to overestimate the relative error. A more realistic estimate of the deviation from the true covariance could be found if we had several covariance estimates that were averaged over the same number of κ-maps as our actual estimate. A method which allows us to estimate this is the so-called *bootstrap* method which was invented by Efron in 1979 (see e.g. [20]). It provides a way to construct bootstrap samples of the complete set of κ-maps from which one can obtain new covariance estimates. A bootstrap sample of size $N_{\rm size}$ is obtained by drawing repeatedly $N_{\rm size}$ κ-maps from the original sample with replacement. In this way one can produce the desired number of bootstrap samples $N_{\rm boot}$ – and thus covariance estimates – to average over. The relative error found should be smaller the larger the bootstrap sample size $N_{\rm size}$. For our purpose it is useful to find bootstrap samples of the same size as the original sample in order to determine the impact of the sample size on the relative error.

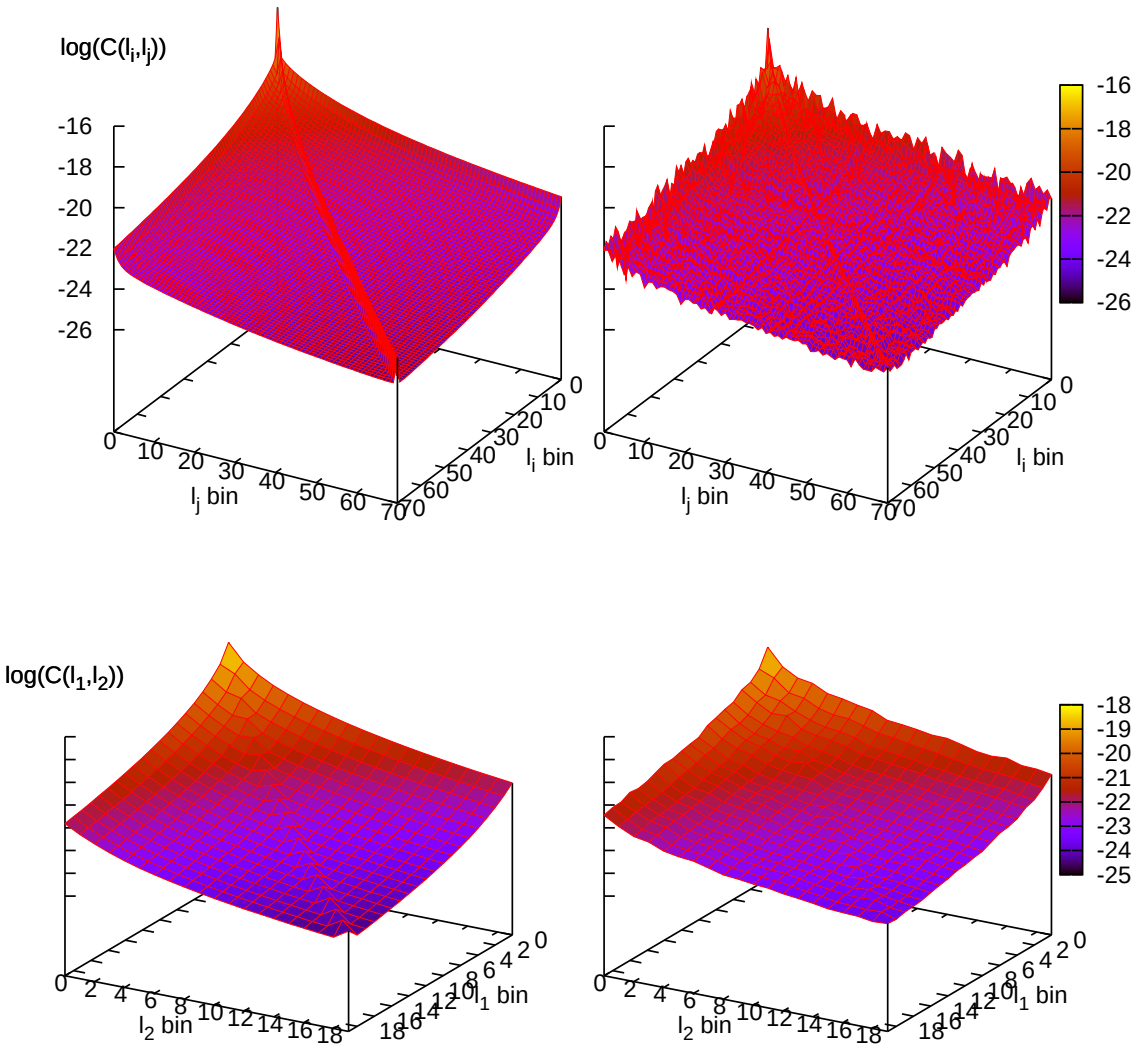

Figure 6.7: *Convergence power spectrum covariance* $\log \mathcal{C}_{ij}$ *against wave-numbers* (l_i, l_j) *for the Borgani (upper plots) and Virgo simulation (lower plots). The left plots correspond to the theoretical predictions as derived with the halo model, whereas the right plots display the simulation estimates for the power spectrum covariance. The finer binning of Borgani is more prone to noise than the coarser binning of the Virgo simulation.*

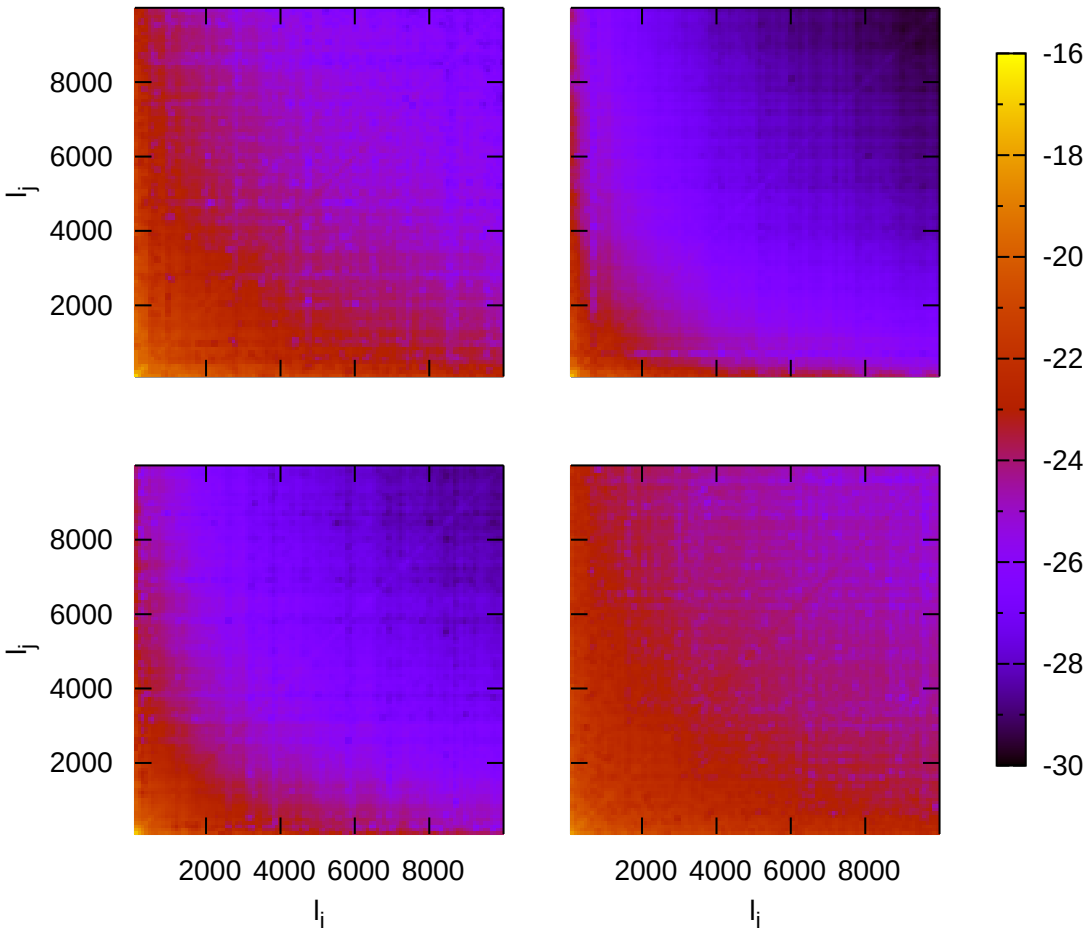

Figure 6.8: *Convergence power spectrum covariance* $\log(\mathcal{C}_{ij})$ *against wave-numbers* (l_i, l_j) *for the Gems simulation. After splitting the* $N_{\text{maps}} = 220$ *available* κ*-maps into ten subsets of equal size, we found ten estimates for the Gems covariance. The plots display four of them, illustrating that the estimates vary strongly.*

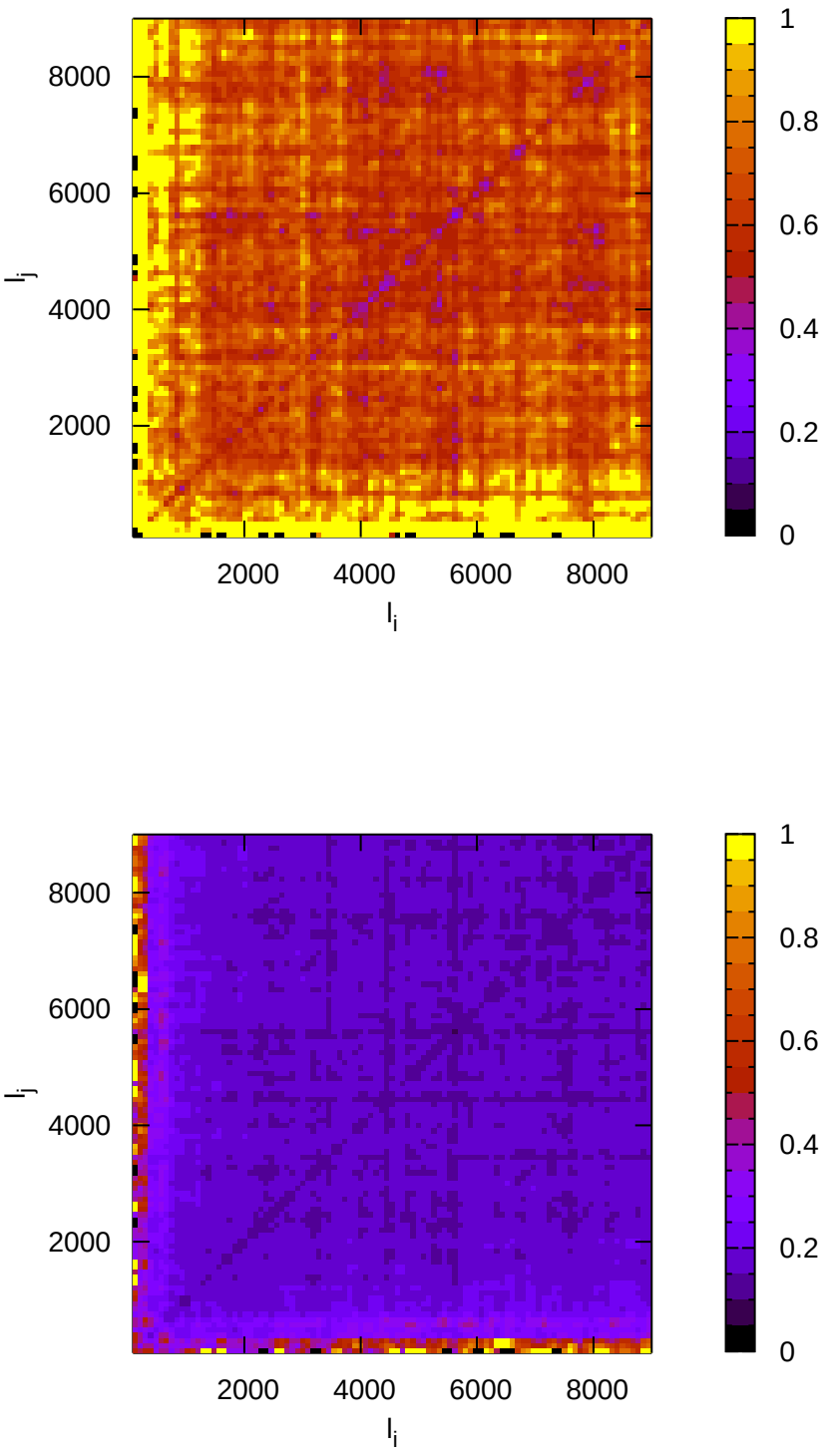

Figure 6.9: *Relative error of the Gems covariance $\Delta\mathcal{C}_{ij}$ as defined in Eq. (6.10) against wave-numbers (l_i, l_j). The upper plot displays the relative error as found from the ten different estimates obtained from $N_{\mathrm{map}} = 22$ convergence maps. The lower plot shows the relative error obtained from 50 bootstrap subsamples of size $N_{\mathrm{boot}} = 219$, which corresponds to the original sample size. For the bootstrap method the overall mean standard deviation is about 20%, whereas for the subsample method we find a mean standard deviation of about 40%.*

In order to determine a more realistic standard error for the covariance mean of the Gems simulation, we produce 50 bootstrap subsamples of size $N_{\rm boot} = 219$ which correspond to the size of the original sample. The resulting standard deviation is presented in the lower plot of Fig. 6.9 against the wave-numbers (l_i, l_j). In comparison to the former method where we found covariance estimates from subsets of the available κ-maps (upper plot) the error as determined with the bootstrap method decreased to approximately one-third of the previous values. Along the diagonal entries the mean standard deviation is around 10%, the overall mean standard deviation is 20%. Varying the number of bootstrap subsamples shows that this result is stable.

6.3.3 Comparing the covariances

Considering the covariances directly allows only a qualitative comparison between halo model and simulation prediction. For a more quantitative analysis, it is useful to consider their relative difference which we define as

$$\Delta\mathcal{C}_{ij} = \frac{\mathcal{C}^{ij}_{\rm halo} - \mathcal{C}^{ij}_{\rm sim}}{\mathcal{C}^{ij}_{\rm sim}}\,, \tag{6.11}$$

where $\mathcal{C}^{ij}_{\rm halo}$ denotes the halo model prediction of the covariance as given by Eq. (5.15) and $\mathcal{C}^{ij}_{\rm sim}$ is the corresponding simulation estimate we defined in Sect. 6.3.1. We assume that the resulting covariance from simulation is more reliable than the halo model prediction and put it in the denominator. Nevertheless, one should be aware that there is some scatter in the simulation covariances as we have seen in the previous section. This effect is most severe along the *border of the covariance*, i.e. in areas where one wave-number has a small value and the second wave-number can take arbitrary values.

The theoretical predictions are calculated as described in Sect. 5.3 and include the 1-halo, 2-halo and Gaussian contributions. Furthermore, we use our fiducial halo model as summarized in Sect. 4.7 and apply different cosmological parameters for each simulation as given in Tab. 6.2. The simulation estimate of the covariance is found as described earlier in this section.

The resulting plots in Figs. 6.10-6.14 clearly show that the theoretical halo model predictions underestimate the covariance estimates from simulations in general. The deviation is larger towards smaller scales. In the worst case the halo model covariance differs from the simulation covariance by a factor of 2. The best results are achieved on large scales and for the diagonal entries, where the halo model predictions differ around 10% from the simulation estimates. Including a scatter in the halo concentration parameter (see Sect. 5.4) increases the amount of correlation on small scales and thus decreases the relative difference between simulation and theoretical prediction. In the

following, we analyze the figures for the different simulations in detail. The scales considered for each simulation are listed in Tab. 6.3.

Table 6.3: *Binning scheme and wave-number ranges of the simulations.*

Simulation	$l_{\min}$	$l_{\max}$	Δl
Virgo	720	14400	720
Gems	90	9000	90
Borgani	84	9884	140
VLS	109	10441	211
Grossi	52	5244	88
Millennium Run	144	3672	72

Virgo

Fig. 6.10 compares the halo model prediction of the lensing power spectrum covariance with that from the Virgo simulation ($z_\mathrm{s} = 1$) by plotting the relative covariance difference $\Delta\mathcal{C}_{ij}$ against the wave-numbers (l_i, l_j). The binning scheme is linear with $\Delta l = 720$ going from $l_0 = 720$ to $l_{19} = 14400$. The upper plots illustrate the relative error for a deterministic relation between halo concentration and mass as described in Sect. 5.4. The left panel shows the full covariance, whereas the right panel shows a zoom of the lower right corner up to $l_5 = 4320$. On small scales, for wave-numbers larger than $l_{10} = 7920$, the halo model underestimates the results from the simulation up to $70\% - 80\%$, whereas on large scales, for wave-numbers smaller than $l_3 = 2160$, it overestimates the simulation results by about 50%. The best agreement is achieved on intermediate scales for l between 2200 and 4500 with a relative difference smaller or equal to 30%. The lower plots of Fig. 6.10 show that this can be improved significantly by considering a dispersion in the concentration parameter as defined in Eq. (5.49). Since we showed in Sect. 6.2.2 that the effect of a concentration dispersion is rather small for the projected power spectrum on the scales considered, we neglect this effect here and in the following in the Gaussian contribution of the covariance. For the halo model prediction of the covariance in the left panel we chose $\sigma_{\ln c} = 0.2$; for the right panel we used $\sigma_{\ln c} = 0.3$. In both cases the relative difference between the covariances is lowered significantly, in particular along the borders. For $\sigma_{\ln c} = 0.3$ the region with a deviation smaller than 30% extends up to $l = 10^4$. For wave-numbers smaller than $l \simeq 5000$ the error becomes 20% or less, which is within the uncertainty of the simulation. We obtain very similar results for sources at a redshift $z_\mathrm{s} = 2$, as shown in Fig. 6.11.

Gems

In Fig. 6.12 we illustrate the relative covariance difference between the halo model prediction and the Gems simulation ($z_\mathrm{s} = 1$). In contrast to the Virgo simulation the binning is much finer with a bin-width $\Delta l = 90$. The upper plots display the halo model prediction for the deterministic concentration-mass relation. The left panel shows the complete range of bins, whereas the right panel shows a zoom of the region with $l < 1800$. Again the simulation covariance is underestimated by the halo model up to 80% for wave-numbers larger than $l \simeq 6300$. The best agreement with $\Delta\mathcal{C}_{ij} < 30\%$ is found on large scales, where $l < 1000$, and along the diagonal, where $\Delta\mathcal{C}_{ij} < 50\%$. Including a scatter in the concentration of $\sigma_{\ln c} = 0.2$ (lower left panel) and $\sigma_{\ln c} = 0.3$ (lower right panel) decreases the relative difference significantly. In case of a concentration dispersion $\sigma_{\ln c} = 0.3$ the region up to $l \simeq 3600$ differs about 50% or less from the simulation results. In contrast to the Virgo simulation the halo model never overestimates the simulation.

Borgani

Fig. 6.13 displays the relative covariance difference between the halo model prediction and the Borgani simulation ($z_\mathrm{s} = 1$). The comparison covers a range from $l_0 = 84$ to $l_{70} = 8840$ using a bin-width $\Delta l = 140$. As one can see, this simulation is strongly affected by noise around the covariance border. The reason for this is twofold: all small wave-numbers have a large sampling variance and this simulation has only $N_\mathrm{map} = 60$ κ-maps to average over. The foregoing two simulations had both around $N_\mathrm{map} \approx 200$ and thus much fewer problems with noise. However, the overall impression of the quantitative analysis is very similar. The best agreement between theoretical prediction and simulation is along the diagonal entries and for wave-numbers smaller than $l \approx 3000$. In this regions the deviation is approximately 50%. Considering an additional scatter in the concentration parameter improves the result only slightly.

Millennium Run

In Fig. 6.14 which depicts the quantitative difference of halo model to Millennium Run simulation covariance the effect of noise is even more severe: The first eight bins are too noisy to make a quantitative analysis. This is due to the finer binning with $\Delta l = 72$ and even more importantly due to the very small number of κ-maps available which amounts to $N_\mathrm{map} = 20$. As a result the resemblance from theory to simulation is seldom better than 50%.

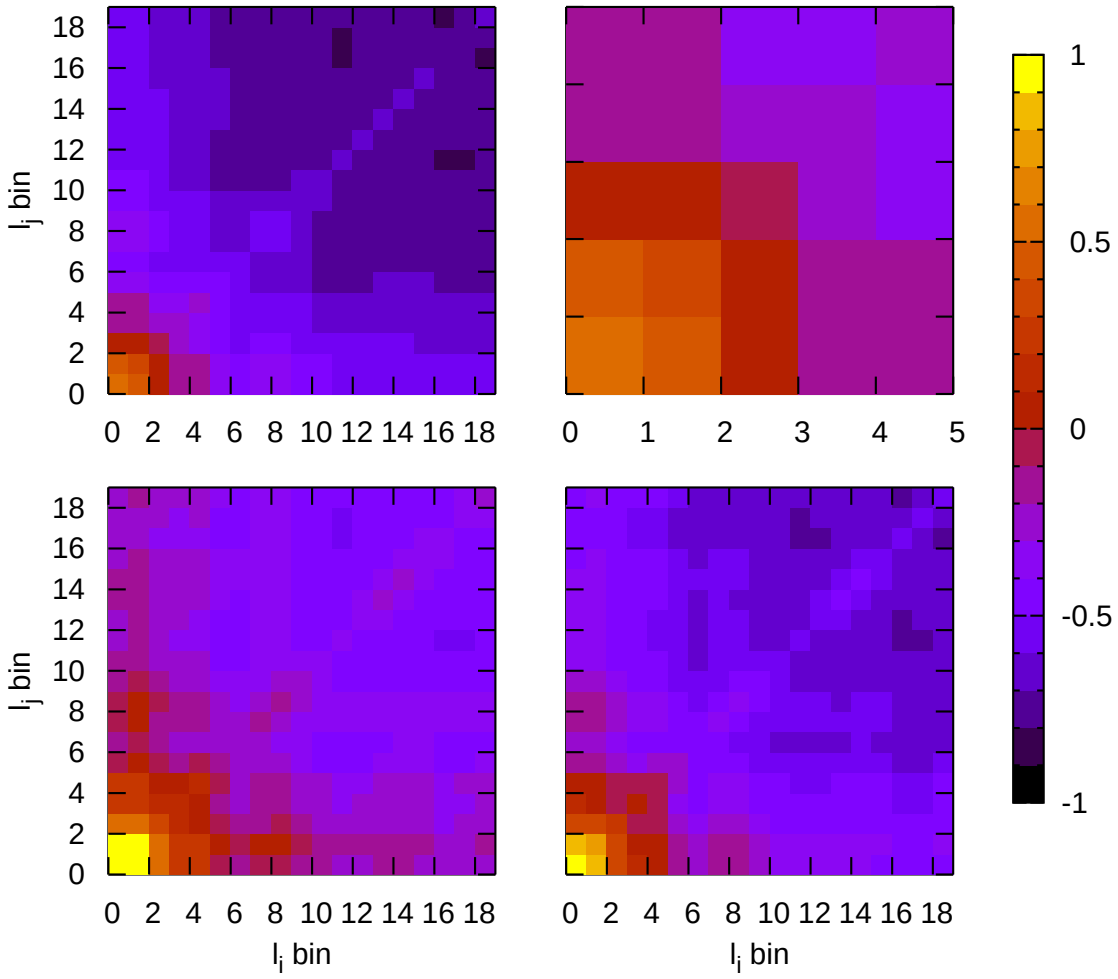

Figure 6.10: *Relative error $\Delta\mathcal{C}_{ij}$ of the theoretical halo model prediction for the projected power spectrum covariance in comparison with the results from the Virgo simulation ($z_\mathrm{s} = 1$) against wave-numbers (l_i, l_j). The binning scheme is linear with $\Delta l = 720$ going from $l_0 = 720$ to $l_{19} = 14400$. The upper plots illustrate the relative error as found when considering the 1-halo, 2-halo and Gaussian term with $\sigma_{\ln c} = 0$ for the halo model. Left and right plot differ only in the number of considered bins. In the lower left panel, we consider $\sigma_{\ln c} = 0.2$ and the lower right panel displays a variance $\sigma_{\ln c} = 0.3$.*

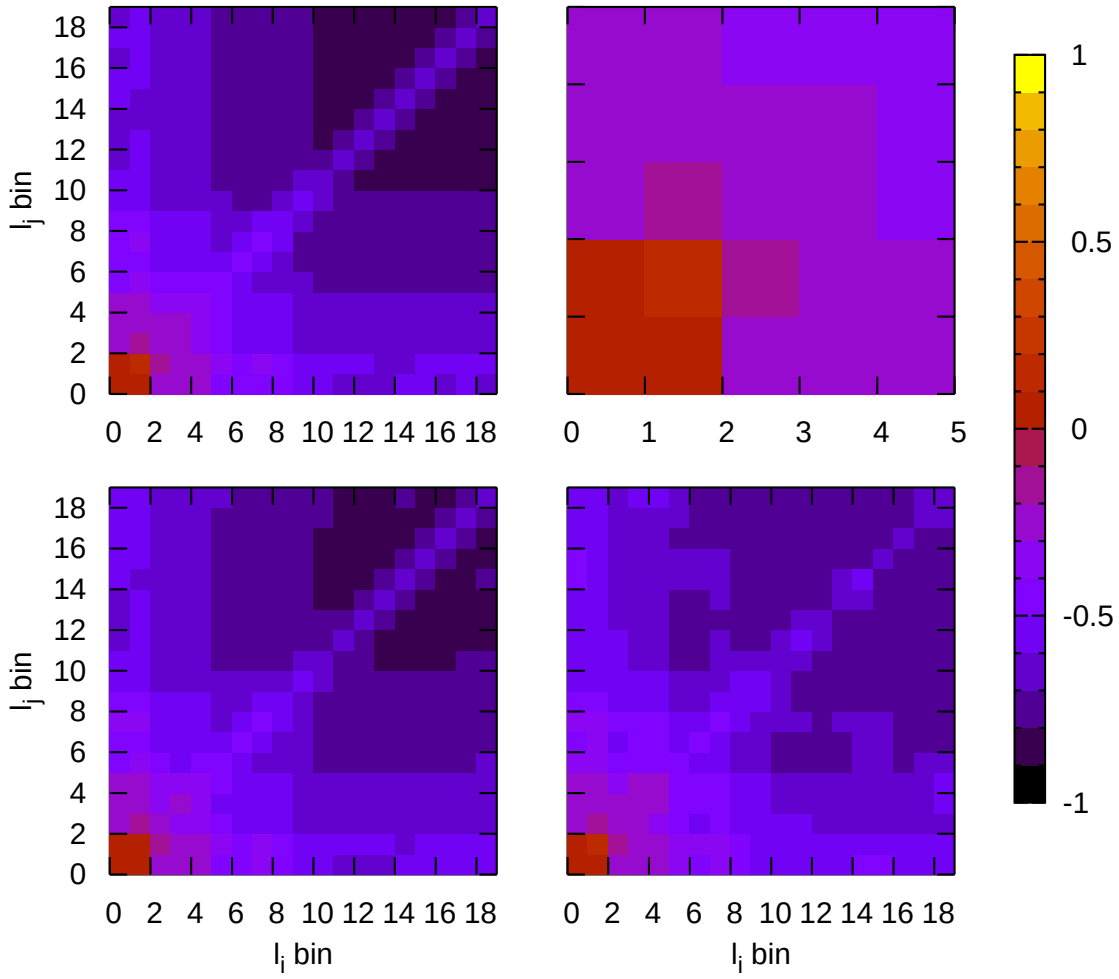

Figure 6.11: *Relative error $\Delta\mathcal{C}_{ij}$ of the theoretical halo model prediction for the projected power spectrum covariance in comparison with the results from the Virgo simulation ($z_\mathrm{s} = 2$) against wave-numbers (l_i, l_j). The binning scheme is linear with $\Delta l = 720$ going from $l_0 = 720$ to $l_{19} = 14400$. The upper plots illustrate the relative error as found when considering the 1-halo, 2-halo and Gaussian term with $\sigma_{\ln c} = 0$ for the halo model. The right plot is a zoom into the left plot. In the lower left panel, we consider $\sigma_{\ln c} = 0.2$ and the lower right panel displays a variance $\sigma_{\ln c} = 0.3$.*

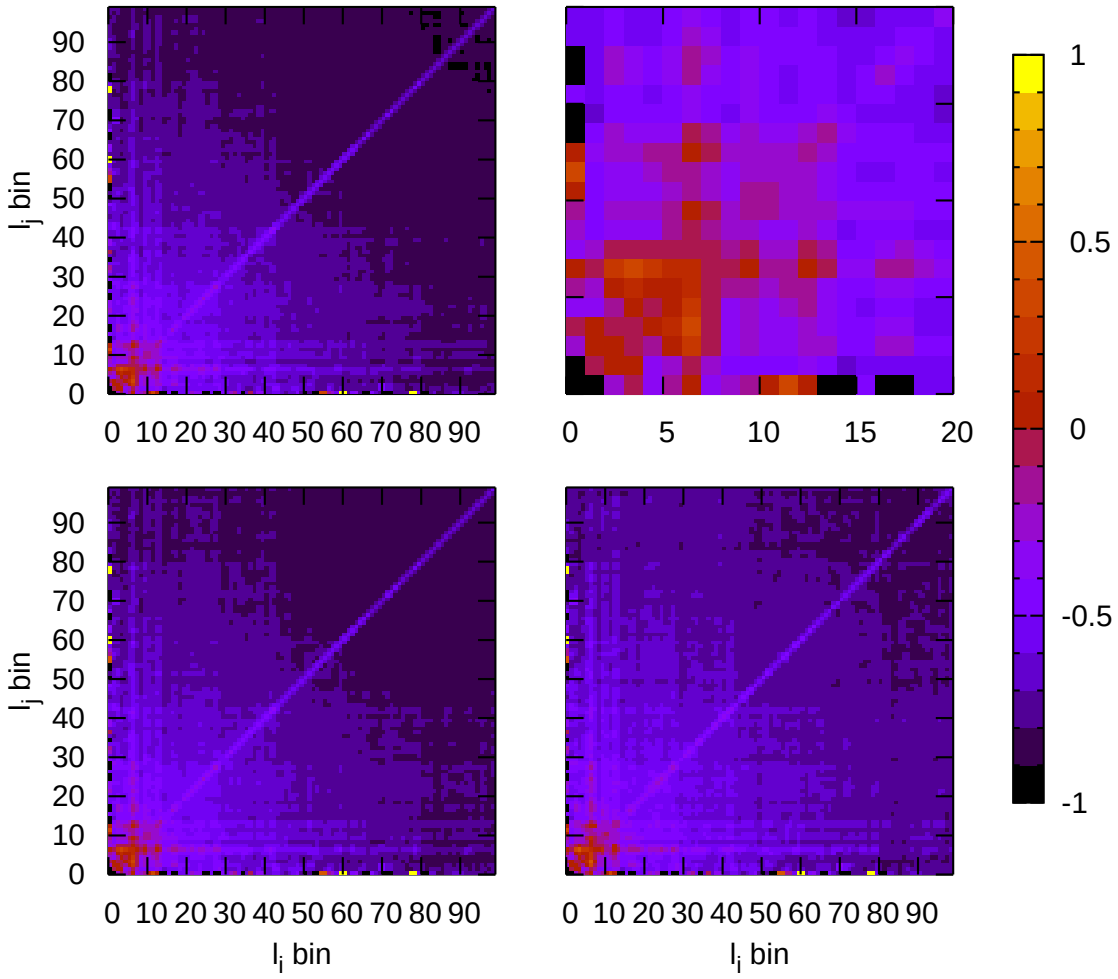

Figure 6.12: *Relative error $\Delta\mathcal{C}_{ij}$ of the theoretical halo model prediction for the projected power spectrum covariance in comparison with the results from the Gems simulation ($z_\mathrm{s} = 1$) against wave-numbers (l_i, l_j). The binning scheme is linear with $\Delta l = 90$ going from $l_0 = 90$ to $l_{99} = 9000$. The upper plots illustrate the relative error as found when considering the 1-halo, 2-halo and Gaussian term with $\sigma_{\ln c} = 0$ for the halo model. The right plot is a zoom into the left plot. In the lower left panel, we consider $\sigma_{\ln c} = 0.2$ and the lower right panel displays a variance $\sigma_{\ln c} = 0.4$.*

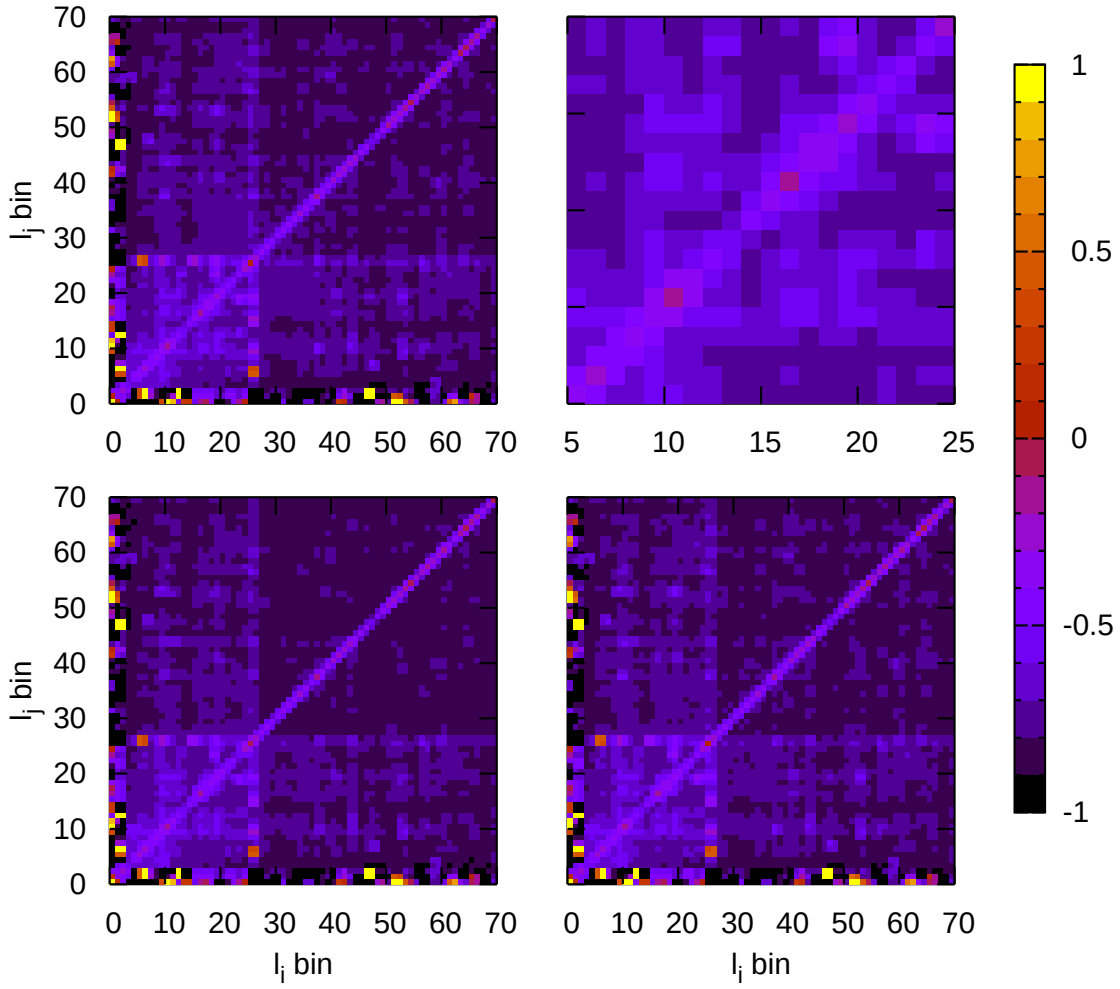

Figure 6.13: *Relative error $\Delta\mathcal{C}_{ij}$ of the theoretical halo model prediction for the projected power spectrum covariance in comparison with the results from the Borgani simulation ($z_\mathrm{s} = 1$) against wave-numbers (l_i, l_j). The binning scheme is linear with $\Delta l = 140$ going from $l_0 = 84$ to $l_{70} = 9884$. The upper plots illustrate the relative error as found when considering the 1-halo, 2-halo and Gaussian term with $\sigma_{\ln c} = 0$ for the halo model. The right plot is a zoom into the left plot. In the lower left panel, we consider $\sigma_{\ln c} = 0.2$ and the lower right panel displays a variance $\sigma_{\ln c} = 0.3$. The simulation shows a lot of noise close to the covariance borders.*

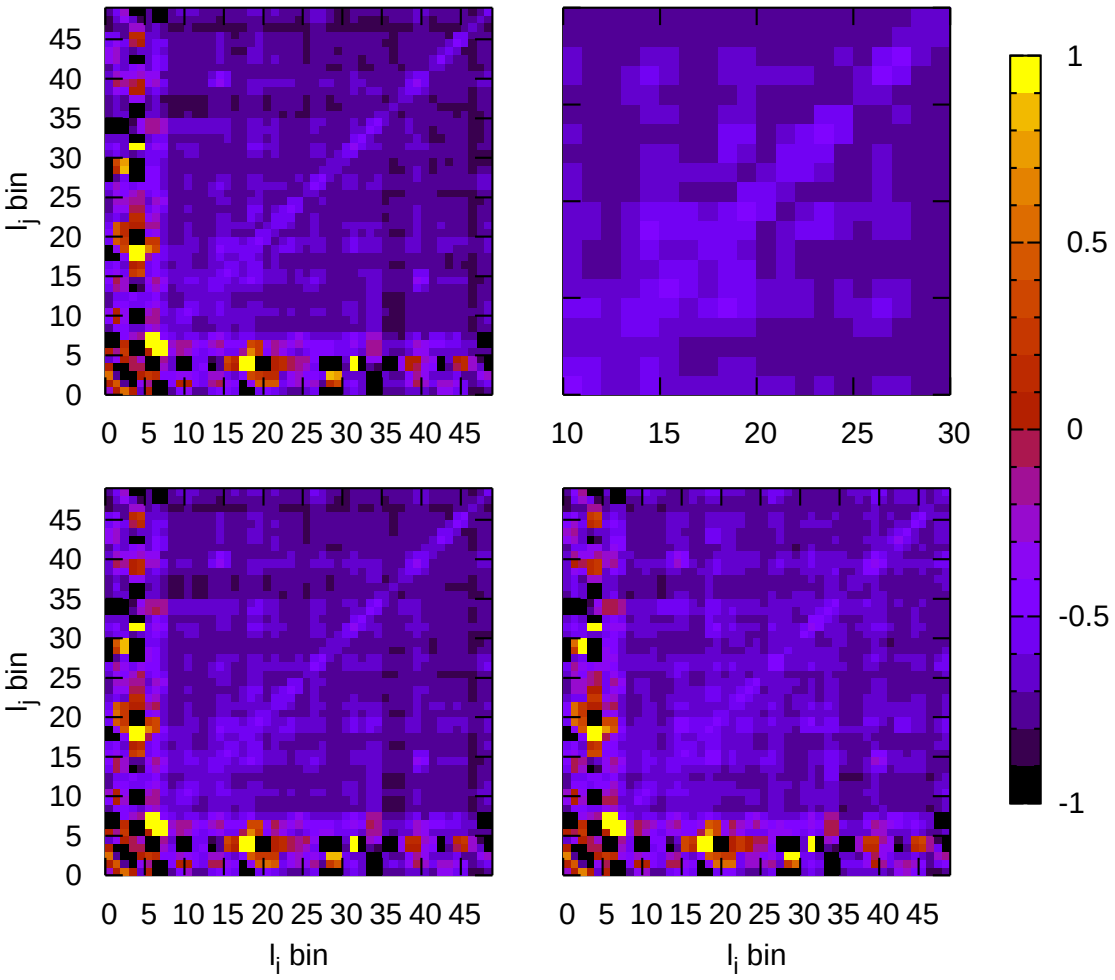

Figure 6.14: *Relative error $\Delta\mathcal{C}_{ij}$ of the theoretical halo model prediction for the projected power spectrum covariance in comparison with the results from the Millennium Run simulation ($z_\mathrm{s} = 1$) against wave-numbers (l_i, l_j). The binning scheme is linear with $\Delta l = 72$ going from $l_0 = 144$ to $l_{49} = 3672$. The upper plots illustrate the relative error as found when considering the 1-halo, 2-halo and Gaussian term with $\sigma_{\ln c} = 0$ for the halo model. The right plot is a zoom into the left plot. In the lower left panel, we consider $\sigma_{\ln c} = 0.2$ and the lower right panel displays a variance $\sigma_{\ln c} = 0.4$. The simulation shows a lot of noise close to the covariance borders.*

6.4 Non-Gaussian to Gaussian ratio

In order to understand how the non-Gaussian contribution scales with respect to the Gaussian part of the convergence covariance, we focus in this section on the ratio between the two

$$R_l \equiv \frac{\mathcal{C}^{\mathrm{NG}}(l,l)}{\mathcal{C}^{\mathrm{G}}(l,l)} = \frac{\mathcal{C}(l,l)}{\mathcal{C}^{G}(l,l)} - 1\,, \tag{6.12}$$

where $\mathcal{C}^{\mathrm{NG}}$ denotes the non-Gaussian and $\mathcal{C}^{\mathrm{G}}$ the Gaussian contribution of the covariance as defined in Sect. 5.3.3. Note that only one wave-number l occurs, since the Gaussian part contributes only to the diagonal of the covariance. In the halo model approach we can calculate this ratio as

$$R_l = \frac{\bar{T}_\kappa(l,l)}{4\pi P_\kappa^2(l)}\,\Delta l \cdot l\,, \tag{6.13}$$

where $\bar{T}_\kappa(l,l)$ is the bin-averaged convergence trispectrum, $P_\kappa(l)$ the convergence power spectrum and Δl the bin-width of the simulation (see Sect. 5.6.2). In the following, we will compare this with estimates obtained from simulations.

6.4.1 Estimating the ratio from simulations

The simplest way to find an estimator for Eq. (6.12) is to formulate separate estimators for each contribution, i.e. for the complete convergence covariance $\mathcal{C}$ and the Gaussian contribution $\mathcal{C}^{\mathrm{G}}$. The first part is straightforward, as we can make use of the covariance estimator from Sect. 6.3.1:

$$\widehat{\mathcal{C}}(l,l) = \widehat{\mathrm{Cov}}[\hat{P}_\kappa(l), \hat{P}_\kappa(l)] = \widehat{\mathrm{Var}}[\hat{P}_\kappa(l)] \tag{6.14}$$

For the Gaussian covariance contribution we construct the estimator in the following way:

$$\widehat{\mathcal{C}^{\mathrm{G}}}(l,l) = B\langle \hat{P}_\kappa(l)^2\rangle = B(\widehat{\mathrm{Var}}[\hat{P}_\kappa(l)] + \langle \hat{P}_\kappa(l)\rangle^2)\,, \tag{6.15}$$

where $B = \frac{2}{A}\,\frac{(2\pi)^2}{A_{\mathrm{r}}(l)}$ as defined in Eq. (5.57). This choice has the advantage that we can reuse the results for the convergence covariance and power spectrum estimators of the previous sections. Putting together Eqs. (6.14) and (6.15), the estimator for the ratio is

$$\hat{R}_l = \frac{\widehat{\mathcal{C}}(l,l)}{\widehat{\mathcal{C}^{\mathrm{G}}}(l,l)} - 1 = \frac{\widehat{\mathrm{Var}}[\hat{P}_\kappa(l)]}{B(\widehat{\mathrm{Var}}[\hat{P}_\kappa(l)]) + \langle \hat{P}_\kappa(l)\rangle^2)} - 1\,. \tag{6.16}$$

As the power spectrum variance is much smaller than the average power spectrum squared, we approximate the denominator and obtain

$$\hat{R}_l \simeq \frac{\widehat{\mathrm{Var}}[\hat{P}_\kappa(l)]}{B\langle \hat{P}_\kappa(l)\rangle^2} - 1\,. \tag{6.17}$$

6.4.2 Results

Figs. 6.15 and 6.16 show the estimates obtained from different N-body simulations and compare them to the halo model prediction. From the plots one can see that both estimates are of similar magnitude, but differ in the behavior on small scales. For the Virgo and the Gems simulations the constructed estimator from Eq. (6.17) is too noisy for a reliable comparison. One factor for this might be that the estimator is constructed from two other estimators which suffer from noise effects as well. The VLS and Millennium Run simulations show an increasing amount of non-Gaussianity on small scales contrary to the decreasing slope predicted by the halo model. This behavior is most likely the result of discreteness effects as the amount of increase is very steep. This is presumedly due to three effects: shot noise and force softening (see Sect. 6.1.3) become important for wave-numbers larger than $l \simeq 10^4$. Additionally, *aliasing* might occur on small scales causing a decrease of the power spectrum estimate [39]. As the power spectrum is inversely proportional to the ratio this can explain the rise of the ratio on small scales. The overall impression is that the halo model prediction and the simulation estimates do not coincide well. Currently, one cannot say whether this is because the halo model misinterprets the non-Gaussian contribution or because the present simulations are not reliable enough on small scales.

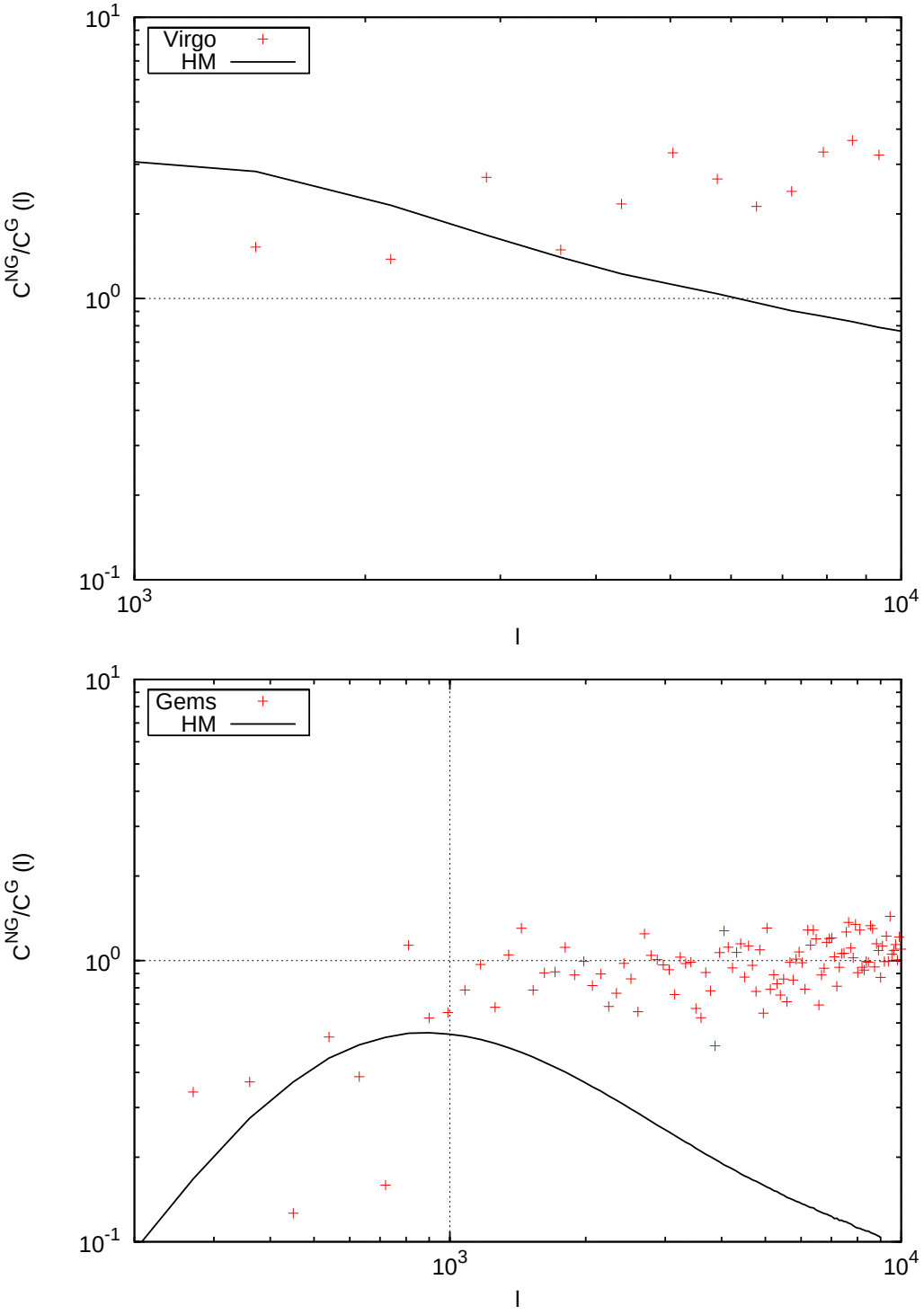

Figure 6.15: *Ratio of non-Gaussian to Gaussian contribution of the convergence power spectrum covariance R_l against wave-number l. The halo model (HM, black line) predicts a reduction of non-Gaussianity on small scales. The Virgo and Gems simulation estimates (red points) are very noisy and do not show a clear behavior.*

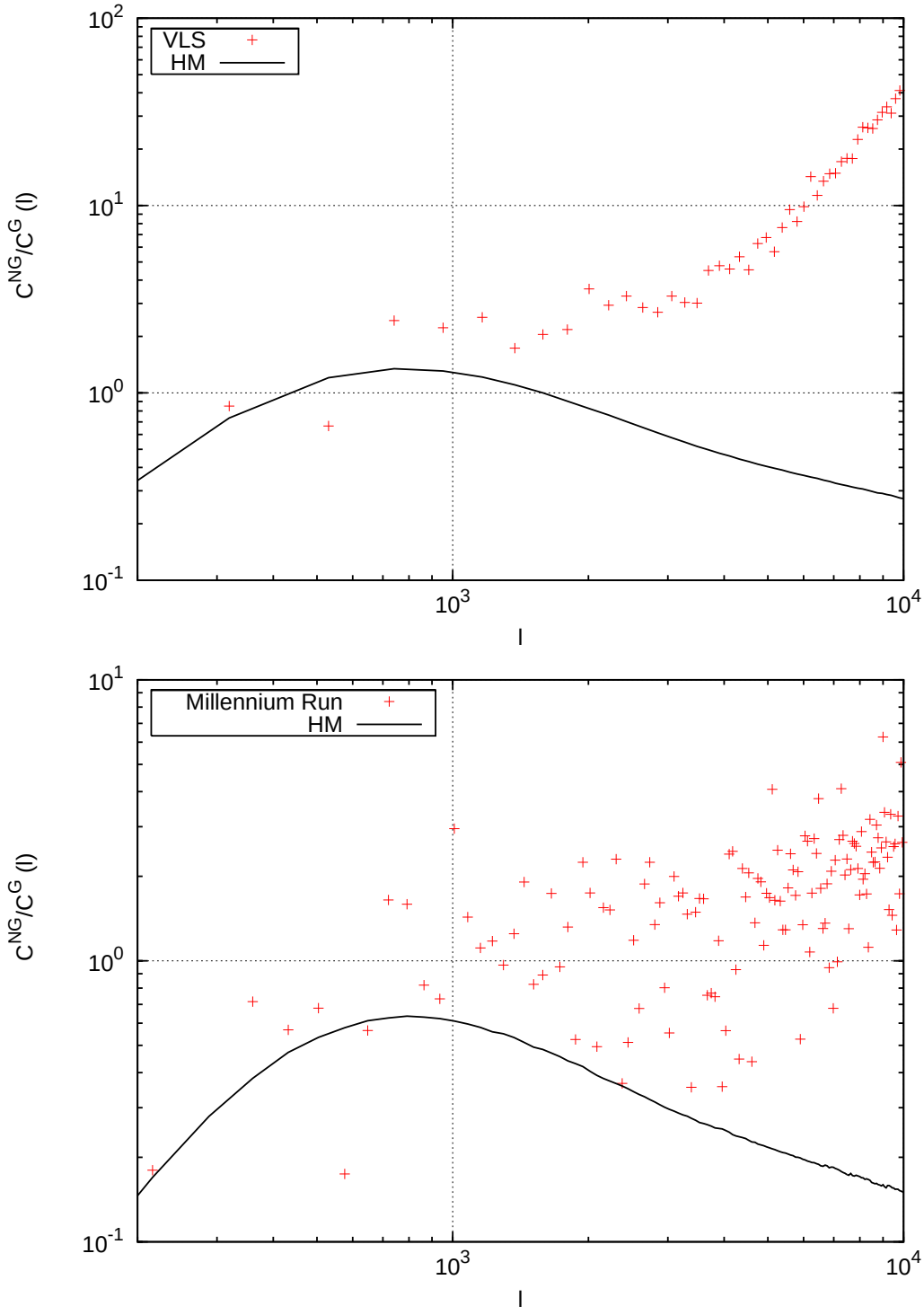

Figure 6.16: *Ratio of non-Gaussian to Gaussian contribution of the convergence power spectrum covariance R_l against wave-number l. The halo model (HM, black line) predicts a reduction of non-Gaussianity on small scales. The VLS and Millennium Run simulation estimates (red points) predict the opposite behavior.*

Summary and conclusions

The motivation behind this thesis was to provide an analytical treatment of higher-order correlation functions in the cosmic matter density field and to compare the results obtained with numerical N-body simulations. As the perturbative description holds only for densities of the order $|\delta| \simeq 1$, the tool of choice was a semi-analytic halo model which combines results from both perturbation theory and numerical simulations. In this work, we focused on the fourth-order correlation function and its Fourier counterpart, the trispectrum, since it allows us to study the non-Gaussianities of the matter field and to calculate the full non-Gaussian covariance of the power spectrum. This provides a way to estimate the error and mode coupling in the matter power spectrum to higher accuracy than has been previously.

To calculate the trispectrum and, subsequently, the full non-Gaussian covariance of the convergence power spectrum in the halo model approach, we had to combine results from different areas in mathematics, physics and cosmology. We have summarized the most important ones in the first chapters of this thesis. This includes a detailed overview of the standard model of cosmology, perturbation theory and the properties of cosmological random fields. Additionally, Chapter 4 gives a comprehensive overview of the halo model description of dark matter and its ingredients such as the halo mass abundance, halo profile and clustering of halos.

With the halo model at hand, one has a recipe to calculate matter correlation functions of arbitrary order. We used the halo model to find the expectation value both for the three-dimensional and the convergence power spectrum covariance and confirmed the analytical results of Cooray and Hu [16] and Scoccimarro et al. [72]. In order to find a fast way to calculate the full non-Gaussian covariance, we studied the accuracy of different approximations to the complete trispectrum in the halo model approach. As a result, we found that the combination of 1-halo and 2-halo contributions of the trispectrum yields an error smaller than 10% on intermediate and large scales. More precisely, in the three-dimensional case this approximation is accurate for $k \gtrsim 0.3\,h\,\mathrm{Mpc}^{-1}$ and in the projected case for $l \gtrsim 300$ and $z_\mathrm{s} = 1$ for our fiducial ΛCDM model. Furthermore, we extended this result to the non-Gaussian contribution of the covariance. The aforementioned approximation yields for the same wave-numbers a comparable error to the total trispectrum. Thus a combination of 1-halo and 2-halo terms of the bin-averaged trispectrum together with the Gaussian contribution of the covariance provides an accurate and efficient approximation to the full non-Gaussian covariance of the matter power spectrum.

Since recent results from numerical N-body simulations suggest that the concentration-mass relation is probabilistic [38], i.e. there is scatter in the concentration parameter for a fixed mass m, we investigated the impact of this on the 1- and 2-halo term contributions to the power spectrum and trispectrum. In the three-dimensional case, we found a significant deviation from a deterministic concentration relation on scales smaller than $k \simeq 5\,h\,\mathrm{Mpc}^{-1}$, which is less pronounced for the power spectrum. Additionally, we were able to show that this effect is also present in the projected versions of the spectra on scales smaller than $l \simeq 2000$. In contrast to the three-dimensional case, the increase in the small scale tails of the spectra is diminished, but alters the covariance on small scales up to 12% for reasonable concentration dispersions.

In order to understand how different wave-number modes in the power spectrum covariance couple, we analyzed the angular dependence of the normalized trispectrum for each halo term. We found the strongest mode coupling in the 4-halo term for colinear configurations of the trispectrum wave-vectors. Towards orthogonal configurations the angular dependence vanishes and there is almost no coupling visible. We discovered a similar but less pronounced behavior for the 3-halo term. The 2-halo and 1-halo terms show no visible angular dependence of the modes. Furthermore, we studied the mode coupling for trispectrum configurations with different wave-vector length. This revealed a dependence on both the length of the wave-vectors and their ratio. The minimum wave-length, in particular, affects the amplitude of the trispectrum.

Studying the non-Gaussian-to-Gaussian ratio of the covariance, we found an increase of non-Gaussianity towards small scales for the three-dimensional case. In the projected case, the non-Gaussian contribution reaches a maximum amplitude around $l \simeq 800$ and decreases on small scales.

With the results from the halo model consideration, we developed a fitting formula for the non-Gaussian contribution of the convergence power spectrum covariance. The fitting formula is valid in a range $1000 < l < 5000$ and provides on average a 10% accuracy to the corresponding halo model prediction. For the diagonal of the non-Gaussian covariance we achieve in this way a 15% accuracy.

An additional aspect of this thesis was the comparison of the halo model to results from simulations. For this we calculated the power spectrum and covariance from six different numerical N-body simulations that spanned a wide range of possible designs. Our comparison with the halo model revealed a good correspondence with the power spectra. However, the convergence power spectrum covariance with a deterministic concentration parameter consistently underestimates the covariance estimate from simulations on small scales. This effect is mitigated if we include in the halo model prediction a probabilistic concentration relation with a dispersion $\sigma_{\ln c} = 0.2 - 0.3$ in the logarithmic halo concentration. Nevertheless, the halo model prediction of the covariance underestimates still the simulation estimate on small scales. This discrepancy hints that the halo model description does not reflect well enough the underlying physics of higher-order correlations on small scales. One also has to be aware that the limited size and resolution of simulations can lead to high noise and unreliable predic-

tions as well. This was particularly visible in our analysis of the ratio of non-Gaussian-to-Gaussian contributions to the convergence power spectrum covariance. While the halo model predicted a decrease towards small scales, the simulations showed an increase. A final conclusion whether the halo model has to be corrected for this will require future simulations of improved quality.

Appendix A

Halo model trispectrum

In Sect. 5 we showed that for the calculation of the power spectrum covariance in the halo model description, we only have to consider parallelogram configurations of the trispectrum wave-vectors. The restriction to these configurations allows us to simplify the expressions for the terms of the trispectrum. In the following, we perform this calculation in detail after summarizing the most important properties of the second order coupling functions.

A.1 Second-order coupling functions

From the recursion relations (2.26) and (2.27) for the n-th order density contrast δ_n and divergence velocity field θ_n that solve in a perturbative approach the collisionless Boltzmann equation for an ideal dark matter fluid, we find for the second-order coupling functions the following expressions:

$$F_2(\boldsymbol{q}_1, \boldsymbol{q}_2) = \frac{5}{7}\,\alpha(\boldsymbol{k}_1, \boldsymbol{k}_2) + \frac{2}{7}\,\beta(\boldsymbol{k}_1, \boldsymbol{k}_2)\,, \tag{A.1}$$

$$G_2(\boldsymbol{q}_1, \boldsymbol{q}_2) = \frac{5}{7}\,\alpha(\boldsymbol{k}_1, \boldsymbol{k}_2) + \frac{2}{7}\,\beta(\boldsymbol{k}_1, \boldsymbol{k}_2)\,, \tag{A.2}$$

where

$$\alpha(\boldsymbol{k}_1, \boldsymbol{k}_2) \equiv \frac{(\boldsymbol{k}_1 + \boldsymbol{k}_2)\cdot\boldsymbol{k}_1}{k_1^2} \quad \text{and} \quad \beta(\boldsymbol{k}_1, \boldsymbol{k}_2) \equiv \frac{k_{12}^2(\boldsymbol{k}_1\cdot\boldsymbol{k}_2)}{2k_1^2k_2^2}\,, \tag{A.3}$$

denote the mode coupling functions (see Sect. 2.4.3). The symmetrized versions of the second-order coupling functions are:

$$F_2^{(s)}(\boldsymbol{q}_1, \boldsymbol{q}_2) = \frac{5}{7} + \frac{2}{7}\frac{(\boldsymbol{q}_1\cdot\boldsymbol{q}_2)^2}{q_1^2q_2^2} + \frac{1}{2}\frac{\boldsymbol{q}_1\cdot\boldsymbol{q}_2}{q_1q_2}\left(\frac{q_1}{q_2} + \frac{q_2}{q_1}\right), \tag{A.4}$$

$$G_2^{(s)}(\boldsymbol{q}_1, \boldsymbol{q}_2) = \frac{3}{7} + \frac{4}{7}\frac{(\boldsymbol{q}_1\cdot\boldsymbol{q}_2)^2}{q_1^2q_2^2} + \frac{1}{2}\frac{\boldsymbol{q}_1\cdot\boldsymbol{q}_2}{q_1q_2}\left(\frac{q_1}{q_2} + \frac{q_2}{q_1}\right). \tag{A.5}$$

Since these expressions play an important role for the subsequent calculation of the trispectrum in the halo model description, we summarize their properties which follow immediately from their definitions [26]:

- $F_2^{(s)}(\boldsymbol{q}_1, -\boldsymbol{q}_1) = G_2^{(s)}(\boldsymbol{q}_1, -\boldsymbol{q}_1) = 0$,
- $F_2^{(s)}(\boldsymbol{q}_1, \boldsymbol{q}_1) = G_2^{(s)}(\boldsymbol{q}_1, \boldsymbol{q}_1) = 2$,
- $F_2^{(s)}(\boldsymbol{q}_1, \boldsymbol{q}_2) = F_2^{(s)}(-\boldsymbol{q}_1, -\boldsymbol{q}_2)$,
- $G_2^{(s)}(\boldsymbol{q}_1, \boldsymbol{q}_2) = G_2^{(s)}(-\boldsymbol{q}_1, -\boldsymbol{q}_2)$,
- $\lim\limits_{|\boldsymbol{\epsilon}|\to 0} F_2^{(s)}(\boldsymbol{q}_1, \boldsymbol{\epsilon}) = \lim\limits_{|\boldsymbol{\epsilon}|\to 0} G_2^{(s)}(\boldsymbol{q}_1, \boldsymbol{\epsilon}) = \lim\limits_{|\boldsymbol{\epsilon}|\to 0} \boldsymbol{\epsilon}/|\boldsymbol{\epsilon}|^2 \to \infty.$

A.2 Third-order coupling functions

In order to calculate the lowest non-vanishing order contribution to the perturbation theory trispectrum, we need to determine the symmetrized third-order coupling function $F_3^{(s)}$ as imposed by the recursion relation (2.26). The unsymmetrized third-order coupling function is then given by

$$\begin{aligned} F_3(\boldsymbol{q}_1, \boldsymbol{q}_2, \boldsymbol{q}_3) = {} & \frac{7}{18}\,\alpha(\boldsymbol{q}_1, \boldsymbol{q}_2 + \boldsymbol{q}_3) F_2(\boldsymbol{q}_2, \boldsymbol{q}_3) + \frac{2}{18}\,\beta(\boldsymbol{q}_1, \boldsymbol{q}_2 + \boldsymbol{q}_3) G_2(\boldsymbol{q}_2, \boldsymbol{q}_3) \\ & + G_2(\boldsymbol{q}_1, \boldsymbol{q}_2) \left[\frac{7}{18}\,\alpha(\boldsymbol{q}_1 + \boldsymbol{q}_2, \boldsymbol{q}_3) + \frac{2}{18}\,\beta(\boldsymbol{q}_1 + \boldsymbol{q}_2, \boldsymbol{q}_3) \right] . \end{aligned} \tag{A.6}$$

Note that it is not necessary to derive the corresponding expression for G_3, as we are only interested in the dark matter trispectrum. Symmetrizing Eq. (A.6) yields

$$\begin{aligned} F_3^{(s)}(\boldsymbol{q}_1, \boldsymbol{q}_2, \boldsymbol{q}_3) = {} & \frac{7}{54} \Big[\alpha(\boldsymbol{q}_1, \boldsymbol{q}_2 + \boldsymbol{q}_3) F_2^{(s)}(\boldsymbol{q}_2, \boldsymbol{q}_3) + \alpha(\boldsymbol{q}_2, \boldsymbol{q}_1 + \boldsymbol{q}_3) F_2^{(s)}(\boldsymbol{q}_1, \boldsymbol{q}_3) \\ & + \alpha(\boldsymbol{q}_3, \boldsymbol{q}_1 + \boldsymbol{q}_2) F_2^{(s)}(\boldsymbol{q}_1, \boldsymbol{q}_2) \Big] \\ & + \frac{4}{54} \Big[\beta(\boldsymbol{q}_1, \boldsymbol{q}_2 + \boldsymbol{q}_3) G_2^{(s)}(\boldsymbol{q}_2, \boldsymbol{q}_3) + \beta(\boldsymbol{q}_2, \boldsymbol{q}_1 + \boldsymbol{q}_3) G_2^{(s)}(\boldsymbol{q}_1, \boldsymbol{q}_3) \\ & + \beta(\boldsymbol{q}_3, \boldsymbol{q}_1 + \boldsymbol{q}_2) G_2^{(s)}(\boldsymbol{q}_1, \boldsymbol{q}_2) \Big] \\ & + \frac{7}{54} \Big[\alpha(\boldsymbol{q}_1 + \boldsymbol{q}_2, \boldsymbol{q}_3) G_2^{(s)}(\boldsymbol{q}_1, \boldsymbol{q}_2) + \alpha(\boldsymbol{q}_1 + \boldsymbol{q}_3, \boldsymbol{q}_2) G_2^{(s)}(\boldsymbol{q}_1, \boldsymbol{q}_3) \\ & + \alpha(\boldsymbol{q}_2 + \boldsymbol{q}_3, \boldsymbol{q}_1) G_2^{(s)}(\boldsymbol{q}_2, \boldsymbol{q}_3) \Big] \end{aligned} \tag{A.7}$$

From the symmetry properties of the F_2 and G_2 kernels and the mode coupling functions, we find $F_3^{(s)}(\boldsymbol{q}_1, \boldsymbol{q}_2, \boldsymbol{q}_3) = F_3^{(s)}(-\boldsymbol{q}_1, -\boldsymbol{q}_2, -\boldsymbol{q}_3)$.

Since we are only interested in parallelogram configurations of the trispectrum, we have to consider the situation $\boldsymbol{k}_2 = -\boldsymbol{k}_1$. In this case $F_3^{(s)}$ simplifies to

$$\begin{aligned} F_3^{(s)}(\boldsymbol{q}_1, -\boldsymbol{q}_1, \boldsymbol{q}_2) = \frac{7}{54}[\alpha(\boldsymbol{q}_1, \boldsymbol{q}_2 - \boldsymbol{q}_1)F_2^{(s)}(-\boldsymbol{q}_1, \boldsymbol{q}_2) + \alpha(-\boldsymbol{q}_1, \boldsymbol{q}_1 + \boldsymbol{q}_2)F_2^{(s)}(\boldsymbol{q}_1, \boldsymbol{q}_2)] \\ + \frac{4}{54}[\beta(\boldsymbol{q}_1, \boldsymbol{q}_2 - \boldsymbol{q}_1)G_2^{(s)}(-\boldsymbol{q}_1, \boldsymbol{q}_2) + \beta(-\boldsymbol{q}_1, \boldsymbol{q}_1 + \boldsymbol{q}_2)G_2^{(s)}(\boldsymbol{q}_1, \boldsymbol{q}_2)] \\ + \frac{7}{54}[\alpha(\boldsymbol{q}_2 - \boldsymbol{q}_1, \boldsymbol{q}_1)G_2^{(s)}(-\boldsymbol{q}_1, \boldsymbol{q}_2) + \alpha(\boldsymbol{q}_1 + \boldsymbol{q}_2, -\boldsymbol{q}_1)G_2^{(s)}(\boldsymbol{q}_1, \boldsymbol{q}_2)] \end{aligned} \tag{A.8}$$

since for each square brace in Eq. (A.7) the third term vanishes.

A.3 Perturbative Bispectrum

The lowest order, non-vanishing contribution to the perturbation theory bispectrum is

$$\begin{aligned} B_{\rm pt}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) &= 2\,[F_2(\boldsymbol{k}_1, \boldsymbol{k}_2)P(k_1)P(k_2) + F_2(\boldsymbol{k}_2, \boldsymbol{k}_3)P(k_2)P(k_3) \\ &\quad + F_2(\boldsymbol{k}_3, \boldsymbol{k}_1)P(k_3)P(k_1)]\ . \\ &= 2F_2(\boldsymbol{k}_1, \boldsymbol{k}_2)P(k_1)P(k_2) + 2\,\text{perms.}\,, \end{aligned} \tag{A.9}$$

where the permutations are considered with respect to the wave-vectors $\boldsymbol{k}_i$. For the calculation of parallelogram configurations of the trispectrum wave-vectors, we have to consider the case $\lim_{\epsilon\to\boldsymbol{0}} B_{\rm pt}(\boldsymbol{k}_i, -\boldsymbol{k}_i, \epsilon)$ and check what happens with the infrared divergences, which arise in the second-order coupling functions for $\lim\limits_{\epsilon\to\boldsymbol{0}} F_2(\boldsymbol{k}_i, \boldsymbol{\epsilon})$ (see Sect. A.1). From the tree-level bispectrum (A.9), we find

$$\begin{aligned} B_{\rm pt}(\boldsymbol{k}_i, -\boldsymbol{k}_i, \epsilon) &= 2F_2(\boldsymbol{k}_i, -\boldsymbol{k}_i)P^2(k_i) \\ &\quad + 2[F_2(\boldsymbol{k}_i, \boldsymbol{\epsilon}) + F_2(-\boldsymbol{k}_i, \boldsymbol{\epsilon})]P(k_i)P(\epsilon) \\ &= 2\left[\frac{10}{7} + \frac{4}{7}\,\frac{(\boldsymbol{k}_i \cdot \boldsymbol{\epsilon})^2}{k_i^2\epsilon^2}\right] P(k_i)P(\epsilon)\,, \end{aligned} \tag{A.10}$$

where the first term is 0, since $F_2(\boldsymbol{k}_i, -\boldsymbol{k}_i) = 0$ and the terms responsible for the infrared divergencies cancel each other. In the limit $\epsilon \to 0$, Eq. (A.10) vanishes, since the factor in front of the power spectra is finite and the $P(\epsilon)$ goes to 0 in the considered limit.

A.4 Perturbative Trispectrum

Using a perturbative approach as described in Sect. 3.5.2, we found for the trispectrum the following result:

$$
\begin{aligned}
T_{\rm pt}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\boldsymbol{k}_4) = 4\times[\quad & F_2(\boldsymbol{k}_1,-\boldsymbol{k}_{12})F_2(\boldsymbol{k}_3,\boldsymbol{k}_{12})P_1P_{12}P_3 \\
+ \quad & F_2(\boldsymbol{k}_1,-\boldsymbol{k}_{12})F_2(\boldsymbol{k}_4,\boldsymbol{k}_{12})P_1P_{12}P_4 \\
+ \quad & F_2(\boldsymbol{k}_2,-\boldsymbol{k}_{12})F_2(\boldsymbol{k}_3,\boldsymbol{k}_{12})P_2P_{12}P_3 \\
+ \quad & F_2(\boldsymbol{k}_2,-\boldsymbol{k}_{12})F_2(\boldsymbol{k}_4,\boldsymbol{k}_{12})P_2P_{12}P_4 \\
+ \quad & F_2(\boldsymbol{k}_1,-\boldsymbol{k}_{13})F_2(\boldsymbol{k}_2,\boldsymbol{k}_{13})P_1P_{13}P_2 \\
+ \quad & F_2(\boldsymbol{k}_1,-\boldsymbol{k}_{13})F_2(\boldsymbol{k}_4,\boldsymbol{k}_{13})P_1P_{13}P_4 \\
+ \quad & F_2(\boldsymbol{k}_3,-\boldsymbol{k}_{13})F_2(\boldsymbol{k}_2,\boldsymbol{k}_{13})P_3P_{13}P_2 \\
+ \quad & F_2(\boldsymbol{k}_3,-\boldsymbol{k}_{13})F_2(\boldsymbol{k}_4,\boldsymbol{k}_{13})P_3P_{13}P_4 \\
+ \quad & F_2(\boldsymbol{k}_1,-\boldsymbol{k}_{14})F_2(\boldsymbol{k}_2,\boldsymbol{k}_{14})P_1P_{14}P_2 \\
+ \quad & F_2(\boldsymbol{k}_1,-\boldsymbol{k}_{14})F_2(\boldsymbol{k}_3,\boldsymbol{k}_{14})P_1P_{14}P_3 \\
+ \quad & F_2(\boldsymbol{k}_4,-\boldsymbol{k}_{14})F_2(\boldsymbol{k}_2,\boldsymbol{k}_{14})P_4P_{14}P_2 \\
+ \quad & F_2(\boldsymbol{k}_4,-\boldsymbol{k}_{14})F_2(\boldsymbol{k}_3,\boldsymbol{k}_{14})P_4P_{14}P_3] \\
+ \quad & 6\times[F_3(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3)P_1P_2P_3 \\
+ \quad & F_3(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_4)P_1P_2P_4 \\
+ \quad & F_3(\boldsymbol{k}_1,\boldsymbol{k}_3,\boldsymbol{k}_4)P_1P_3P_4 \\
+ \quad & F_3(\boldsymbol{k}_2,\boldsymbol{k}_3,\boldsymbol{k}_4)P_2P_3P_4]\,.
\end{aligned}
$$

For the subsequent consideration, it is convenient to use the following short notation:

$$
T_{\rm pt}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\boldsymbol{k}_4) = 4T_a + 6T_b\,, \tag{A.11}
$$

where

$$
\begin{aligned}
T_a &= P(k_1)P(k_2)\,[P(k_{13})F_2(\boldsymbol{k}_1,-\boldsymbol{k}_{13})F_2(\boldsymbol{k}_2,\boldsymbol{k}_{13}) \\
&\quad + P(k_{14})F_2(\boldsymbol{k}_1,-\boldsymbol{k}_{14})F_2(\boldsymbol{k}_2,\boldsymbol{k}_{14})] + 5\,\text{perms.}\,, \qquad \text{(A.12)} \\
T_b &= F_3(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3)P(k_1)P(k_2)P(k_3) + 3\,\text{perms.}\,. \qquad \text{(A.13)}
\end{aligned}
$$

If we consider only parallelogram configurations of the trispectrum wave-vectors, we can restrict our calculations to the case where $\boldsymbol{k}_3 = -\boldsymbol{k}_1$ and $\boldsymbol{k}_4 = -\boldsymbol{k}_2$. As a consequence, making use of the properties of the coupling functions, one finds

$$
\begin{aligned}
T(\boldsymbol{k}_1,-\boldsymbol{k}_1,\boldsymbol{k}_2,-\boldsymbol{k}_2) = 4 \times [\quad & [F_2(\boldsymbol{k}_1,-(\boldsymbol{k}_1+\boldsymbol{k}_2))]^2 P_1^2 P(|\boldsymbol{k}_1+\boldsymbol{k}_2|) \\
+ \quad & 2F_2(\boldsymbol{k}_1,-(\boldsymbol{k}_1+\boldsymbol{k}_2))F_2(\boldsymbol{k}_2,-(\boldsymbol{k}_1+\boldsymbol{k}_2))P_1 P(|\boldsymbol{k}_1+\boldsymbol{k}_2|)P_2 \\
+ \quad & [F_2(\boldsymbol{k}_2,-(\boldsymbol{k}_1+\boldsymbol{k}_2))]^2 P_2^2 P(|\boldsymbol{k}_1+\boldsymbol{k}_2|) \\
+ \quad & [F_2(\boldsymbol{k}_1,-(\boldsymbol{k}_1-\boldsymbol{k}_2))]^2 P_1^2 P(|\boldsymbol{k}_1-\boldsymbol{k}_2|) \\
+ \quad & 2F_2(\boldsymbol{k}_1,-(\boldsymbol{k}_1-\boldsymbol{k}_2))F_2(-\boldsymbol{k}_2,-(\boldsymbol{k}_1-\boldsymbol{k}_2))P_1 P(|\boldsymbol{k}_1-\boldsymbol{k}_2|)P_2 \\
+ \quad & [F_2(-\boldsymbol{k}_2,-(\boldsymbol{k}_1-\boldsymbol{k}_2))]^2 P_2^2 P(|\boldsymbol{k}_1-\boldsymbol{k}_2|)] \\
+ \quad & 6 \times [F_3(\boldsymbol{k}_1,-\boldsymbol{k}_1,\boldsymbol{k}_2)P_1^2 P_2 \\
+ \quad & F_3(\boldsymbol{k}_1,-\boldsymbol{k}_1,-\boldsymbol{k}_2)P_1^2 P_2 \\
+ \quad & F_3(\boldsymbol{k}_1,\boldsymbol{k}_2,-\boldsymbol{k}_2)P_1 P_2^2 \\
+ \quad & F_3(-\boldsymbol{k}_1,\boldsymbol{k}_2,-\boldsymbol{k}_2)P_1 P_2^2] \,.
\end{aligned}
$$

A.5 1-halo term

The 1-halo term in the parallelogram configuration is trivial as it consists of only one building block term depending only on the length of the wave-vectors. If we set $\boldsymbol{k}_3 = -\boldsymbol{k}_1$ and $\boldsymbol{k}_4 = -\boldsymbol{k}_2$, we find

$$T_{1\mathrm{h}} = M_{04}(k_1, k_1, k_2, k_2) \,. \tag{A.14}$$

Note that we dropped the wave-vector dependencies on the left hand side, which we will continue to do for the subsequent halo terms as well.

A.6 2-halo term

The 2-halo term can be further subdivided into $T_{2\mathrm{h}}^{31}$, corresponding to correlations of three points in one halo and a fourth point in a second halo and $T_{2\mathrm{h}}^{22}$, describing correlations of two points in one and another two points in the second halo such that

$$T_{2\mathrm{h}} = T_{2\mathrm{h}}^{31} + T_{2\mathrm{h}}^{22} \,, \tag{A.15}$$

where

$$
\begin{aligned}
T_{\rm 2h}^{31} = & M_{13}(k_2,k_3,k_4)M_{11}(k_1)P_{\rm lin}(k_1) \\
& + M_{13}(k_1,k_2,k_3)M_{11}(k_4)P_{\rm lin}(k_4) \\
& + M_{13}(k_4,k_1,k_2)M_{11}(k_3)P_{\rm lin}(k_3) \\
& + M_{13}(k_3,k_4,k_1)M_{11}(k_2)P_{\rm lin}(k_2)
\end{aligned} \tag{A.16}
$$

and

$$
\begin{aligned}
T_{\rm 2h}^{22} = & M_{12}(k_1,k_2)M_{12}(k_3,k_4)P_{\rm lin}(|\boldsymbol{k}_1+\boldsymbol{k}_2|) \\
& + M_{12}(k_1,k_3)M_{12}(k_2,k_4)P_{\rm lin}(|\boldsymbol{k}_1+\boldsymbol{k}_3|) \\
& + M_{12}(k_1,k_4)M_{12}(k_2,k_3)P_{\rm lin}(|\boldsymbol{k}_1+\boldsymbol{k}_4|)\,.
\end{aligned} \tag{A.17}
$$

If we consider parallelogram configurations of the trispectrum where $\boldsymbol{k}_3 = -\boldsymbol{k}_1$ and $\boldsymbol{k}_4 = -\boldsymbol{k}_2$, the 2-halo terms simplify, since the 4 terms of Eq. (A.16) merge into 2

$$
\begin{aligned}
T_{\rm 2h}^{31} = & 2M_{13}(k_1,k_2,k_2)M_{11}(k_1)P_{\rm lin}(k_1) \\
& + 2M_{13}(k_1,k_1,k_2)M_{11}(k_2)P_{\rm lin}(k_2)
\end{aligned} \tag{A.18}
$$

and one term of Eq. (A.17) vanishes in the limit $\lim_{\boldsymbol{k}_3\to-\boldsymbol{k}_1} P(|\boldsymbol{k}_1+\boldsymbol{k}_3|) \to 0$.

$$
T_{\rm 2h}^{22} = M_{12}^2(k_1,k_2)\left[P_{\rm lin}(|\boldsymbol{k}_1+\boldsymbol{k}_2|) + P_{\rm lin}(|\boldsymbol{k}_1-\boldsymbol{k}_2|)\right]\,.
$$

A.7 3-halo term

The 3-halo term splits into two types of terms

$$
T_{\rm 3h} = T_{\rm 3h}^1 + T_{\rm 3h}^2\,, \tag{A.19}
$$

where

$$
\begin{aligned}
T_{\rm 3h}^1 = & M_{11}(k_3)M_{11}(k_4)M_{12}(k_1,k_2)B_{\rm pt}(\boldsymbol{k}_3,\boldsymbol{k}_4,\boldsymbol{k}_1+\boldsymbol{k}_2) \\
& + M_{11}(k_2)M_{11}(k_4)M_{12}(k_1,k_3)B_{\rm pt}(\boldsymbol{k}_2,\boldsymbol{k}_4,\boldsymbol{k}_1+\boldsymbol{k}_3) \\
& + M_{11}(k_2)M_{11}(k_3)M_{12}(k_1,k_4)B_{\rm pt}(\boldsymbol{k}_2,\boldsymbol{k}_3,\boldsymbol{k}_1+\boldsymbol{k}_4) \\
& + M_{11}(k_1)M_{11}(k_4)M_{12}(k_2,k_3)B_{\rm pt}(\boldsymbol{k}_1,\boldsymbol{k}_4,\boldsymbol{k}_2+\boldsymbol{k}_3) \\
& + M_{11}(k_1)M_{11}(k_3)M_{12}(k_2,k_4)B_{\rm pt}(\boldsymbol{k}_1,\boldsymbol{k}_3,\boldsymbol{k}_2+\boldsymbol{k}_4) \\
& + M_{11}(k_1)M_{11}(k_2)M_{12}(k_3,k_4)B_{\rm pt}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3+\boldsymbol{k}_4)
\end{aligned}
$$

with $B_{\rm pt}$ as defined in Eq. (3.57) and

$$\begin{aligned} T^2_{\rm 3h} = {} & M_{22}(k_3,k_4)M_{11}(k_1)M_{11}(k_2)P_{\rm lin}(k_1)P_{\rm lin}(k_2) \\ & + M_{22}(k_2,k_4)M_{11}(k_1)M_{11}(k_3)P_{\rm lin}(k_1)P_{\rm lin}(k_3) \\ & + M_{22}(k_2,k_3)M_{11}(k_1)M_{11}(k_4)P_{\rm lin}(k_1)P_{\rm lin}(k_4) \\ & + M_{22}(k_1,k_4)M_{11}(k_2)M_{11}(k_3)P_{\rm lin}(k_2)P_{\rm lin}(k_3) \\ & + M_{22}(k_1,k_3)M_{11}(k_2)M_{11}(k_4)P_{\rm lin}(k_2)P_{\rm lin}(k_4) \\ & + M_{22}(k_1,k_2)M_{11}(k_3)M_{11}(k_4)P_{\rm lin}(k_3)P_{\rm lin}(k_4)\,. \end{aligned} \tag{A.20}$$

If we consider parallelogram configurations of the trispectrum where $\boldsymbol{k}_3 = -\boldsymbol{k}_1$ and $\boldsymbol{k}_4 = -\boldsymbol{k}_2$, the 3-halo terms simplify in the following way:

$$\begin{aligned} T^1_{\rm 3h} = {} & 2M_{11}(k_1)M_{11}(k_2)M_{12}(k_1,k_2)B_{\rm pt}(\boldsymbol{k}_1,\boldsymbol{k}_2,-\boldsymbol{k}_1-\boldsymbol{k}_2) \\ & + 2M_{11}(k_1)M_{11}(k_2)M_{12}(k_1,k_2)B_{\rm pt}(-\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_1-\boldsymbol{k}_2) \end{aligned} \tag{A.21}$$

and

$$\begin{aligned} T^2_{\rm 3h} = {} & M_{22}(k_1,k_1)M^2_{11}(k_2)P^2_{\rm lin}(k_2) \\ & + 4M_{22}(k_1,k_2)M_{11}(k_1)M_{11}(k_2)P_{\rm lin}(k_1)P_{\rm lin}(k_2) \\ & + M_{22}(k_2,k_2)M^2_{11}(k_1)P^2_{\rm lin}(k_1)\,, \end{aligned} \tag{A.22}$$

where two terms of $T^1_{\rm 3h}$ vanish in the limits $\boldsymbol{k}_3 \to \boldsymbol{k}_1$ and $\boldsymbol{k}_4 \to \boldsymbol{k}_2$, since $\lim_{\epsilon\to\boldsymbol{0}} B_{\rm pt}(\boldsymbol{k}_i,-\boldsymbol{k}_i,\epsilon) \to 0$ and we made use of the symmetry properties of $B_{\rm pt}$ to merge the terms.

A.8 4-halo term

The 4-halo term describes correlations between points in four different halos and is given by

$$\begin{aligned} T_{\rm 4h} = {} & M_{11}(k_1)M_{11}(k_2)M_{11}(k_3)M_{11}(k_4)T_{\rm pt}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\boldsymbol{k}_4) \\ & + M_{11}(k_1)M_{11}(k_2)M_{11}(k_3)M_{21}(k_4)P_{\rm lin}(k_1)P_{\rm lin}(k_2)P_{\rm lin}(k_2) \\ & + M_{11}(k_4)M_{11}(k_3)M_{11}(k_2)M_{21}(k_1)P_{\rm lin}(k_4)P_{\rm lin}(k_3)P_{\rm lin}(k_2) \\ & + M_{11}(k_1)M_{11}(k_4)M_{11}(k_3)M_{21}(k_2)P_{\rm lin}(k_1)P_{\rm lin}(k_4)P_{\rm lin}(k_3) \\ & + M_{11}(k_2)M_{11}(k_1)M_{11}(k_4)M_{21}(k_3)P_{\rm lin}(k_2)P_{\rm lin}(k_1)P_{\rm lin}(k_4)\,, \end{aligned} \tag{A.23}$$

where $T_{\rm pt}$ denotes the lowest order non-vanishing contributions to the perturbation theory trispectrum as denoted in Eq. (A.11). If we consider parallelogram configurations of the trispectrum where $\boldsymbol{k}_3 = -\boldsymbol{k}_1$ and $\boldsymbol{k}_4 = -\boldsymbol{k}_2$, the 4-halo term simplifies in the following way:

$$\begin{aligned} T_{\rm 4h} = {} & M_{11}^2(k_1) M_{11}^2(k_2) T_{\rm pt}(\boldsymbol{k}_1, \boldsymbol{k}_2, -\boldsymbol{k}_1, -\boldsymbol{k}_2) \\ & + 2 M_{11}^2(k_1) M_{11}(k_2) M_{21}(k_2) P_{\rm lin}^2(k_1) P_{\rm lin}(k_2) \\ & + 2 M_{11}^2(k_2) M_{11}(k_1) M_{21}(k_1) P_{\rm lin}^2(k_2) P_{\rm lin}(k_1) \,, \end{aligned} \tag{A.24}$$

where $T_{\rm pt}$ is for this configuration given by Eq. (3.59).

Appendix B

Cosmology-dependent parameters

This section summarizes the cosmology dependence of the collapse density $\delta_{\rm sc}$, the virial density $\Delta_{\rm vir}$, the linear growth factor D_1 and the luminosity distance $d_{\rm L}$. Equations for the corresponding quantities in a Universe with open cosmology can be found in Henry [31].

B.1 Einstein-de Sitter ($\Omega_m = 1$)

$$\delta_{\rm sc}(z) = \frac{3(12\pi)^{2/3}}{20} \approx 1.686\,, \tag{B.1}$$

$$\Delta_{\rm vir}(z) = 18\pi^2 \approx 177.7\,, \tag{B.2}$$

$$D_1(z) = \frac{1}{1+z}\,, \tag{B.3}$$

$$d_{\rm L}(z) = 2\,\frac{c}{H_0}\left(1+z-\sqrt{1+z}\right)\,, \tag{B.4}$$

B.2 Flat models ($\Omega_m + \Omega_\Lambda = 1$)

$$x \equiv \frac{(\Omega_m^{-1}-1)^{1/3}}{1+z}\,, \tag{B.5}$$

$$\delta_{\rm sc}(z) = \frac{3(12\pi)^{2/3}}{20}\left[1-0.0123\log(1+x^3)\right]\,, \tag{B.6}$$

$$\Delta_{\rm vir}(z) = 18\pi^2(1+0.4093x^{2.71572})\,, \tag{B.7}$$

$$D_1(z) = \frac{x}{x_0}\sqrt{1+x^3}\int_0^1 \mathrm{d}y\left(1+x^3y^{6/5}\right)\,, \tag{B.8}$$

$$d_{\mathrm{L}}(z) = \frac{c}{H_0}\,(1+z)\int_0^z \frac{\mathrm{d}y}{\sqrt{\Omega_m(1+y)^3+(1-\Omega_m)}}\,, \tag{B.9}$$

Bibliography

[1] S. J. Aarseth. Dynamical evolution of clusters of galaxies, I. MNRAS, 126:223–+, 1963.

[2] R. J. Adler. *The Geometry of Random Fields*. The Geometry of Random Fields, Chichester: Wiley, 1981, 1981.

[3] D. J. Bacon, A. R. Refregier, and R. S. Ellis. Detection of weak gravitational lensing by large-scale structure. MNRAS, 318:625–640, October 2000.

[4] J. M. Bardeen, J. R. Bond, N. Kaiser, and A. S. Szalay. The statistics of peaks of Gaussian random fields. ApJ, 304:15–61, May 1986. doi: 10.1086/164143.

[5] M. Bartelmann and P. Schneider. Weak gravitational lensing. Phys. Rep., 340: 291–472, January 2001.

[6] F. Bernardeau, S. Colombi, E. Gaztañaga, and R. Scoccimarro. Large-scale structure of the Universe and cosmological perturbation theory. Phys. Rep., 367:1–3, September 2002.

[7] E. Bertschinger. Simulations of Structure Formation in the Universe. ARA&A, 36:599–654, 1998. doi: 10.1146/annurev.astro.36.1.599.

[8] J. R. Bond and S. T. Myers. The Peak-Patch Picture of Cosmic Catalogs. I. Algorithms. ApJS, 103:1–+, March 1996. doi: 10.1086/192267.

[9] J. R. Bond, S. Cole, G. Efstathiou, and N. Kaiser. Excursion set mass functions for hierarchical Gaussian fluctuations. ApJ, 379:440–460, October 1991. doi: 10.1086/170520.

[10] S. Borgani, G. Murante, V. Springel, A. Diaferio, K. Dolag, L. Moscardini, G. Tormen, L. Tornatore, and P. Tozzi. X-ray properties of galaxy clusters and groups from a cosmological hydrodynamical simulation. MNRAS, 348:1078–1096, March 2004. doi: 10.1111/j.1365-2966.2004.07431.x.

[11] R. G. Bower. The evolution of groups of galaxies in the Press-Schechter formalism. MNRAS, 248:332–352, January 1991.

[12] T. G. Brainerd, R. D. Blandford, and I. Smail. Weak Gravitational Lensing by Galaxies. ApJ, 466:623–+, August 1996. doi: 10.1086/177537.

[13] J. S. Bullock, T. S. Kolatt, Y. Sigad, R. S. Somerville, A. V. Kravtsov, A. A. Klypin, J. R. Primack, and A. Dekel. Profiles of dark haloes: evolution, scatter and environment. MNRAS, 321:559–575, March 2001.

[14] S. M. Carroll, W. H. Press, and E. L. Turner. The cosmological constant. ARA&A, 30:499–542, 1992. doi: 10.1146/annurev.aa.30.090192.002435.

[15] S. Cole and N. Kaiser. Biased clustering in the cold dark matter cosmogony. MNRAS, 237:1127–1146, April 1989.

[16] A. Cooray and W. Hu. Power Spectrum Covariance of Weak Gravitational Lensing. ApJ, 554:56–66, June 2001. doi: 10.1086/321376.

[17] A. Cooray and R. Sheth. Halo models of large scale structure. Phys. Rep., 372: 1–129, December 2002.

[18] S. Dodelson. *Modern cosmology*. Modern cosmology / Scott Dodelson. Amsterdam (Netherlands): Academic Press. ISBN 0-12-219141-2, 2003, XIII + 440 p., 2003.

[19] K. Dolag, S. Borgani, S. Schindler, A. Diaferio, and A. M. Bykov. Simulation techniques for cosmological simulations. *ArXiv e-prints*, 801, January 2008.

[20] B. Efron. *The Jackknife, the Bootstrap and other resampling plans*. CBMS-NSF Regional Conference Series in Applied Mathematics, Philadelphia: Society for Industrial and Applied Mathematics (SIAM), 1982, 1982.

[21] G. Efstathiou, C. S. Frenk, S. D. M. White, and M. Davis. Gravitational clustering from scale-free initial conditions. MNRAS, 235:715–748, December 1988.

[22] D. J. Eisenstein and W. Hu. Baryonic Features in the Matter Transfer Function. ApJ, 496:605–+, March 1998. doi: 10.1086/305424.

[23] V. R. Eke, S. Cole, and C. S. Frenk. Cluster evolution as a diagnostic for Omega. MNRAS, 282:263–280, September 1996.

[24] P. Fosalba, J. Pan, and I. Szapudi. Cosmological Three-Point Function: Testing the Halo Model against Simulations. ApJ, 632:29–48, October 2005. doi: 10.1086/432906.

[25] J. N. Fry. The Galaxy correlation hierarchy in perturbation theory. ApJ, 279: 499–510, April 1984. doi: 10.1086/161913.

[26] M. H. Goroff, B. Grinstein, S.-J. Rey, and M. B. Wise. Coupling of modes of cosmological mass density fluctuations. ApJ, 311:6–14, December 1986. doi: 10.1086/164749.

[27] M. Grossi, K. Dolag, E. Branchini, S. Matarrese, and L. Moscardini. Evolution of massive haloes in non-Gaussian scenarios. MNRAS, 382:1261–1267, December 2007. doi: 10.1111/j.1365-2966.2007.12458.x.

[28] J. E. Gunn and J. R. I. Gott. On the Infall of Matter Into Clusters of Galaxies and Some Effects on Their Evolution. ApJ, 176:1–+, August 1972.

[29] A. J. S. Hamilton. Formulae for growth factors in expanding universes containing matter and a cosmological constant. MNRAS, 322:419–425, April 2001.

[30] J. Hartlap. Studying Galaxy-Galaxy-Lensing using Ray-Tracing Simulations. Master's thesis, University of Bonn, 2005.

[31] J. P. Henry. Measuring Cosmological Parameters from the Evolution of Cluster X-Ray Temperatures. ApJ, 534:565–580, May 2000. doi: 10.1086/308783.

[32] L. Hernquist. An analytical model for spherical galaxies and bulges. ApJ, 356: 359–364, June 1990. doi: 10.1086/168845.

[33] R. W. Hockney and J. W. Eastwood. *Computer Simulation Using Particles*. Computer Simulation Using Particles, New York: McGraw-Hill, 1981, 1981.

[34] B. Jain and E. Bertschinger. Second-order power spectrum and nonlinear evolution at high redshift. ApJ, 431:495–505, August 1994. doi: 10.1086/174502.

[35] B. Jain, U. Seljak, and S. White. Ray-tracing Simulations of Weak Lensing by Large-Scale Structure. ApJ, 530:547–577, February 2000. doi: 10.1086/308384.

[36] A. Jenkins, C. S. Frenk, F. R. Pearce, P. A. Thomas, J. M. Colberg, S. D. M. White, H. M. P. Couchman, J. A. Peacock, G. Efstathiou, and A. H. Nelson. Evolution of Structure in Cold Dark Matter Universes. ApJ, 499:20–+, May 1998. doi: 10.1086/305615.

[37] A. Jenkins, C. S. Frenk, S. D. M. White, J. M. Colberg, S. Cole, A. E. Evrard, H. M. P. Couchman, and N. Yoshida. The mass function of dark matter haloes. MNRAS, 321:372–384, February 2001.

[38] Y. P. Jing. The Density Profile of Equilibrium and Nonequilibrium Dark Matter Halos. ApJ, 535:30–36, May 2000. doi: 10.1086/308809.

[39] Y. P. Jing. Correcting for the Alias Effect When Measuring the Power Spectrum Using a Fast Fourier Transform. ApJ, 620:559–563, February 2005. doi: 10.1086/427087.

[40] N. Kaiser. On the spatial correlations of Abell clusters. ApJ, 284:L9–L12, September 1984. doi: 10.1086/184341.

[41] N. Kaiser. Weak Lensing and Cosmology. ApJ, 498:26–+, May 1998. doi: 10.1086/305515.

[42] N. Kaiser, G. Wilson, and G. A. Luppino. Large-Scale Cosmic Shear Measurements. *ArXiv Astrophysics e-prints*, March 2000.

[43] M. Kendall and A. Stuart. *The advanced theory of statistics. Vol.2: Inference and relationship*. London: Griffin, 1979, 4th ed., 1979.

[44] M. Kendall and A. Stuart. *The advanced theory of statistics. Vol.1: Distribution theory*. London: Griffin, 1977, 4th ed., 1977.

[45] A. Klypin. Numerical Simulations in Cosmology I: Methods. *ArXiv Astrophysics e-prints*, May 2000.

[46] E. W. Kolb and M. S. Turner. *The early universe*. Frontiers in Physics, Reading, MA: Addison-Wesley, 1988, 1990, 1990.

[47] E. Komatsu, J. Dunkley, M. R. Nolta, C. L. Bennett, B. Gold, G. Hinshaw, N. Jarosik, D. Larson, M. Limon, L. Page, D. N. Spergel, M. Halpern, R. S. Hill, A. Kogut, S. S. Meyer, G. S. Tucker, J. L. Weiland, E. Wollack, and E. L. Wright. Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation. *ArXiv e-prints*, 803, March 2008.

[48] A. R. Liddle and D. H. Lyth. *Cosmological Inflation and Large-Scale Structure*. Cosmological Inflation and Large-Scale Structure, by Andrew R. Liddle and David H. Lyth, pp. 414. ISBN 052166022X. Cambridge, UK: Cambridge University Press, April 2000., April 2000.

[49] V. J. Martínez and E. Saar. *Statistics of galaxy clustering*, pages 143–160. Statistical Challenges in Astronomy, 2003.

[50] B. Ménard, T. Hamana, M. Bartelmann, and N. Yoshida. Improving the accuracy of cosmic magnification statistics. A&A, 403:817–828, June 2003. doi: 10.1051/0004-6361:20030406.

[51] H. J. Mo and S. D. M. White. An analytic model for the spatial clustering of dark matter haloes. MNRAS, 282:347–361, September 1996.

[52] H. J. Mo, Y. P. Jing, and S. D. M. White. High-order correlations of peaks and haloes: a step towards understanding galaxy biasing. MNRAS, 284:189–201, January 1997.

[53] P. Monaco. The Mass Function of Cosmic Structures with Nonspherical Collapse. ApJ, 447:23–+, July 1995. doi: 10.1086/175853.

[54] B. Moore, T. Quinn, F. Governato, J. Stadel, and G. Lake. Cold collapse and the core catastrophe. MNRAS, 310:1147–1152, December 1999.

[55] T. T. Nakamura and Y. Suto. Strong Gravitational Lensing and Velocity Function as Tools to Probe Cosmological Parameters — Current Constraints and Future Predictions —. *Progress of Theoretical Physics*, 97:49–+, January 1997.

[56] J. F. Navarro, C. S. Frenk, and S. D. M. White. A Universal Density Profile from Hierarchical Clustering. ApJ, 490:493–+, December 1997. doi: 10.1086/304888.

[57] A. F. Neto, L. Gao, P. Bett, S. Cole, J. F. Navarro, C. S. Frenk, S. D. M. White, V. Springel, and A. Jenkins. The statistics of Λ CDM halo concentrations. MNRAS, 381:1450–1462, November 2007. doi: 10.1111/j.1365-2966.2007.12381.x.

[58] J. Neyman and E. L. Scott. A Theory of the Spatial Distribution of Galaxies. ApJ, 116:144–+, July 1952.

[59] F. Pace, M. Maturi, M. Meneghetti, M. Bartelmann, L. Moscardini, and K. Dolag. Testing the reliability of weak lensing cluster detections. A&A, 471:731–742, September 2007. doi: 10.1051/0004-6361:20077217.

[60] J. A. Peacock. *Cosmological Physics*. Cosmological Physics, by John A. Peacock, pp. 704. ISBN 052141072X. Cambridge, UK: Cambridge University Press, January 1999., January 1999.

[61] J. A. Peacock. Large-scale surveys and cosmic structure, 2003. URL `http://www.citebase.org/abstract?id=oai:arXiv.org:astro-ph/0309240`.

[62] J. A. Peacock and S. J. Dodds. Non-linear evolution of cosmological power spectra. MNRAS, 280:L19–L26, June 1996.

[63] J. A. Peacock and A. F. Heavens. Alternatives to the Press-Schechter cosmological mass function. MNRAS, 243:133–143, March 1990.

[64] P. J. E. Peebles. Structure of the Coma Cluster of Galaxies. AJ, 75:13–+, February 1970.

[65] P. J. E. Peebles. *The large-scale structure of the universe*. Research supported by the National Science Foundation. Princeton, N.J., Princeton University Press, 1980. 435 p., 1980.

[66] W. H. Press and P. Schechter. Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation. ApJ, 187:425–438, February 1974.

[67] D. S. Reed, R. Bower, C. S. Frenk, A. Jenkins, and T. Theuns. The halo mass function from the dark ages through the present day. MNRAS, 374:2–15, January 2007. doi: 10.1111/j.1365-2966.2006.11204.x.

[68] J. Rödiger. *in preparation*. PhD thesis, University of Bonn, 2008.

[69] R. J. Scherrer and E. Bertschinger. Statistics of primordial density perturbations from discrete seed masses. ApJ, 381:349–360, November 1991. doi: 10.1086/170658.

[70] R. Scoccimarro. The Bispectrum: From Theory to Observations. ApJ, 544: 597–615, December 2000. doi: 10.1086/317248.

[71] R. Scoccimarro, S. Colombi, J. N. Fry, J. A. Frieman, E. Hivon, and A. Melott. Nonlinear Evolution of the Bispectrum of Cosmological Perturbations. ApJ, 496: 586–+, March 1998. doi: 10.1086/305399.

[72] R. Scoccimarro, M. Zaldarriaga, and L. Hui. Power Spectrum Correlations Induced by Nonlinear Clustering. ApJ, 527:1–15, December 1999. doi: 10.1086/308059.

[73] R. Scoccimarro, R. K. Sheth, L. Hui, and B. Jain. How Many Galaxies Fit in a Halo? Constraints on Galaxy Formation Efficiency from Spatial Clustering. ApJ, 546:20–34, January 2001. doi: 10.1086/318261.

[74] U. Seljak. Analytic model for galaxy and dark matter clustering. MNRAS, 318: 203–213, October 2000.

[75] U. Seljak and M. S. Warren. Large-scale bias and stochasticity of haloes and dark matter. MNRAS, 355:129–136, November 2004. doi: 10.1111/j.1365-2966.2004.08297.x.

[76] R. K. Sheth and G. Lemson. Biasing and the distribution of dark matter haloes. MNRAS, 304:767–792, April 1999.

[77] R. K. Sheth and G. Tormen. Large-scale bias and the peak background split. MNRAS, 308:119–126, September 1999.

[78] R. K. Sheth, H. J. Mo, and G. Tormen. Ellipsoidal collapse and an improved model for the number and spatial distribution of dark matter haloes. MNRAS, 323:1–12, May 2001. doi: 10.1046/j.1365-8711.2001.04006.x.

[79] R. E. Smith, J. A. Peacock, A. Jenkins, S. D. M. White, C. S. Frenk, F. R. Pearce, P. A. Thomas, G. Efstathiou, and H. M. P. Couchman. Stable clustering, the halo model and non-linear cosmological power spectra. MNRAS, 341:1311–1332, June 2003. doi: 10.1046/j.1365-8711.2003.06503.x.

[80] D. N. Spergel, R. Bean, O. Doré, M. R. Nolta, C. L. Bennett, J. Dunkley, G. Hinshaw, N. Jarosik, E. Komatsu, L. Page, H. V. Peiris, L. Verde, M. Halpern, R. S. Hill, A. Kogut, M. Limon, S. S. Meyer, N. Odegard, G. S. Tucker, J. L. Weiland, E. Wollack, and E. L. Wright. Three-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Implications for Cosmology. ApJS, 170:377–408, June 2007. doi: 10.1086/513700.

[81] V. Springel. The cosmological simulation code GADGET-2. MNRAS, 364:1105–1134, December 2005. doi: 10.1111/j.1365-2966.2005.09655.x.

[82] V. Springel, S. D. M. White, A. Jenkins, C. S. Frenk, N. Yoshida, L. Gao, J. Navarro, R. Thacker, D. Croton, J. Helly, J. A. Peacock, S. Cole, P. Thomas,

H. Couchman, A. Evrard, J. Colberg, and F. Pearce. Simulations of the formation, evolution and clustering of galaxies and quasars. Nature, 435:629–636, June 2005. doi: 10.1038/nature03597.

[83] N. Sugiyama. Cosmic Background Anisotropies in Cold Dark Matter Cosmology. ApJS, 100:281–+, October 1995. doi: 10.1086/192220.

[84] M. Takada and B. Jain. The three-point correlation function in cosmology. MNRAS, 340:580–608, April 2003. doi: 10.1046/j.1365-8711.2003.06321.x.

[85] L. Van Waerbeke, Y. Mellier, T. Erben, J. C. Cuillandre, F. Bernardeau, R. Maoli, E. Bertin, H. J. Mc Cracken, O. Le Fèvre, B. Fort, M. Dantel-Fort, B. Jain, and P. Schneider. Detection of correlated galaxy ellipticities from CFHT data: first evidence for gravitational lensing by large-scale structures. A&A, 358:30–44, June 2000.

[86] R. H. Wechsler, J. S. Bullock, J. R. Primack, A. V. Kravtsov, and A. Dekel. Concentrations of Dark Halos from Their Assembly Histories. ApJ, 568:52–70, March 2002. doi: 10.1086/338765.

[87] S. Weinberg. The cosmological constant problem. *Reviews of Modern Physics*, 61:1–23, January 1989.

[88] E. Witten. The Cosmological Constant from the Viewpoint of String Theory. In D. B. Cline, editor, *Sources and Detection of Dark Matter and Dark Energy in the Universe*, pages 27–+, 2001.

[89] D. M. Wittman, J. A. Tyson, D. Kirkman, I. Dell'Antonio, and G. Bernstein. Detection of weak gravitational lensing distortions of distant galaxies by cosmic dark matter at large scales. Nature, 405:143–148, May 2000.

[90] W.-M. Yao et al. Review of Particle Physics. *Journal of Physics G*, 33:1+, 2006. URL `http://pdg.lbl.gov`.

[91] M. Zaldarriaga and U. Seljak. CMBFAST for Spatially Closed Universes. ApJS, 129:431–434, August 2000. doi: 10.1086/313423.

[92] A. R. Zentner. The Excursion Set Theory of Halo Mass Functions, Halo Clustering, and Halo Growth. *International Journal of Modern Physics D*, 16:763–815, 2007. doi: 10.1142/S0218271807010511.

Acknowledgments

This thesis benefited a lot from the inspiration and help of many people. First of all, I thank Peter Schneider for giving me the opportunity to work in his group on a very exciting topic in cosmology and for supervising my thesis. I am grateful for his advice during my PhD time and for carefully reading the manuscript of my thesis. A special thanks goes to Peter Watts who introduced me to the halo model description of matter in the first place and helped me to become acquainted with the field of cosmology.

I gained a lot from my collaboration with Jens Rödiger and the fruitful discussions with Jan Hartlap, Martin Kilbinger and the lens group.

I thank my colleague Jan Hartlap and Francesco Pace from the University of Heidelberg for sharing their ray-tracing simulation results with me.

For the tedious work of proof-reading my thesis and making useful suggestions for improvements I thank Peter Schneider, Jens Rödiger, Martin Kilbinger, Christoph Lampert, Jan Hartlap, Jacqueline Chen, Nadya Ben Bekhti and Xinzhong Er.

Thanks to Kathy Schrüfer, Christina Stein-Schmitz and Elisabeth Danne for their help with administrative matters and to Oliver Cordes, Ole Marggraf and Günther Lay for their computer support.

I am grateful to Cristiano Porciani for accepting the co-referee of my thesis and to Manuel Drees and Martin Rumpf who agreed to participate in the commitee for the final examination.

A big thank-you to all the members from the AIfA for creating a pleasant and warm atmosphere that made work enjoyable.

This work was supported by the Deutsche Forschungsgemeinschaft through the TRR 33, “The Dark Universe”.