



ELSEVIER

Physics Reports 367 (2002) 1–248

PHYSICS REPORTS

www.elsevier.com/locate/physrep

Large-scale structure of the Universe and cosmological perturbation theory

F. Bernardeau^a, S. Colombi^b, E. Gaztañaga^{c,d}, R. Scoccimarro^{e,f,*}

^a*Service de Physique Théorique, C.E. de Saclay, F-91191 Gif-sur-Yvette Cedex, France*

^b*Institut d'Astrophysique de Paris, CNRS, 98 bis Boulevard Arago, F-75014 Paris, France*

^c*Instituto Nacional de Astrofísica, Óptica y Electrónica (INAOE), Luis Enrique Erro 1, Tonantzintla, Chohula, 78840 Puebla, Mexico*

^d*Institut d'Estudis Espacials de Catalunya, ICE/CSIC, Edf. nexus-201-c/Gran Capità 2-4, 08034 Barcelona, Spain*

^e*Department of Physics, New York University, 4 Washington Place, New York, NY 10003, USA*

^f*Institute for Advanced Study, School of Natural Sciences, Einstein Drive, Princeton, NJ 08540, USA*

Received 15 December 2001

editor: M.P. Kamionkowski

Abstract

We review the formalism and applications of non-linear perturbation theory (PT) to understanding the large-scale structure of the Universe. We first discuss the dynamics of gravitational instability, from the linear to the non-linear regime. This includes Eulerian and Lagrangian PT, non-linear approximations, and a brief description of numerical simulation techniques. We then cover the basic statistical tools used in cosmology to describe cosmic fields, such as correlation functions in real and Fourier space, probability distribution functions, cumulants and generating functions. In subsequent sections we review the use of PT to make quantitative predictions about these statistics according to initial conditions, including effects of possible non-Gaussianity of the primordial fields. Results are illustrated by detailed comparisons of PT predictions with numerical simulations. The last sections deal with applications to observations. First, we review in detail practical estimators of statistics in galaxy catalogs and related errors, including traditional approaches and more recent developments. Then, we consider the effects of the bias between the galaxy distribution and the matter distribution, the treatment of redshift distortions in three-dimensional surveys and of projection effects in angular catalogs, and some applications to weak gravitational lensing. We finally review the current observational situation regarding statistics in galaxy catalogs and what the future generation of galaxy surveys promises to deliver. © 2002 Elsevier Science B.V. All rights reserved.

PACS: 98.80.-k; 98.80.Hw

* Corresponding author. Department of Physics, New York University, 4 Washington Place, New York, NY 10003, USA.

E-mail address: rs123@nyu.edu (R. Scoccimarro).

Contents

1. Introduction and notation	5
2. Dynamics of gravitational instability	7
2.1. The Vlasov equation	10
2.2. Eulerian dynamics	11
2.3. Eulerian linear perturbation theory	12
2.4. Eulerian non-linear perturbation theory	14
2.4.1. The equations of motion in the Fourier representation	15
2.4.2. General solutions in Einstein–de Sitter cosmology	15
2.4.3. Cosmology dependence of non-linear growth factors	18
2.4.4. Approximate solutions in arbitrary cosmology	19
2.4.5. The density and velocity fields up to third order	20
2.4.6. Non-linear growing and decaying modes	21
2.5. Lagrangian dynamics	22
2.6. Linear solutions and the Zel’dovich approximation	23
2.7. Lagrangian perturbation theory	24
2.8. Non-linear approximations	25
2.9. Numerical simulations	27
2.9.1. Introduction	27
2.9.2. Direct summation	28
2.9.3. The tree algorithm	28
2.9.4. The PM algorithm	29
2.9.5. Hybrid methods	29
3. Random cosmic fields and their statistical description	30
3.1. The need for a statistical approach	30
3.1.1. Physical origin of fluctuations from inflation	31
3.1.2. Physical origin of fluctuations from topological defects	32
3.2. Correlation functions and power spectra	32
3.2.1. Statistical homogeneity and isotropy	32
3.2.2. Two-point correlation function and power spectrum	33
3.2.3. The Wick theorem for Gaussian fields	33
3.2.4. Higher-order correlators: diagrammatics	34
3.2.5. Probabilities and correlation functions	36
3.3. Moments, cumulants and their generating functions	36
3.3.1. Moments and cumulants	36
3.3.2. Smoothing	37
3.3.3. Generating functions	37
3.4. Probability distribution functions	38
3.5. Weakly non-Gaussian distributions: Edgeworth expansion	38
4. From dynamics to statistics: N -point results	40
4.1. The weakly non-linear regime: “tree-level” PT	40
4.1.1. Emergence of non-Gaussianity	40
4.1.2. Power spectrum evolution in linear PT	42
4.1.3. The bispectrum induced by gravity	43
4.1.4. The three-point correlation function	45
4.2. The transition to the non-linear regime: “loop corrections”	46
4.2.1. One-loop PT and previrialization	46
4.2.2. The one-loop power spectrum	47
4.2.3. The one-loop bispectrum	50
4.3. The power spectrum in the Zel’dovich approximation	54

4.4.	Non-Gaussian initial conditions	55
4.4.1.	General results	55
4.4.2.	χ^2 initial conditions	56
4.5.	The strongly non-linear regime	59
4.5.1.	The self-similar solution	59
4.5.2.	Stable clustering	60
4.5.3.	Scale invariance	62
4.5.4.	The non-linear evolution of two-point statistics	63
4.5.5.	The hierarchical models	64
4.5.6.	Hyperextended perturbation theory	66
5.	From dynamics to statistics: the local cosmic fields	68
5.1.	The density field third moment: skewness	68
5.1.1.	The unsmoothed case	68
5.1.2.	The smoothed case	69
5.1.3.	Physical interpretation of smoothing	71
5.1.4.	Dependence of the skewness on cosmological parameters	72
5.1.5.	The skewness of the local velocity divergence	73
5.2.	The fourth-order density cumulant: kurtosis	73
5.3.	Results for Gaussian smoothing filters	74
5.4.	The density cumulants hierarchy	75
5.4.1.	The unsmoothed density cumulant generating function	75
5.4.2.	Geometrical properties of smoothing in Lagrangian space	77
5.4.3.	Lagrangian to Eulerian space mapping: smoothed case	79
5.5.	One-loop corrections to one-point moments	80
5.5.1.	Exact results	81
5.5.2.	The spherical collapse model approximation	82
5.6.	Evolution from non-Gaussian initial conditions	84
5.7.	Transients from initial conditions	87
5.8.	The density PDF	92
5.8.1.	Reconstruction of the PDF from the generating function	92
5.8.2.	Dependence on cosmological parameters	94
5.8.3.	The PDF in the Zel'dovich approximation	95
5.9.	Two-dimensional dynamics	96
5.10.	The velocity divergence PDF	97
5.10.1.	The velocity divergence cumulants hierarchy	97
5.10.2.	The shape of the PDF	99
5.10.3.	Comparison with N -body simulations	99
5.11.	The velocity–density relation	100
5.12.	The two-point density PDF	103
5.13.	Extended perturbation theories	105
6.	From theory to observations: estimators and errors	107
6.1.	Introduction	107
6.2.	Basic concepts	110
6.2.1.	Cosmic bias and cosmic error	110
6.2.2.	The covariance matrix	110
6.3.	Fair sample hypothesis and local Poisson approximation	111
6.3.1.	The fair sample hypothesis	111
6.3.2.	Poisson realization of a continuous field	112
6.4.	The two-point correlation function	113
6.4.1.	Estimators	114
6.4.2.	Cosmic bias and integral constraint of the LS estimator	116

6.4.3.	Cosmic error of the LS estimator	117
6.4.4.	The covariance matrix	118
6.4.5.	Recipes for error calculations	119
6.5.	The power spectrum	120
6.5.1.	Simple estimators	120
6.5.2.	Cosmic bias and integral constraint	121
6.5.3.	The cosmic error	122
6.5.4.	The covariance matrix	123
6.6.	Generalization to higher-order correlation functions	125
6.7.	One-point distributions: counts-in-cells	126
6.7.1.	Definitions	126
6.7.2.	Estimators	127
6.7.3.	Error propagation: cosmic bias vs. cosmic error	128
6.7.4.	Cosmic error and cross-correlations of factorial moments	129
6.7.5.	Cosmic error and cosmic bias of cumulants	131
6.8.	Multivariate count-in-cells	134
6.9.	Optimal weighting	135
6.10.	Cosmic distribution function and cross-correlations	138
6.10.1.	Cosmic distribution function and likelihood	138
6.10.2.	Cross-correlations between different statistics	140
6.10.3.	Validity of the Gaussian approximation	140
6.11.	Optimal techniques for Gaussian random fields	143
6.11.1.	Maximum likelihood estimates	143
6.11.2.	Quadratic estimators	145
6.11.3.	Uncorrelated error bars	147
6.11.4.	Data compression and the Karhunen–Loève transform	148
6.12.	Measurements in N -body simulations	149
6.12.1.	Cosmic error and cosmic bias in simulations	149
6.12.2.	N -body relaxation and force softening	150
7.	Applications to observations	150
7.1.	The problem of galaxy biasing	150
7.1.1.	Some general results	151
7.1.2.	Halo clustering in the tree hierarchical model	156
7.1.3.	Halo clustering in the extended Press–Schechter approach	158
7.1.4.	Galaxy clustering	159
7.2.	Projection effects	161
7.2.1.	The projected density contrast	162
7.2.2.	The small-angle approximation	163
7.2.3.	Projection in the hierarchical model	164
7.2.4.	The correlation hierarchy for the projected density	166
7.2.5.	Comparison with numerical simulations	168
7.2.6.	Reconstructing the PDF of the projected density	169
7.3.	Weak gravitational lensing	170
7.3.1.	The convergence field as a projected mass map	171
7.3.2.	Statistical properties	172
7.3.3.	Next to leading order effects	172
7.3.4.	Biasing from weak gravitational lensing	173
7.4.	Redshift distortions	173
7.4.1.	The density field in redshift space	174
7.4.2.	The redshift-space power spectrum	176
7.4.3.	The redshift-space bispectrum	178

7.4.4. The higher-order moments in redshift space	180
7.4.5. Cosmological distortions	181
8. Results from galaxy surveys	182
8.1. Galaxies as cosmological tracers	182
8.2. Results from angular galaxy surveys	184
8.2.1. Angular catalogs	184
8.2.2. The angular correlation function and power spectrum	187
8.2.3. Inversion from angular to 3D clustering	189
8.2.4. Three-point statistics and higher order	192
8.2.5. Skewness, kurtosis and higher-order cumulants	198
8.2.6. Constraints on biasing and primordial non-Gaussianity	199
8.3. Results from redshift galaxy surveys	201
8.3.1. Redshift catalogs	201
8.3.2. Two-point statistics	203
8.3.3. Three-point statistics	206
8.3.4. Skewness, kurtosis and higher-order cumulants	209
8.3.5. Constraints on biasing and primordial non-Gaussianity	212
8.4. Recent results from 2dFGRS and SDSS	214
9. Summary and conclusions	216
Acknowledgements	218
Appendix A. The spherical collapse dynamics	218
Appendix B. Tree summations	220
B.1. For one field	220
B.2. For two fields	221
B.3. The large separation limit	222
Appendix C. Geometrical properties of top-hat window functions	223
Appendix D. One-loop calculations: dimensional regularization	224
Appendix E. PDF construction from cumulant generating function	226
E.1. Counts-in-cells and generating functions	226
E.2. The continuous limit	227
E.3. Approximate forms for $P(\rho)$ when $\zeta_{\text{int}} \ll 1$	227
E.4. Approximate forms for $P(\rho)$ when $\zeta_{\text{int}} \gg 1$	228
E.5. Numerical computation of the Laplace inverse transform	229
Appendix F. Cosmic errors: expressions for the factorial moments	230
F.1. Method	230
F.2. Analytic results	231
References	233

1. Introduction and notation

Understanding the large-scale structure of the universe is one of the main goals of cosmology. In the last two decades it has become widely accepted that gravitational instability plays a central role in giving rise to the remarkable structures seen in galaxy surveys. Extracting the wealth of information contained in galaxy clustering to learn about cosmology thus requires a quantitative understanding of the dynamics of gravitational instability and application of sophisticated statistical tools that can best be used to test theoretical models against observations.

In this work we review the use of non-linear cosmological perturbation theory (hereafter PT) to accomplish this goal. The usefulness of PT in interpreting results from galaxy surveys is based on the fact that in the gravitational instability scenario density fluctuations become small enough at large scales (the so-called “weakly non-linear regime”) that a perturbative approach suffices to understand their evolution. Since early developments in the 1980s, PT has gone through a period of rapid evolution in the last decade which gave rise to numerous useful results. Given the imminent completion of next-generation large-scale galaxy surveys ideal for applications of PT, it seems timely to provide a comprehensive review of the subject.

The purpose of this review is twofold:

- (1) To summarize the most important theoretical results, which are sometimes rather technical and appeared somewhat scattered in the literature with often fluctuating notation, in a clear, consistent and unified fashion. We tried, in particular, to unveil approximations that might have been overlooked in the original papers, and to highlight the outstanding theoretical issues that remain to be addressed.
- (2) To present the state-of-the-art observational knowledge of galaxy clustering with particular emphasis on constraints derived from higher-order statistics on galaxy biasing and primordial non-Gaussianity, and give a rigorous basis for the confrontation of theoretical results with observational data from upcoming galaxy catalogues.

We assume throughout this review that the universe satisfies the standard homogeneous and isotropic big-bang model. The framework of gravitational instability, in which PT is based, assumes that gravity is the only agent at large scales responsible for the formation of structures in a universe with density fluctuations dominated by dark matter. This assumption is in very good agreement with observations of galaxy clustering, in particular, as we discuss in detail here, from higher-order statistics which are sensitive to the detailed structure of the dynamics responsible for large-scale structures.¹ The non-gravitational effects associated with galaxy formation may alter the distribution of luminous matter compared to that of the underlying dark matter, in particular at small scales: such “galaxy biasing” can be probed with the techniques reviewed in this work.

Inevitably, we had to make some decisions in the choice of topics to be covered. Our presentation is definitely focused on the density field, with much less coverage on peculiar velocities. This choice is, in particular, motivated by the comparatively still preliminary stage of cosmic velocity fields, at least from an observational point of view (see however [607,160] for a review). On the other hand, note that since velocity field results are often obtained by identical techniques to those used for the density field, we mention some of these results but without giving them their due importance.

In order to fully characterize the density field, we choose to follow the traditional approach of using statistical methods, in particular, N -point correlation functions [508]. Alternative methods include morphological descriptors such as Minkowski functionals (of which the genus is perhaps the most widely known), percolation analysis, etc. Unlike correlation functions, however, these other statistics are not as directly linked to dynamics as correlation functions, and thus are not as easy

¹ As opposed to just properties of the linearized equations of motion, which can be mimicked by non-gravitational theories of structure formation in some cases [10].

to predict from theoretical models. Furthermore, applications of PT to make predictions of these quantities is still in its infancy (see e.g. [441] and references therein for recent work).

Given that PT is an approximate method to solve the dynamics of gravitational clustering, it is desirable to test the validity of the results with other techniques. In particular, we resort to numerical simulations, which involve different approximations in solving the equations of motion that are not restricted to the weakly non-linear regime. There is a strong and healthy interplay between PT and N -body simulations which we extensively illustrate throughout this review. At large scales PT can be used to test quantitatively for spurious effects in numerical simulations (e.g. finite volume effects, transients from initial conditions), whereas at smaller, non-linear scales N -body simulations can be used to investigate the regime of validity of PT predictions.

Although reviewing the current understanding of clustering at small scales is beyond the scope of this review, we have also included a discussion of the predictions of non-linear clustering amplitudes because connections between PT and strongly non-linear behavior have been suggested in the literature. We also include a discussion about stable clustering at small scales which, when coupled with self-similarity, leads to a connection between the large- and small-scale scaling behavior of correlations functions.

This review is structured so that different chapters can be read independently, although there are inevitable relations. Section 2 deals with the basic equations of motion and their solution in PT, including a brief summary of numerical simulations. Section 3 is a review of the basics of statistics; we have made it as succinct as possible to swiftly introduce the reader to the core of the review. For a more in-depth treatment we refer the reader to [609,61]. The next two sections represent the main theoretical results; Section 4 deals with N -point functions, whereas Section 5 reviews results for the smoothed one-point moments and PDFs. These two sections heavily rely on material covered in Sections 2 and 3.

In Section 6 we describe in detail the standard theory of estimators and errors for application to galaxy surveys, with particular attention to the issue of cosmic bias and errors of estimators of the two-point correlation function, power spectrum and higher-order moments such as the skewness. Section 7 deals with theoretical issues related to surveys, such as redshift distortions, projection effects, galaxy biasing and weak gravitational lensing. Section 8 presents the current observational status of galaxy clustering, including future prospects in upcoming surveys, with particular emphasis on higher-order statistics. Section 9 contains our conclusions and outlook. A number of appendices extend the material in the main text for those interested in carrying out detailed calculations. Finally, to help the reader, Tables 1–4 list the main abbreviations and notations used for various cosmological variables, fields and statistics.

2. Dynamics of gravitational instability

The most natural explanation for the large-scale structures seen in galaxy surveys (e.g. superclusters, walls and filaments) is that they are the result of gravitational amplification of small primordial fluctuations due to the gravitational interaction of collisionless cold dark matter (CDM) particles in an expanding universe [509,75,173,174]. Throughout this review we will assume this framework and discuss how PT can be used to understand the physics of gravitational instability and test this hypothesis against observations.

Table 1
Abbreviations

PT	Perturbation theory
2LPT	Second-order Lagrangian perturbation theory
EPT	Extended perturbation theory
HEPT	Hyperextended perturbation theory
ZA	Zel'dovich approximation
SC	Spherical collapse
CDM	Cold dark matter (model)
SCDM	Standard CDM model
Λ CDM	Flat CDM model with a cosmological constant
PDF	Probability distribution function
CPDF	Count probability distribution function

Table 2
Notation for various cosmological variables

Ω_m	The total matter density in units of critical density
Ω_Λ	The reduced cosmological constant
Ω_{tot}	The total energy density of the universe in units of critical density, $\Omega_{\text{tot}} = \Omega_m + \Omega_\Lambda$
H	The Hubble constant
h	The Hubble constant at present time, in units of 100 km/s/Mpc, $h \equiv H_0/100$
a	The scale factor
τ	The conformal time, $d\tau = dt/a$
\mathcal{H}	The conformal expansion rate, $\mathcal{H} = aH$
D_1	The linear growth factor
D_n	The n th-order growth factor
$f(\Omega_m, \Omega_\Lambda)$	The logarithmic derivative of (the fastest growing mode of) the linear growth factor with respect to a : $f(\Omega_m, \Omega_\Lambda) \equiv d \ln D_1 / d \ln a$

Table 3
Notation for the cosmic fields

\tilde{X}	The Fourier transform of field X ; $\tilde{X}(\mathbf{k}) = (2\pi)^{-3} \int d^3\mathbf{x} e^{-i\mathbf{k}\cdot\mathbf{x}} X(\mathbf{x})$ (except in Section 6.5)
\mathbf{x}	The comoving position in real space
$\rho(\mathbf{x})$	The local cosmic density
$\delta(\mathbf{x})$	The local density contrast, $\delta = \rho/\bar{\rho} - 1$
$\Phi(\mathbf{x})$	The gravitational potential
$\mathbf{u}(\mathbf{x})$	The local peculiar velocity field
$\theta(\mathbf{x})$	The local velocity divergence in units of $\mathcal{H} = aH$
$F_p(\mathbf{k}_1, \dots, \mathbf{k}_p)$	The p th-order density field kernel
$G_p(\mathbf{k}_1, \dots, \mathbf{k}_p)$	The p th-order velocity divergence field kernel
$\psi(\mathbf{q})$	The Lagrangian displacement field
$J(\mathbf{q})$	The Jacobian of the Lagrangian–Eulerian mapping

Although the nature of dark matter has not yet been identified, all candidates for CDM particles are extremely light compared to the mass scale of typical galaxies, with expected number densities of at least 10^{50} particles/Mpc³ [383]. In this limit where the number of particles $N \gg 1$, discreteness effects

Table 4
Notation for statistical quantities

$P(k)$	The density power spectrum
$\Delta(k)$	The dimensionless power, $\Delta = 4\pi k^3 P(k)$
$B(k_1, k_2, k_3)$	The bispectrum
$P_N(\mathbf{k}_1, \dots, \mathbf{k}_N)$	The N -point polyspectrum
P_N	The count-in-cell probability distribution function
$p(\delta) d\delta$	The cosmic density probability distribution function
F_k	The factorial moment of order k
$\xi_2(\mathbf{x}_1, \mathbf{x}_2) \equiv \xi_{12} \equiv \xi$	The two-point correlation function, $\xi_2(\mathbf{x}_1, \mathbf{x}_2) = \langle \delta(\mathbf{x}_1)\delta(\mathbf{x}_2) \rangle = \langle \delta(\mathbf{x}_1)\delta(\mathbf{x}_2) \rangle_c$
$\sigma^2 \equiv \bar{\xi} \equiv \bar{\xi}_2$	The cell-average <i>two</i> -point correlation function
σ_8	The value of the (linearly extrapolated) σ in a sphere of $8h^{-1}$ Mpc radius
Γ	Shape parameter of the linear power spectrum, $\Gamma \simeq \Omega_m h$
$\xi_N(\mathbf{x}_1, \dots, \mathbf{x}_N)$	The N -point correlation functions $\xi_N(\mathbf{x}_1, \dots, \mathbf{x}_N) = \langle \delta(\mathbf{x}_1) \cdots \delta(\mathbf{x}_N) \rangle_c$
$w_N(\theta_1, \dots, \theta_N)$	The angular N -point correlation functions
$\bar{\xi}_N$	The cell-averaged N -point correlation functions $\bar{\xi}_N = \langle \delta_R^N \rangle_c$
\bar{w}_N	The cell-averaged angular N -point correlation functions
S_p	The density normalized cumulants, $S_p = \langle \delta_R^p \rangle_c / \langle \delta_R^2 \rangle_c^{p-1} = \bar{\xi}_p / \bar{\xi}^{p-1}$
S_3, S_4	The (reduced) skewness/kurtosis
s_p	The projected density normalized cumulants
$Q \equiv Q_3, \tilde{Q} \equiv \tilde{Q}_3$	The three-point hierarchical amplitude in real/Fourier space
Q_N, \tilde{Q}_N	The N -point hierarchical amplitude in real/Fourier space; Q_N can also stand for S_N/N^{N-2} (Section 6)
q_N, \tilde{q}_N	The projected N -point hierarchical amplitude in real/Fourier space can also stand for s_N/N^{N-2} (Section 6)
T_p	The velocity divergence normalized cumulants
C_{pq}	The two-point density normalized cumulants, $C_{pq} = \langle \delta_1^p \delta_2^q \rangle_c / (\xi_{12} \langle \delta^2 \rangle_c^{p+q-2})$
$\varphi(y)$	The one-point cumulant generating function, $\varphi(y) = \sum_p S_p (-y)^p / p!$
v_p, μ_p	The density/velocity field vertices
$\mathcal{G}_\delta(\tau) \equiv \mathcal{G}_\delta^L(\tau),$ $\mathcal{G}_\theta(\tau) \equiv \mathcal{G}_\theta^L(\tau)$	The vertex generating function for the density/velocity field, $\mathcal{G}_\delta(\tau) \equiv \sum_{p \geq 1} v_p (-\tau)^p / p!$ and $\mathcal{G}_\theta(\tau) \equiv -f(\Omega_m, \Omega_\Lambda) \sum_{p \geq 1} v_p (-\tau)^p / p!$
$\langle X \rangle$	The ensemble average of statistic X
\hat{X}	The estimator of statistic X
$\Upsilon(\hat{X}) d\hat{X}$	The cosmic distribution function of estimator \hat{X}
ΔX	The cosmic error on estimator \hat{X}

such as two-body relaxation (important e.g. in globular clusters) are negligible, and collisionless dark matter² obeys the Vlasov equation for the distribution function in phase space, Eq. (12). This is the master equation from which all subsequent calculations of gravitational instability are derived.

Since CDM particles are non-relativistic, at scales much smaller than the Hubble radius the equations of motion reduce essentially to those of Newtonian gravity.³ The expansion of the universe simply calls for a redefinition of the variable used to describe the position and momentum of particles, and a redefinition of the gravitational potential. For a detailed discussion of the Newtonian

² There has been recently a renewed interest in studying *collisional* dark matter [600,700,170], which may help solve some problems with collisionless CDM at small scales, of order few kpc.

³ A detailed treatment of relativistic linear PT of gravitational instability can be found in [19,466,400].

limit from general relativity see e.g. [508]. We will simply motivate the results without giving a derivation.

2.1. The Vlasov equation

Let us consider a set of particles of mass m that interact only gravitationally in an expanding universe. The equation of motion for a particle of velocity \mathbf{v} at position \mathbf{r} is thus,

$$\frac{d\mathbf{v}}{dt} = Gm \sum_i \frac{\mathbf{r}_i - \mathbf{r}}{|\mathbf{r}_i - \mathbf{r}|^3}, \quad (1)$$

where the summation is made over all other particles at position r_i .

In the limit of a large number of particles, this equation can be rewritten in terms of a smooth gravitational potential due to the particle distribution,

$$\frac{d\mathbf{v}}{dt} = -\frac{\partial\phi}{\partial\mathbf{r}}, \quad (2)$$

where ϕ is the Newtonian potential induced by the local mass density $\rho(\mathbf{r})$,

$$\phi(\mathbf{r}) = G \int d^3\mathbf{r}' \frac{\rho(\mathbf{r}')}{|\mathbf{r}' - \mathbf{r}|}. \quad (3)$$

In the context of gravitational instabilities in an expanding universe we have to consider the departures from the homogeneous Hubble expansion. Positions of particles are described by their *comoving* coordinates \mathbf{x} such that the physical coordinates are $\mathbf{r} = a(\tau)\mathbf{x}$ where a is the cosmological scale factor. We choose to describe the equations of motion in terms of the conformal time τ related to cosmic time by $dt = a(\tau)d\tau$. The equations of motion that follow are valid in an arbitrary homogeneous and isotropic background universe, which evolves according to Friedmann equations:

$$\frac{\partial\mathcal{H}(\tau)}{\partial\tau} = -\frac{\Omega_m(\tau)}{2} \mathcal{H}^2(\tau) + \frac{\Lambda}{3} a^2(\tau) \equiv \left(\Omega_\Lambda(\tau) - \frac{\Omega_m(\tau)}{2} \right) \mathcal{H}^2(\tau), \quad (4)$$

$$(\Omega_{\text{tot}}(\tau) - 1) \mathcal{H}^2(\tau) = k, \quad (5)$$

where $\mathcal{H} \equiv d\ln a/d\tau = Ha$ is the conformal expansion rate, H is the Hubble constant, Ω_m is the ratio of matter density to critical density, Λ is the cosmological constant and $k = -1, 0, 1$ for $\Omega_{\text{tot}} < 1$, $\Omega_{\text{tot}} = 1$ and $\Omega_{\text{tot}} > 1$, respectively ($\Omega_{\text{tot}} \equiv \Omega_m + \Omega_\Lambda$). Note that Ω_m and Ω_Λ are time dependent.

We then define the density contrast $\delta(\mathbf{x})$ by

$$\rho(\mathbf{x}, \tau) \equiv \bar{\rho}(\tau)[1 + \delta(\mathbf{x}, \tau)], \quad (6)$$

the peculiar velocity \mathbf{u} with

$$\mathbf{v}(\mathbf{x}, \tau) \equiv \mathcal{H}\mathbf{x} + \mathbf{u}(\mathbf{x}, \tau) \quad (7)$$

and the cosmological gravitational potential Φ with

$$\phi(\mathbf{x}, \tau) \equiv -\frac{1}{2} \frac{\partial \mathcal{H}}{\partial \tau} x^2 + \Phi(\mathbf{x}, \tau) \quad (8)$$

so that the latter is sourced only by density fluctuations, as expected; indeed the Poisson equation reads

$$\nabla^2 \Phi(\mathbf{x}, \tau) = \frac{3}{2} \Omega_m(\tau) \mathcal{H}^2(\tau) \delta(\mathbf{x}, \tau) . \quad (9)$$

In the following we will only use comoving coordinates as the spatial variable so that all space derivatives should be understood as done with respect to \mathbf{x} .

The equation of motion (Eq. (2)) then reads

$$\frac{d\mathbf{p}}{d\tau} = -am \nabla \Phi(\mathbf{x}) \quad (10)$$

with

$$\mathbf{p} = am \mathbf{u} . \quad (11)$$

Let us now define the particle number density in phase space by $f(\mathbf{x}, \mathbf{p}, \tau)$; phase-space conservation implies the Vlasov equation

$$\frac{df}{d\tau} = \frac{\partial f}{\partial \tau} + \frac{\mathbf{p}}{ma} \cdot \nabla f - am \nabla \Phi \cdot \frac{\partial f}{\partial \mathbf{p}} = 0 . \quad (12)$$

Needless to say, this equation is very difficult to solve, being a non-linear partial differential equation involving seven variables. The non-linearity is induced by the fact that the potential Φ depends through Poisson equation on the integral of the distribution function over momentum (which gives the density field, see Eq. (13)).

2.2. Eulerian dynamics

In practice, however, we are usually not interested in solving the full phase-space dynamics, but rather the evolution of the spatial distribution. This can be conveniently obtained by taking momentum moments of the distribution function. The zeroth-order moment simply relates the phase-space density to the local mass density field,

$$\int d^3 \mathbf{p} f(\mathbf{x}, \mathbf{p}, \tau) \equiv \rho(\mathbf{x}, \tau) . \quad (13)$$

The next order moments,

$$\int d^3 \mathbf{p} \frac{\mathbf{p}}{am} f(\mathbf{x}, \mathbf{p}, \tau) \equiv \rho(\mathbf{x}, \tau) \mathbf{u}(\mathbf{x}, \tau) , \quad (14)$$

$$\int d^3 \mathbf{p} \frac{p_i p_j}{a^2 m^2} f(\mathbf{x}, \mathbf{p}, \tau) \equiv \rho(\mathbf{x}, \tau) \mathbf{u}_i(\mathbf{x}, \tau) \mathbf{u}_j(\mathbf{x}, \tau) + \sigma_{ij}(\mathbf{x}, \tau) , \quad (15)$$

define the *peculiar velocity flow* $\mathbf{u}(\mathbf{x}, \tau)$ and the *stress tensor* $\sigma_{ij}(\mathbf{x}, \tau)$. The equation for these fields follow from taking moments of the Vlasov equation. The zeroth moment gives the continuity equation

$$\frac{\partial \delta(\mathbf{x}, \tau)}{\partial \tau} + \nabla \cdot \{ [1 + \delta(\mathbf{x}, \tau)] \mathbf{u}(\mathbf{x}, \tau) \} = 0, \quad (16)$$

which describes conservation of mass. Taking the first moment of Eq. (12) and subtracting $\mathbf{u}(\mathbf{x}, \tau)$ times the continuity equation we obtain the Euler equation

$$\frac{\partial \mathbf{u}(\mathbf{x}, \tau)}{\partial \tau} + \mathcal{H}(\tau) \mathbf{u}(\mathbf{x}, \tau) + \mathbf{u}(\mathbf{x}, \tau) \cdot \nabla \mathbf{u}(\mathbf{x}, \tau) = -\nabla \Phi(\mathbf{x}, \tau) - \frac{1}{\rho} \nabla_j (\rho \sigma_{ij}), \quad (17)$$

which describes conservation of momentum. Note that the continuity equation couples the zeroth (ρ) to the first moment (\mathbf{u}) of the distribution function, the Euler equation couples the first moment (\mathbf{u}) to the second moment (σ_{ij}), and so on. However, having integrated out the phase-space information, we are here in a more familiar ground, and we have reasonable phenomenological models to close the hierarchy by postulating an ansatz for the stress tensor σ_{ij} , i.e. the *equation of state* of the cosmological fluid. For example, standard fluid dynamics [392] gives $\sigma_{ij} = -p\delta_{ij} + \eta(\nabla_i u_j + \nabla_j u_i - \frac{2}{3} \delta_{ij} \nabla \cdot \mathbf{u}) + \zeta \delta_{ij} \nabla \cdot \mathbf{u}$, where p denotes the pressure and η and ζ are viscosity coefficients.

The equation of state basically relies on the assumption that cosmological structure formation is driven by matter with negligible velocity dispersion or pressure, as for example CDM. Note that from its definition, Eq. (15), the stress tensor characterizes the deviation of particle motions from a single coherent flow (single stream), for which the first term will be the dominant contribution. Therefore, it is a good approximation to set $\sigma_{ij} \approx 0$, at least in the first stages of gravitational instability when structures did not have time to collapse and virialize. As time goes on, this approximation will break down at progressively larger scales, but we will see that at present times at the scales relevant to large-scale structure, a great deal can be explored and understood using this simple approximation. In particular, the breakdown of $\sigma_{ij} \approx 0$ describes the generation of velocity dispersion (or even anisotropic pressure) due to multiple streams, generically known as *shell crossing*. We will discuss this issue further below.

We now turn to a systematic investigation of the solutions of Eqs. (9), (16) and (17) for vanishing stress tensor.

2.3. Eulerian linear perturbation theory

At large scales, where we expect the universe to become smooth, the fluctuation fields in Eqs. (6)–(8) can be assumed to be small compared to the homogeneous contribution described by the first terms. Therefore, it follows that we can linearize Eqs. (9), (16) and (17) to obtain the equations of motion in the *linear regime*:

$$\frac{\partial \delta(\mathbf{x}, \tau)}{\partial \tau} + \theta(\mathbf{x}, \tau) = 0, \quad (18)$$

$$\frac{\partial \mathbf{u}(\mathbf{x}, \tau)}{\partial \tau} + \mathcal{H}(\tau) \mathbf{u}(\mathbf{x}, \tau) = -\nabla \Phi(\mathbf{x}, \tau), \quad (19)$$

where $\theta(\mathbf{x}, \tau) \equiv \nabla \cdot \mathbf{u}(\mathbf{x}, \tau)$ is the divergence of the velocity field. These equations are now straightforward to solve. The velocity field, as any vector field, can be completely described by its divergence

$\theta(\mathbf{x}, \tau)$ and its vorticity $\mathbf{w}(\mathbf{x}, \tau) \equiv \nabla \times \mathbf{u}(\mathbf{x}, \tau)$, whose equations of motion follow from Eq. (19):

$$\frac{\partial \theta(\mathbf{x}, \tau)}{\partial \tau} + \mathcal{H}(\tau)\theta(\mathbf{x}, \tau) + \frac{3}{2} \Omega_m(\tau)\mathcal{H}^2(\tau)\delta(\mathbf{x}, \tau) = 0, \quad (20)$$

$$\frac{\partial \mathbf{w}(\mathbf{x}, \tau)}{\partial \tau} + \mathcal{H}(\tau)\mathbf{w}(\mathbf{x}, \tau) = 0. \quad (21)$$

The vorticity evolution readily follows from Eq. (21), $\mathbf{w}(\tau) \propto a^{-1}$, i.e. in the linear regime any initial vorticity decays away due to the expansion of the universe. The density contrast evolution follows by taking the time derivative of Eq. (20) and replacing in Eq. (18):

$$\frac{d^2 D_1(\tau)}{d\tau^2} + \mathcal{H}(\tau) \frac{dD_1(\tau)}{d\tau} = \frac{3}{2} \Omega_m(\tau)\mathcal{H}^2(\tau)D_1(\tau), \quad (22)$$

where we wrote $\delta(\mathbf{x}, \tau) = D_1(\tau)\delta(\mathbf{x}, 0)$, with $D_1(\tau)$ the *linear growth factor*. This equation, together with the Friedmann equations, Eqs. (4) and (5), determines the growth of density perturbations in the linear regime as a function of cosmology. Since it is a second-order differential equation, it has two independent solutions. Let us denote the fastest growing mode $D_1^{(+)}(\tau)$ and the slowest one $D_1^{(-)}(\tau)$. The evolution of the density is then

$$\delta(\mathbf{x}, \tau) = D_1^{(+)}(\tau)A(\mathbf{x}) + D_1^{(-)}(\tau)B(\mathbf{x}), \quad (23)$$

where $A(\mathbf{x})$ and $B(\mathbf{x})$ are two arbitrary functions of position describing the initial density field configuration, whereas the velocity divergence [using Eq. (18)] is given by

$$\theta(\mathbf{x}, \tau) = -\mathcal{H}(\tau)[f(\Omega_m, \Omega_A)A(\mathbf{x}) + g(\Omega_m, \Omega_A)B(\mathbf{x})], \quad (24)$$

$$f(\Omega_m, \Omega_A) \equiv \frac{d \ln D_1^{(+)}}{d \ln a} = \frac{1}{\mathcal{H}} \frac{d \ln D_1^{(+)}}{d \tau}, \quad g(\Omega_m, \Omega_A) = \frac{1}{\mathcal{H}} \frac{d \ln D_1^{(-)}}{d \tau}. \quad (25)$$

The most important cases are

(1) When $\Omega_m = 1$, $\Omega_A = 0$, we have the simple solution

$$D_1^{(+)} = a, \quad D_1^{(-)} = a^{-3/2}, \quad f(1, 0) = 1, \quad (26)$$

thus density fluctuations grow as the scale factor.

(2) When $\Omega_m < 1$, $\Omega_A = 0$ we have ($x \equiv 1/\Omega_m - 1$) [504]

$$D_1^{(+)} = 1 + \frac{3}{x} + 3\sqrt{\frac{1+x}{x^3}} \ln[\sqrt{1+x} - \sqrt{x}], \quad D_1^{(-)} = \sqrt{\frac{1+x}{x^3}}, \quad (27)$$

and the logarithmic derivative can be approximated by [506]

$$f(\Omega_m, 0) \approx \Omega_m^{3/5}. \quad (28)$$

As $\Omega_m \rightarrow 0$ ($x \gg 1$), $D_1^{(+)} \rightarrow 1$ and $D_1^{(-)} \rightarrow x^{-1}$ and perturbations cease to grow.

- (3) In the case where there is only matter and vacuum energy, the linear growth factor admits the integral representation [305] as a function of Ω_m and Ω_Λ :

$$D_1^{(+)} = a^3 H(a) \frac{5\Omega_m}{2} \int_0^a \frac{da}{a^3 H^3(a)}, \quad (29)$$

where $H(a) = \sqrt{\Omega_m^0 a^{-3} + (1 - \Omega_m^0 - \Omega_\Lambda^0) a^{-2} + \Omega_\Lambda^0}$. In general, it is not possible to solve analytically for $D_1^{(+)}$ (unlike $D_1^{(-)}$, see [305]), but can be approximated by [390,114]

$$D_1^{(+)} \approx \left(\frac{5}{2}\right) \frac{a\Omega_m}{\Omega_m^{4/7} - \Omega_\Lambda + (1 + \Omega_m/2)(1 + \Omega_\Lambda/70)}, \quad (30)$$

$$D_1^{(-)} = \frac{\mathcal{H}}{a}, \quad (31)$$

$$f(\Omega_m, \Omega_\Lambda) \approx \frac{1}{[1 - (\Omega_0 + \Omega_\Lambda^0 - 1)a + \Omega_\Lambda^0 a^3]^{0.6}}, \quad (32)$$

where $\Omega_\Lambda^0 \equiv \Omega_\Lambda(a=1)$. When $\Omega_m + \Omega_\Lambda = 1$, we have

$$f(\Omega_m, 1 - \Omega_m) \approx \Omega_m^{5/9}. \quad (33)$$

Due to Eqs. (31) and (4), $g(\Omega_m, \Omega_\Lambda) = \Omega_m - \Omega_\Lambda/2 - 1$ holds for arbitrary Ω_m and Ω_Λ .

2.4. Eulerian non-linear perturbation theory

We will now consider the evolution of density and velocity fields beyond the linear approximation. To do so, we shall first make a *self-consistent* approximation, that is, we will characterize the velocity field by its divergence, and neglect the vorticity degrees of freedom. This can be justified as follows. From Eq. (17) we can write the vorticity equation of motion

$$\frac{\partial \mathbf{w}(\mathbf{x}, \tau)}{\partial \tau} + \mathcal{H}(\tau) \mathbf{w}(\mathbf{x}, \tau) - \nabla \times [\mathbf{u}(\mathbf{x}, \tau) \times \mathbf{w}(\mathbf{x}, \tau)] = \nabla \times \left(\frac{1}{\rho} \nabla \cdot \vec{\sigma} \right), \quad (34)$$

where we have temporarily restored the stress tensor contribution (σ_{ij}) to the conservation of momentum. We see that if $\sigma_{ij} \approx 0$, as in the case of a pressureless perfect fluid, if the primordial vorticity vanishes, it remains zero at all times. On the other hand, if the initial vorticity is non-zero, we saw in the previous section that in the linear regime vorticity decays due to the expansion of the universe; however, it can be amplified non-linearly through the third term in Eq. (34). In what follows, we shall assume that the initial vorticity vanishes, thus Eq. (34) together with the equation of state $\sigma_{ij} \approx 0$ guarantees that vorticity remains zero throughout the evolution. We must note, however, that this assumption is self-consistent only as long as the condition $\sigma_{ij} \approx 0$ remains valid; in particular, multi-streaming and shocks can generate vorticity (see for instance [521]). This is indeed expected to happen at small enough scales. We will come back to this point in order to interpret the breakdown of PT at small scales.

The assumption of PT is that it is possible to expand the density and velocity fields about the linear solutions, effectively treating the variance of the linear fluctuations as a small parameter (and assuming no vorticity in the velocity field). Linear solutions correspond to simple (time dependent) scalings of the *initial* density field; thus we can write

$$\delta(\mathbf{x}, t) = \sum_{n=1}^{\infty} \delta^{(n)}(\mathbf{x}, t), \quad \theta(\mathbf{x}, t) = \sum_{n=1}^{\infty} \theta^{(n)}(\mathbf{x}, t), \quad (35)$$

where $\delta^{(1)}$ and $\theta^{(1)}$ are linear in the initial density field, $\delta^{(2)}$ and $\theta^{(2)}$ are quadratic in the initial density field, etc.

2.4.1. The equations of motion in the Fourier representation

At large scales, when fluctuations are small, linear PT provides an adequate description of cosmological fields. In this regime, different Fourier modes evolve independently conserving the primordial statistics. Therefore, it is natural to Fourier transform Eqs. (9), (16) and (17) and work in Fourier space. Our convention for the Fourier transform of a field $A(\mathbf{x}, \tau)$ is

$$\tilde{A}(\mathbf{k}, \tau) = \int \frac{d^3\mathbf{x}}{(2\pi)^3} \exp(-i\mathbf{k} \cdot \mathbf{x}) A(\mathbf{x}, \tau). \quad (36)$$

When non-linear terms in the perturbation series are taken into account, the equations of motion in Fourier space show the coupling between different Fourier modes characteristic of non-linear theories. Taking the divergence of Eq. (17) and Fourier transforming the resulting equations of motion we get

$$\frac{\partial \tilde{\delta}(\mathbf{k}, \tau)}{\partial \tau} + \tilde{\theta}(\mathbf{k}, \tau) = - \int d^3\mathbf{k}_1 d^3\mathbf{k}_2 \delta_D(\mathbf{k} - \mathbf{k}_{12}) \alpha(\mathbf{k}_1, \mathbf{k}_2) \tilde{\theta}(\mathbf{k}_1, \tau) \tilde{\delta}(\mathbf{k}_2, \tau), \quad (37)$$

$$\begin{aligned} \frac{\partial \tilde{\theta}(\mathbf{k}, \tau)}{\partial \tau} + \mathcal{H}(\tau) \tilde{\theta}(\mathbf{k}, \tau) + \frac{3}{2} \Omega_m \mathcal{H}^2(\tau) \tilde{\delta}(\mathbf{k}, \tau) = & - \int d^3\mathbf{k}_1 d^3\mathbf{k}_2 \delta_D(\mathbf{k} - \mathbf{k}_{12}) \\ & \times \beta(\mathbf{k}_1, \mathbf{k}_2) \tilde{\theta}(\mathbf{k}_1, \tau) \tilde{\theta}(\mathbf{k}_2, \tau) \end{aligned} \quad (38)$$

(δ_D denotes the three-dimensional Dirac delta distribution) where the functions

$$\alpha(\mathbf{k}_1, \mathbf{k}_2) \equiv \frac{\mathbf{k}_{12} \cdot \mathbf{k}_1}{k_1^2}, \quad \beta(\mathbf{k}_1, \mathbf{k}_2) \equiv \frac{k_{12}^2 (\mathbf{k}_1 \cdot \mathbf{k}_2)}{2k_1^2 k_2^2} \quad (39)$$

encode the non-linearity of the evolution (mode coupling) and come from the non-linear terms in continuity equation (16) and Euler equation (17), respectively. From Eqs. (37) and (38) we see that the evolution of $\tilde{\delta}(\mathbf{k}, \tau)$ and $\tilde{\theta}(\mathbf{k}, \tau)$ is determined by the mode coupling of the fields at all pairs of wave vectors \mathbf{k}_1 and \mathbf{k}_2 whose sum is \mathbf{k} , as required by translation invariance in a spatially homogeneous universe.

2.4.2. General solutions in Einstein–de Sitter cosmology

Let us first consider an Einstein–de Sitter universe, for which $\Omega_m = 1$ and $\Omega_A = 0$. In this case the Friedmann equation, Eq. (4), implies $a(\tau) \propto \tau^2$, $\mathcal{H}(\tau) = 2/\tau$, and scaling out an overall factor of \mathcal{H} from the velocity field brings Eqs. (37) and (38) into homogeneous form in τ or, equivalently,

in $a(\tau)$. As a consequence, these equations can formally be solved with the following perturbative expansion [270,334,428]:

$$\tilde{\delta}(\mathbf{k}, \tau) = \sum_{n=1}^{\infty} a^n(\tau) \delta_n(\mathbf{k}), \quad \tilde{\theta}(\mathbf{k}, \tau) = -\mathcal{H}(\tau) \sum_{n=1}^{\infty} a^n(\tau) \theta_n(\mathbf{k}), \quad (40)$$

where only the fastest growing mode is taken into account. Remarkably, it implies that the PT expansions defined in Eq. (35) are actually expansions with respect to the linear density field with time-independent coefficients. At small a , the series are dominated by their first term, and since $\theta_1(\mathbf{k}) = \delta_1(\mathbf{k})$ from the continuity equation, $\delta_1(\mathbf{k})$ completely characterizes the linear fluctuations.

The equations of motion, Eqs. (37) and (38) determine $\delta_n(\mathbf{k})$ and $\theta_n(\mathbf{k})$ in terms of the linear fluctuations to be

$$\delta_n(\mathbf{k}) = \int d^3 \mathbf{q}_1 \cdots \int d^3 \mathbf{q}_n \delta_{\mathbb{D}}(\mathbf{k} - \mathbf{q}_{1\dots n}) F_n(\mathbf{q}_1, \dots, \mathbf{q}_n) \delta_1(\mathbf{q}_1) \cdots \delta_1(\mathbf{q}_n), \quad (41)$$

$$\theta_n(\mathbf{k}) = \int d^3 \mathbf{q}_1 \cdots \int d^3 \mathbf{q}_n \delta_{\mathbb{D}}(\mathbf{k} - \mathbf{q}_{1\dots n}) G_n(\mathbf{q}_1, \dots, \mathbf{q}_n) \delta_1(\mathbf{q}_1) \cdots \delta_1(\mathbf{q}_n), \quad (42)$$

where F_n and G_n are homogeneous functions of the wave vectors $\{\mathbf{q}_1, \dots, \mathbf{q}_n\}$ with degree zero. They are constructed from the fundamental mode coupling functions $\alpha(\mathbf{k}_1, \mathbf{k}_2)$ and $\beta(\mathbf{k}_1, \mathbf{k}_2)$ according to the recursion relations ($n \geq 2$, see [270,334] for a derivation)

$$F_n(\mathbf{q}_1, \dots, \mathbf{q}_n) = \sum_{m=1}^{n-1} \frac{G_m(\mathbf{q}_1, \dots, \mathbf{q}_m)}{(2n+3)(n-1)} [(2n+1)\alpha(\mathbf{k}_1, \mathbf{k}_2) F_{n-m}(\mathbf{q}_{m+1}, \dots, \mathbf{q}_n) + 2\beta(\mathbf{k}_1, \mathbf{k}_2) G_{n-m}(\mathbf{q}_{m+1}, \dots, \mathbf{q}_n)], \quad (43)$$

$$G_n(\mathbf{q}_1, \dots, \mathbf{q}_n) = \sum_{m=1}^{n-1} \frac{G_m(\mathbf{q}_1, \dots, \mathbf{q}_m)}{(2n+3)(n-1)} [3\alpha(\mathbf{k}_1, \mathbf{k}_2) F_{n-m}(\mathbf{q}_{m+1}, \dots, \mathbf{q}_n) + 2n\beta(\mathbf{k}_1, \mathbf{k}_2) G_{n-m}(\mathbf{q}_{m+1}, \dots, \mathbf{q}_n)] \quad (44)$$

(where $\mathbf{k}_1 \equiv \mathbf{q}_1 + \dots + \mathbf{q}_m$, $\mathbf{k}_2 \equiv \mathbf{q}_{m+1} + \dots + \mathbf{q}_n$, $\mathbf{k} \equiv \mathbf{k}_1 + \mathbf{k}_2$ and $F_1 = G_1 \equiv 1$).

For $n = 2$, we have

$$F_2(\mathbf{q}_1, \mathbf{q}_2) = \frac{5}{7} + \frac{1}{2} \frac{\mathbf{q}_1 \cdot \mathbf{q}_2}{q_1 q_2} \left(\frac{q_1}{q_2} + \frac{q_2}{q_1} \right) + \frac{2}{7} \frac{(\mathbf{q}_1 \cdot \mathbf{q}_2)^2}{q_1^2 q_2^2}, \quad (45)$$

$$G_2(\mathbf{q}_1, \mathbf{q}_2) = \frac{3}{7} + \frac{1}{2} \frac{\mathbf{q}_1 \cdot \mathbf{q}_2}{q_1 q_2} \left(\frac{q_1}{q_2} + \frac{q_2}{q_1} \right) + \frac{4}{7} \frac{(\mathbf{q}_1 \cdot \mathbf{q}_2)^2}{q_1^2 q_2^2}. \quad (46)$$

Explicit expressions for the kernels F_3 and F_4 are given in [270]. Note that the symmetrized kernels, $F_n^{(s)}$ (obtained by a summation of F_n with all possible permutations of the variables), have the following properties [270,692]:

- (1) As $\mathbf{k} = \mathbf{q}_1 + \dots + \mathbf{q}_n$ goes to zero, but the individual \mathbf{q}_i do not, $F_n^{(s)} \propto k^2$. This is a consequence of momentum conservation in center-of-mass coordinates.
- (2) As some of the arguments of $F_n^{(s)}$ get large but the total sum $\mathbf{k} = \mathbf{q}_1 + \dots + \mathbf{q}_n$ stays fixed, the kernels vanish in inverse square law. That is, for $p \gg q_i$, we have

$$F_n^{(s)}(\mathbf{q}_1, \dots, \mathbf{q}_{n-2}, \mathbf{p}, -\mathbf{p}) \propto k^2/p^2, \quad (47)$$

and similarly for $G_n^{(s)}$.

- (3) If one of the arguments \mathbf{q}_i of $F_n^{(s)}$ or $G_n^{(s)}$ goes to zero, there is an infrared divergence of the form \mathbf{q}_i/q_i^2 . This comes from the infrared behavior of the mode-coupling functions $\alpha(\mathbf{k}_1, \mathbf{k}_2)$ and $\beta(\mathbf{k}_1, \mathbf{k}_2)$. There are no infrared divergences as partial sums of several wave vectors go to zero.

A simple application of the recursion relations is to derive the corresponding recursion relation for vertices v_n and μ_n which correspond to the spherical average of the PT kernels:

$$v_n \equiv n! \int \frac{d\Omega_1}{4\pi} \dots \frac{d\Omega_n}{4\pi} F_n(\mathbf{k}_1, \dots, \mathbf{k}_n), \quad (48)$$

$$\mu_n \equiv n! \int \frac{d\Omega_1}{4\pi} \dots \frac{d\Omega_n}{4\pi} G_n(\mathbf{k}_1, \dots, \mathbf{k}_n). \quad (49)$$

Since the kernels F_n and G_n depend only on the ratios k_i/k_j , the vertices depend a priori on these quantities as well. Considering Eqs. (43) and (44), one can see that the angle integrations can be done recursively: it is possible to integrate first on the angle between the vectors $\mathbf{k}_1 = \mathbf{q}_1 + \dots + \mathbf{q}_m$ and $\mathbf{k}_2 = \mathbf{q}_{m+1} + \dots + \mathbf{q}_n$, which amounts to replace $\alpha(\mathbf{k}_1, \mathbf{k}_2)$ and $\beta(\mathbf{k}_1, \mathbf{k}_2)$ by their angular averages $\bar{\alpha} = 1$ and $\bar{\beta} = \frac{1}{3}$. As a result we have

$$v_n = \sum_{m=1}^{n-1} \binom{n}{m} \frac{\mu_m}{(2n+3)(n-1)} \left[(2n+1)v_{n-m} + \frac{2}{3} \mu_{n-m} \right], \quad (50)$$

$$\mu_n = \sum_{m=1}^{n-1} \binom{n}{m} \frac{\mu_m}{(2n+3)(n-1)} \left[3v_{n-m} + \frac{2}{3} n\mu_{n-m} \right], \quad (51)$$

and the vertices are thus pure numbers, e.g.

$$v_1 = \mu_1 = 1, \quad v_2 = \frac{34}{21}, \quad v_3 = \frac{682}{189}, \quad \mu_2 = -\frac{26}{21}, \quad \mu_3 = \frac{142}{63}. \quad (52)$$

This recursion relation plays a central role for the derivation of many results in PT [43].

In particular, it can be shown that it is directly related to the spherical collapse dynamics [43,222]. In this case the initial density field is such that it has a spherical symmetry around $\mathbf{x} = 0$. As a consequence, the Fourier transform of the linear density field $\delta_1(\mathbf{k})$ depends only on the norm of \mathbf{k} , and this property remains valid at any stage of the dynamics. Then the central density for such initial conditions, δ_{sc} , can be written (assuming $\Omega_m = 1$ for definiteness) as

$$\delta_{sc}(a) = \sum_n a^n \int d^3\mathbf{q}_1 \dots \int d^3\mathbf{q}_n F_n(\mathbf{q}_1, \dots, \mathbf{q}_n) \delta_1(|\mathbf{q}_1|) \dots \delta_1(|\mathbf{q}_n|). \quad (53)$$

Performing first the integration over the angles of the wave vectors, one recovers

$$\delta_{\text{sc}}(a) = \sum_n \frac{v_n}{n!} a^n \epsilon^n \quad (54)$$

with $\epsilon = \int d^3\mathbf{q} \delta_1(|\mathbf{q}|)$. Similarly, the central velocity divergence for the spherical collapse is expanded in terms of the μ_n parameters. The angular averages of the PT kernels are thus directly related to the spherical collapse dynamics. This result is valid for any cosmological model.

2.4.3. Cosmology dependence of non-linear growth factors

In general, the PT expansion is more complicated because the solutions at each order become non-separable functions of τ and \mathbf{k} [91,93,46,118]. In particular, the growing mode at order n does not scale as $D_1^n(\tau)$ (or $a^n(\tau)$ as in Eq. (40)).

However, using the recursion relations, we can easily find the full dependence on cosmological parameters for the vertices, that is, the dependence that arises in the spherical collapse approximation. The PT kernels can then be constructed order by order in terms of these solutions [46]. In the spherical model, we can write

$$\delta(\tau) = \sum_{n=1}^{\infty} \frac{v_n(\tau)}{n!} [D_1(\tau) \epsilon]^n, \quad (55)$$

$$\theta(\tau) = -\mathcal{H}(\tau) f(\Omega_m, \Omega_\Lambda) \sum_{n=1}^{\infty} \frac{\mu_n(\tau)}{n!} [D_1(\tau) \epsilon]^n. \quad (56)$$

From the Fourier space equations of motion, Eqs. (37) and (38), and taking into account that the spherical averages of α and β can be taken at once, one gets

$$\frac{dv_n}{d \log D_1} + n v_n - \mu_n = \sum_{m=1}^{n-1} \binom{n}{m} v_{n-m} \mu_m, \quad (57)$$

$$\frac{d\mu_n}{d \log D_1} + n \mu_n + \left(\frac{3\Omega_m}{2f^2} - 1 \right) \mu_n - \frac{3\Omega_m}{2f^2} v_n = \frac{1}{3} \sum_{m=1}^{n-1} \binom{n}{m} \mu_{n-m} \mu_m, \quad (58)$$

noting that $d \log D_1 = \mathcal{H} f d\tau$. This hierarchy of differential equations must then be solved numerically at each order. The results for $n=2, 3$ show that indeed the dependence of the vertices on cosmological parameters is a few percent effect at most [46,223].

For a perfect fluid with a equation of state $p = \eta\rho$ we have [259]

$$v_2 = \frac{2(17 + 48\eta + 27\eta^2)}{3(1 + \eta)(7 + 15\eta)} \quad (59)$$

for an Einstein–de Sitter universe. Of course, this reduces to Eq. (52) as $\eta \rightarrow 0$. For the Brans–Dicke cosmology [98], with a coupling ω to gravity

$$v_2 = \frac{34\omega + 56}{21\omega + 36}, \quad (60)$$

which reduces to the standard result $v_2 = \frac{34}{21}$ in the limit $\omega \rightarrow \infty$ (see [259] for details and results for v_4). Even in these extreme cosmologies, the possible variations of v_2 are quite small given the observational constraints on η and ω [259].

2.4.4. Approximate solutions in arbitrary cosmology

This quite remarkable result is asking for an explanation. It is indeed possible to show that a simple approximation to the equations of motion for general Ω_m and Ω_Λ leads to separable solutions to arbitrary order in PT and the same recursion relations as in the Einstein–de Sitter case [560]. All the information on the dependence of the PT solutions on the cosmological parameters Ω_m and Ω_Λ is then encoded in the linear growth factor, $D_1(\tau)$.

In linear PT, the growing-mode solution to the equations of motion (37) and (38) reads

$$\delta(\mathbf{k}, \tau) = D_1(\tau)\delta_1(\mathbf{k}), \quad (61)$$

$$\theta(\mathbf{k}, \tau) = -\mathcal{H}(\tau)f(\Omega_m, \Omega_\Lambda)D_1(\tau)\delta_1(\mathbf{k}), \quad (62)$$

where $D_1(\tau)$ is linear growing mode. As mentioned before, we look for separable solutions of the form (compare with Eq. (40))

$$\delta(\mathbf{k}, \tau) = \sum_{n=1}^{\infty} D_n(\tau)\delta_n(\mathbf{k}), \quad (63)$$

$$\theta(\mathbf{k}, \tau) = -\mathcal{H}(\tau)f(\Omega_m, \Omega_\Lambda) \sum_{n=1}^{\infty} E_n(\tau)\theta_n(\mathbf{k}). \quad (64)$$

From equations of motion (37) and (38) we get, for the n th-order solutions,

$$\begin{aligned} \frac{dD_n}{d\log D_1} \delta_n - E_n\theta_n &= \int d^3k_1 d^3k_2 \delta_D(\mathbf{k} - \mathbf{k}_{12})\alpha(\mathbf{k}, \mathbf{k}_1) \\ &\quad \times \sum_{m=1}^{n-1} D_{n-m}E_m\theta_m(\mathbf{k}_1)\delta_{n-m}(\mathbf{k}_2), \end{aligned} \quad (65)$$

$$\begin{aligned} \frac{dE_n}{d\log D_1} \theta_n + \left(\frac{3\Omega_m}{2f^2} - 1 \right) E_n\theta_n - \frac{3\Omega_m}{2f^2} D_n\delta_n \\ = \int d^3\mathbf{k}_1 d^3\mathbf{k}_2 \delta_D(\mathbf{k} - \mathbf{k}_{12})\beta(\mathbf{k}, \mathbf{k}_1, \mathbf{k}_2) \sum_{m=1}^{n-1} E_{n-m}E_m\theta_m(\mathbf{k}_1)\theta_{n-m}(\mathbf{k}_2). \end{aligned} \quad (66)$$

By simple inspection, we see that if $f(\Omega_m, \Omega_\Lambda) = \Omega_m^{1/2}$, then the system of equations becomes indeed separable, with $D_n = E_n = (D_1)^n$. In fact, the recursion relations then reduce to the standard $\Omega_m = 1$, $\Omega_\Lambda = 0$ case, shown in Eqs. (43) and (44). Then $\Omega_m/f^2 = 1$ leads to separability of the PT solutions to any order, generalizing what has been noted before in the case of second-order PT [432]. From Section 2.3, the approximation $f(\Omega_m, \Omega_\Lambda) \approx \Omega_m^{1/2}$ is actually very good in practice. As a result, for example, as we review in the next section, the exact solution for the $\Omega_\Lambda = 0$ case gives $D_2/(D_1)^2 = 1 + 3/17(\Omega_m^{-2/63} - 1)$, extremely insensitive to Ω_m , even more than what the approximation

$f(\Omega_m, \Omega_\Lambda) = \Omega_m^{3/5} \approx \Omega_m^{1/2}$ would suggest, since for most of the time evolution Ω_m and Ω_Λ are close to their Einstein–de Sitter values.

2.4.5. The density and velocity fields up to third order

The computations of the local density field can be done order by order for any cosmological model. We give here their explicit expression up to third order. The detailed calculations can be found in [46]. Different approaches have been used in the literature to do such calculations [105,118,93]. The direct calculation appears to be the most secure, if not the rapid or most instructive.

The time dependence of the solutions can be written as a function of $D_1(\tau)$, $v_2(\tau)$, $v_3(\tau)$ and an auxiliary function $\lambda_3(\tau)$ which satisfies

$$\frac{d^2(\lambda_3 D_1^3)}{d\tau^2} + \mathcal{H} \frac{d(\lambda_3 D_1^3)}{d\tau} - \frac{3}{2} \mathcal{H}^2 \Omega_m \lambda_3 D_1^3 = \frac{3}{2} \mathcal{H}^2 \Omega_m D_1^3 \quad (67)$$

with $\lambda_3 \sim \frac{9}{10}$ when $\tau \rightarrow 0$. The geometrical dependences can all be expressed in terms of the two functions $\alpha(\mathbf{q}_i, \mathbf{q}_j)$ [see Eq. (39)] and

$$\gamma(\mathbf{q}_i, \mathbf{q}_j) = \frac{1}{2} [\alpha(\mathbf{q}_i, \mathbf{q}_j) + \alpha(\mathbf{q}_j, \mathbf{q}_i)] - \beta(\mathbf{q}_i, \mathbf{q}_j) = 1 - \frac{(\mathbf{q}_i \cdot \mathbf{q}_j)^2}{(q_i^2 q_j^2)}, \quad (68)$$

which for short will be denoted $\alpha_{i,j}$ and $\gamma_{i,j}$ respectively. Then we have

$$F_2(\mathbf{q}_1, \mathbf{q}_2) = \left(\frac{3}{4}v_2 - \frac{3}{2}\right)\gamma_{1,2} + \alpha_{1,2}, \quad (69)$$

$$G_2(\mathbf{q}_1, \mathbf{q}_2) = -f(\Omega_m, \Omega_\Lambda) \left[\left(\frac{3}{4}\mu_2 - \frac{3}{2}\right)\gamma_{1,2} + \alpha_{1,2}\right] \quad (70)$$

for the second-order solutions. Their symmetrized parts can be shown to take the form (see Section 2.7)

$$F_2^{(s)}(\mathbf{q}_1, \mathbf{q}_2) = \frac{1}{2}(1 + \varepsilon) + \frac{1}{2} \frac{\mathbf{q}_1 \cdot \mathbf{q}_2}{q_1 q_2} \left(\frac{q_1}{q_2} + \frac{q_2}{q_1}\right) + \left(\frac{2}{7} - \frac{\varepsilon}{2}\right) \frac{(\mathbf{q}_1 \cdot \mathbf{q}_2)^2}{q_1^2 q_2^2}, \quad (71)$$

$$G_2^{(s)}(\mathbf{q}_1, \mathbf{q}_2) = \varepsilon + \frac{1}{2} \frac{\mathbf{q}_1 \cdot \mathbf{q}_2}{q_1 q_2} \left(\frac{q_1}{q_2} + \frac{q_2}{q_1}\right) + (1 - \varepsilon) \frac{(\mathbf{q}_1 \cdot \mathbf{q}_2)^2}{q_1^2 q_2^2}, \quad (72)$$

where $\varepsilon \approx \left(\frac{3}{7}\right)\Omega_m^{-2/63}$ for $\Omega_m \gtrsim 0.1$ [93]. At third order the kernel reads

$$F_3(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) = \mathcal{R}_1 + v_2 \mathcal{R}_2 + v_3 \mathcal{R}_3 + \lambda_3 \mathcal{R}_4, \quad (73)$$

where, using the simplified notation $\alpha_{ij,k} = \alpha(\mathbf{q}_i + \mathbf{q}_j, \mathbf{q}_k)$, $\alpha_{i,jk} = \alpha(\mathbf{q}_i, \mathbf{q}_j + \mathbf{q}_k)$ and similar definitions for $\gamma_{ij,k}$ and $\gamma_{i,jk}$, we have

$$\mathcal{R}_1 = \left(\frac{1}{2}\alpha_{3,12} + \frac{1}{2}\alpha_{12,3} - \frac{1}{3}\gamma_{3,12}\right)\alpha_{1,2} + \left(-\frac{3}{2}\alpha_{12,3} - \frac{4}{3}\alpha_{3,12} + \frac{5}{2}\gamma_{3,12}\right)\gamma_{1,2}, \quad (74)$$

$$\mathcal{R}_2 = \frac{3}{4}(\alpha_{3,12} + \alpha_{12,3} - 3\gamma_{3,12})\gamma_{1,2}, \quad (75)$$

$$\mathcal{R}_3 = \frac{3}{8}\gamma_{3,12}\gamma_{1,2}, \quad (76)$$

$$\mathcal{R}_4 = \frac{2}{3}\gamma_{3,12}\alpha_{1,2} - \left(\frac{1}{3}\alpha_{3,12} + \frac{1}{2}\gamma_{3,12}\right)\gamma_{1,2}. \quad (77)$$

These results exhibit the explicit time and geometrical dependence of the density field up to third order (a similar expression can be found for G_3 , see [46]). In Section 5 we examine the consequences of these results for the statistical properties of the cosmic fields.

2.4.6. Non-linear growing and decaying modes

Perturbation theory describes the non-linear dynamics as a collection of linear waves, $\delta_1(\mathbf{k})$, interacting through the mode-coupling functions α and β in Eq. (39). Even if the initial conditions are set in the growing mode, after scattering due to non-linear interactions waves do not remain purely in the growing mode. In the standard treatment, described above, the sub-dominant time dependencies that necessarily appear due to this process have been neglected, i.e. only the fastest growing mode (proportional to D_1^n) is taken into account at each order n in PT. Here we discuss how one can generalize the standard results to include the full time dependence of the solutions at every order in PT [561,569]. This is necessary, for example, to properly address the problem of transients in N -body simulations in which initial conditions are set up using the Zel'dovich approximation (see Section 2.5). This is reviewed in Section 5.7. In addition, the approach presented here can be useful to address evolution from non-Gaussian initial conditions.

The equations of motion can be rewritten in a more symmetric form by defining a two-component “vector” $\Psi_a(\mathbf{k}, z)$, where $a = 1, 2$, $z \equiv \ln a$ (we assume $\Omega_m = 1$ for definiteness), and

$$\Psi_a(\mathbf{k}, z) \equiv (\delta(\mathbf{k}, z), -\theta(\mathbf{k}, z)/\mathcal{H}) , \quad (78)$$

which leads to the following equations of motion (we henceforth use the convention that repeated Fourier arguments are integrated over):

$$\partial_z \Psi_a(\mathbf{k}, z) + \Omega_{ab} \Psi_b(\mathbf{k}, z) = \gamma_{abc}(\mathbf{k}, \mathbf{k}_1, \mathbf{k}_2) \Psi_b(\mathbf{k}_1, z) \Psi_c(\mathbf{k}_2, z) , \quad (79)$$

where γ_{abc} is a matrix whose only non-zero elements are $\gamma_{121}(\mathbf{k}, \mathbf{k}_1, \mathbf{k}_2) = \delta_D(\mathbf{k} - \mathbf{k}_1 - \mathbf{k}_2)\alpha(\mathbf{k}, \mathbf{k}_1)$ and $\gamma_{222}(\mathbf{k}, \mathbf{k}_1, \mathbf{k}_2) = \delta_D(\mathbf{k} - \mathbf{k}_1 - \mathbf{k}_2)\beta(\mathbf{k}_1, \mathbf{k}_2)$ and

$$\Omega_{ab} \equiv \begin{bmatrix} 0 & -1 \\ -\frac{3}{2} & \frac{1}{2} \end{bmatrix} . \quad (80)$$

The somewhat complicated expressions for the PT kernels recursion relations in Section 2.4.2 can be easily derived in this formalism. The perturbative solutions read [see Eq. (40)]

$$\Psi_a(\mathbf{k}, z) = \sum_{n=1}^{\infty} e^{nz} \psi_a^{(n)}(\mathbf{k}) , \quad (81)$$

which leads to

$$(n\delta_{ab} + \Omega_{ab})\psi_b^{(n)}(\mathbf{k}) = \gamma_{abc}(\mathbf{k}, \mathbf{k}_1, \mathbf{k}_2) \sum_{m=1}^{n-1} \psi_b^{(n-m)}(\mathbf{k}_1) \psi_c^{(m)}(\mathbf{k}_2) . \quad (82)$$

Now, let $\sigma_{ab}^{-1}(n) \equiv n\delta_{ab} + \Omega_{ab}$, then we have

$$\psi_a^{(n)}(\mathbf{k}) = \sigma_{ab}(n) \gamma_{bcd}(\mathbf{k}, \mathbf{k}_1, \mathbf{k}_2) \sum_{m=1}^{n-1} \psi_c^{(n-m)}(\mathbf{k}_1) \psi_d^{(m)}(\mathbf{k}_2) , \quad (83)$$

where

$$\sigma_{ab}(n) = \frac{1}{(2n+3)(n-1)} \begin{bmatrix} 2n+1 & 2 \\ 3 & 2n \end{bmatrix}. \quad (84)$$

Eq. (83) is the equivalent of the recursion relations in Eqs. (43) and (44) for the n th-order Fourier amplitude solutions $\psi_a^{(n)}(\mathbf{k})$.

To go beyond this, that is, to incorporate the transient behavior before the asymptotics of solutions in Eq. (81) are valid, it turns out to be convenient to write down the equation of motion, Eq. (79), in integral form. Laplace transformation in the variable z leads to

$$\sigma_{ab}^{-1}(\omega) \Psi_b(\mathbf{k}, \omega) = \phi_a(\mathbf{k}) + \gamma_{abc}(\mathbf{k}, \mathbf{k}_1, \mathbf{k}_2) \oint \frac{d\omega_1}{2\pi i} \Psi_b(\mathbf{k}_1, \omega_1) \Psi_c(\mathbf{k}_2, \omega - \omega_1), \quad (85)$$

where $\phi_a(\mathbf{k})$ denote the initial conditions, that is $\Psi_a(\mathbf{k}, z=0) \equiv \phi_a(\mathbf{k})$. Multiplying by the matrix σ_{ab} , and performing the inversion of the Laplace transform gives [569]

$$\Psi_a(\mathbf{k}, z) = g_{ab}(z) \phi_b(\mathbf{k}) + \int_0^z ds g_{ab}(z-s) \gamma_{bcd}(\mathbf{k}, \mathbf{k}_1, \mathbf{k}_2) \Psi_c(\mathbf{k}_1, s) \Psi_d(\mathbf{k}_2, s), \quad (86)$$

where the *linear propagator* $g_{ab}(z)$ is defined as ($c > 1$ to pick out the standard retarded propagator [561])

$$g_{ab}(z) = \oint_{c-i\infty}^{c+i\infty} \frac{d\omega}{2\pi i} \sigma_{ab}(\omega) e^{\omega z} = \frac{e^z}{5} \begin{bmatrix} 3 & 2 \\ 3 & 2 \end{bmatrix} - \frac{e^{-3z/2}}{5} \begin{bmatrix} -2 & 2 \\ 3 & -3 \end{bmatrix} \quad (87)$$

for $z \geq 0$, whereas $g_{ab}(z) = 0$ for $z < 0$ due to causality, $g_{ab}(z) \rightarrow \delta_{ab}$ as $z \rightarrow 0^+$. The first term in Eq. (87) represents the propagation of linear growing-mode solutions, where the second corresponds to the decaying modes propagation. Eq. (86) can be thought as an equation for $\Psi_a(\mathbf{k}, z)$ in the presence of an “external source” $\phi_b(\mathbf{k})$ with prescribed statistics given by the initial conditions.⁴ It contains the full time dependence of non-linear solutions, as will be discussed in detail in Section 5.7. To recover the standard (asymptotic) time dependence one must take the initial conditions to be set in the growing mode, $\phi_b \propto (1, 1)$, which vanishes upon contraction with the second term in Eq. (87), and reduces to the familiar linear scaling $\phi_a(z) = e^z \phi_a(0) = a(\tau) \phi_a(0)$; and, in addition, set the lower limit of integration in Eq. (86) to $s = -\infty$, to place initial conditions “infinitely far away” in the past.

2.5. Lagrangian dynamics

So far we have dealt with density and velocity fields and their equations of motion. However, it is possible to develop non-linear PT in a different framework, the so-called Lagrangian scheme, by following the trajectories of particles or fluid elements [705,102,465], rather than studying the

⁴ This is essentially a field-theoretic description of gravitational instability, non-linear corrections can be thought as loop corrections to the propagator and the vertex given by the γ_{abc} matrix, see [569] for details.

dynamics of density and velocity fields.⁵ In Lagrangian PT,⁶ the object of interest is the displacement field $\Psi(\mathbf{q})$ which maps the initial particle positions \mathbf{q} into the final Eulerian particle positions \mathbf{x} ,

$$\mathbf{x}(\tau) = \mathbf{q} + \Psi(\mathbf{q}, \tau) . \quad (88)$$

The equation of motion for particle trajectories $\mathbf{x}(\tau)$ is then

$$\frac{d^2\mathbf{x}}{d\tau^2} + \mathcal{H}(\tau) \frac{d\mathbf{x}}{d\tau} = -\nabla\Phi , \quad (89)$$

where Φ denotes the gravitational potential and ∇ the gradient operator in Eulerian coordinates \mathbf{x} . Taking the divergence of this equation we obtain

$$J(\mathbf{q}, \tau) \nabla \cdot \left[\frac{d^2\Psi}{d\tau^2} + \mathcal{H}(\tau) \frac{d\Psi}{d\tau} \right] = \frac{3}{2} \Omega_m \mathcal{H}^2 (J - 1) , \quad (90)$$

where we have used Poisson equation together with the fact that the density field obeys $\bar{\rho} (1 + \delta(\mathbf{x})) d^3x = \bar{\rho} d^3q$, thus

$$1 + \delta(\mathbf{x}) = \frac{1}{\text{Det}(\delta_{ij} + \Psi_{i,j})} \equiv \frac{1}{J(\mathbf{q}, \tau)} , \quad (91)$$

where $\Psi_{i,j} \equiv \partial\Psi_i/\partial q_j$ and $J(\mathbf{q}, \tau)$ is the Jacobian of the transformation between Eulerian and Lagrangian space. Note that when there is shell crossing, i.e. fluid elements with different initial positions \mathbf{q} end up at the same Eulerian position \mathbf{x} through the mapping in Eq. (88), the Jacobian vanishes and the density field becomes singular. At these points the description of dynamics in terms of a mapping does not hold anymore.

Eq. (90) can be fully rewritten in terms of Lagrangian coordinates by using that $\nabla_i = (\delta_{ij} + \Psi_{i,j})^{-1} \nabla_{q_j}$, where $\nabla_q \equiv \partial/\partial\mathbf{q}$ denotes the gradient operator in Lagrangian coordinates. The resulting non-linear equation for $\Psi(\mathbf{q})$ is then solved perturbatively, expanding about its linear solution.

2.6. Linear solutions and the Zel'dovich approximation

The linear solution of Eq. (90)

$$\nabla_q \cdot \Psi^{(1)} = -D_1(\tau) \delta(\mathbf{q}) , \quad (92)$$

where $\delta(\mathbf{q})$ denotes the density field imposed by the initial conditions and $D_1(\tau)$ is the linear growth factor, which obeys Eq. (22). We implicitly assume that vorticity vanishes, then Eq. (92) completely determines the displacement field to linear order. Linear Lagrangian solutions have the property that they become exact for local one-dimensional motion, i.e. when the two eigenvalues of the velocity gradient along the trajectory vanish [102]. Note that the evolution of fluid elements at this order is *local*, i.e. it does not depend on the behavior of the rest of fluid elements.

⁵ It is also possible to study Lagrangian dynamics of density and velocity fields following the fluid elements, by using the convective derivative $D/Dt \equiv \partial/\partial t + \mathbf{u} \cdot \nabla$ in the equations of motion, Eqs. (16) and (17). We will not discuss this possibility here, but e.g. see [62,327].

⁶ For reviews of Lagrangian PT, see e.g. [107,94].

The *Zel'dovich approximation* (hereafter ZA) [705] consists in using the linear displacement field as an approximate solution for the dynamical equations.⁷ It follows from Eq. (91) that the local density field reads

$$1 + \delta(\mathbf{x}, \tau) = \frac{1}{[1 - \lambda_1 D_1(\tau)][1 - \lambda_2 D_1(\tau)][1 - \lambda_3 D_1(\tau)]}, \quad (93)$$

where λ_i are the local eigenvalues of the tidal tensor $\Psi_{i,j}$. From this expression we can see that depending on the relative magnitude of these eigenvalues, the ZA leads to planar collapse (one positive eigenvalue larger than the rest), filamentary collapse (two positive eigenvalues larger than the third) or spherical collapse (all eigenvalues positive and equal). If all eigenvalues are negative, then the evolution corresponds to an underdense region, eventually reaching $\delta = -1$. For Gaussian initial conditions, it is possible to work out the probability distribution for the eigenvalues [190], which leads through the non-linear transformation in Eq. (93) to a characterization of the one-point statistical properties of the density field. These results will be discussed in Section 5.8.3.

2.7. Lagrangian perturbation theory

Unlike in Eulerian PT, there is no known recursive solution for the expression of the order-by-order cosmic fields in Lagrangian PT, even for the Einstein–de Sitter case. One reason for that is that beyond second order, even though one can assume an irrotational flow in Eulerian space, this does not imply that the displacement field is irrotational [105]. It has been stressed that already second-order Lagrangian PT for the displacement field (hereafter 2LPT), does provide a remarkable improvement over the ZA in describing the global properties of density and velocity fields [106,455,93] and in most practical cases the improvement brought by third-order Lagrangian PT is marginal [106,455].

One way to understand this situation is to recall that the Lagrangian picture is intrinsically non-linear in the density field (e.g. see Eq. (91)), and a small perturbation in Lagrangian fluid element paths carries a considerable amount of non-linear information about the corresponding Eulerian density and velocity fields. In particular, as we shall see below, a truncation of Lagrangian PT at a fixed order, yields non-zero Eulerian PT kernels at every order. However, as we shall review in the next few sections, this is not always an advantage, particularly when dealing with initial conditions with enough small-scale power where shell crossing is significant. In these cases, Lagrangian PT generally breaks down at scales larger than Eulerian PT.

The reason for the remarkable improvement of 2LPT over ZA is in fact not surprising. The solution of Eq. (90) to second order describes the correction to the ZA displacement due to gravitational tidal effects, that is, it takes into account the fact that gravitational instability is non-local. It reads

$$\nabla_q \cdot \Psi^{(2)} = \frac{1}{2} D_2(\tau) \sum_{i \neq j} (\Psi_{i,i}^{(1)} \Psi_{j,j}^{(1)} - \Psi_{i,j}^{(1)} \Psi_{j,i}^{(1)}), \quad (94)$$

where $D_2(\tau)$ denotes the second-order growth factor, which for $0.1 \leq \Omega_m \leq 3$ ($\Omega_\Lambda = 0$) obeys

$$D_2(\tau) \approx -\frac{3}{7} D_1^2(\tau), \quad (95)$$

⁷ Rigorously, the ZA results from using the linear displacement field with the constraint that at large scales one recovers linear Eulerian PT [103].

or more precisely

$$D_2(\tau) \approx -\frac{3}{7} D_1^2(\tau) \Omega_m^{-2/63}, \quad (96)$$

to better than 7% and 0.5% respectively [91], whereas for flat models with non-zero cosmological constant Ω_Λ we have for $0.01 \leq \Omega_m \leq 1$

$$D_2(\tau) \approx -\frac{3}{7} D_1^2(\tau) \Omega_m^{-1/143}, \quad (97)$$

to better than 0.6% [93]. Since Lagrangian solutions up to second order are curl free,⁸ it is convenient to define Lagrangian potentials $\phi^{(1)}$ and $\phi^{(2)}$ so that in 2LPT

$$\mathbf{x}(\mathbf{q}) = \mathbf{q} - D_1 \nabla_q \phi^{(1)} + D_2 \nabla_q \phi^{(2)}, \quad (98)$$

and the velocity field then reads

$$\mathbf{u} = -D_1 f_1 \mathcal{H} \nabla_q \phi^{(1)} + D_2 f_2 \mathcal{H} \nabla_q \phi^{(2)}, \quad (99)$$

where the logarithmic derivatives of the growth factors $f_i \equiv (\mathrm{d} \ln D_i) / (\mathrm{d} \ln a)$ can be approximated for open models with $0.1 \leq \Omega_m \leq 1$ by

$$f_1 \approx \Omega_m^{3/5}, \quad f_2 \approx 2\Omega_m^{4/7}, \quad (100)$$

to better than 2% [506] and 5% [93], respectively. For flat models with non-zero cosmological constant Ω_Λ we have for $0.01 \leq \Omega_m \leq 1$

$$f_1 \approx \Omega_m^{5/9}, \quad f_2 \approx 2\Omega_m^{6/11}, \quad (101)$$

to better than 10% and 12%, respectively [93]. The accuracy of these two fits improves significantly for $\Omega_m \geq 0.1$, in the relevant range according to present observations. Summarizing, the time-independent potentials in Eqs. (98) and (99) obey the following Poisson equations [106]:

$$\nabla_q^2 \phi^{(1)}(\mathbf{q}) = \delta(\mathbf{q}), \quad (102)$$

$$\nabla_q^2 \phi^{(2)}(\mathbf{q}) = \sum_{i>j} [\phi_{,ii}^{(1)}(\mathbf{q}) \phi_{,jj}^{(1)}(\mathbf{q}) - (\phi_{,ij}^{(1)}(\mathbf{q}))^2]. \quad (103)$$

It is possible to improve on 2LPT by going to third order in the displacement field (3LPT), however it becomes more costly due to the need of solving three additional Poisson equations [105,117]. Third-order results give a better behavior in underdense regions [93] and lead to additional substructure in high-density regions [108]. Detailed comparison of Lagrangian PT at different orders against numerical simulations is given in [93,367].

2.8. Non-linear approximations

When density fluctuations become strongly non-linear, PT breaks down and one has to resort to numerical simulations to study their evolution. However, numerical simulations provide limited physical insight into the physics of gravitational clustering. On the other hand, many non-linear approximations to the equations of motion have been suggested in the literature which allow calculations to be extrapolated to the non-linear regime. However, as we shall see, it seems fair to say

⁸ This is assuming that initial conditions are in the growing mode, for a more general treatment see [104].

that these approximations have mostly been useful to gain understanding about different aspects of gravitational clustering while quantitatively none of them seem to be accurate enough for practical use. Rigorous PT has provided a very useful way to benchmark these different approximations in the weakly non-linear regime.

In general, most non-linear approximations can be considered as different assumptions (valid in linear PT) that replace Poisson's equation [470]. These modified dynamics, are often local, in the sense described above for the ZA, in order to provide a simpler way of calculating the evolution of perturbations than the full non-local dynamics.

Probably, the best known of non-linear approximations is the ZA, which in Eulerian space is equivalent to replacing the Poisson equation by the following ansatz [470,327]:

$$\mathbf{u}(\mathbf{x}, \tau) = -\frac{2f}{3\Omega_m \mathcal{H}(\tau)} \nabla \Phi(\mathbf{x}, \tau), \quad (104)$$

which is the relation between velocity and gravitational potential valid in linear PT. Conservation of momentum (assuming for definiteness $\Omega_m = 1$) then becomes [see Eq. (17)]

$$\frac{\partial \mathbf{u}(\mathbf{x}, \tau)}{\partial \tau} - \frac{\mathcal{H}(\tau)}{2} \mathbf{u}(\mathbf{x}, \tau) + \mathbf{u}(\mathbf{x}, \tau) \cdot \nabla \mathbf{u}(\mathbf{x}, \tau) = 0. \quad (105)$$

It is straightforward to find the PT recursion relations using these equations of motion [557]; the result for the density field kernel is particularly simple [274],

$$F_n^{(s)}(\mathbf{q}_1, \dots, \mathbf{q}_n) = \frac{1}{n!} \frac{\mathbf{k} \cdot \mathbf{q}_1}{q_1^2} \dots \frac{\mathbf{k} \cdot \mathbf{q}_n}{q_n^2}, \quad (106)$$

where $\mathbf{k} \equiv \mathbf{q}_1 + \dots + \mathbf{q}_n$. As we mentioned before, the ZA is a local approximation and becomes the exact dynamics in one-dimensional collapse. It is also possible to formulate local approximations that besides being exact for planar collapse like the ZA, are also exact for spherical [62] and even cylindrical collapse [327]. However, their implementation for the calculation of statistical properties of density and velocity fields is not straightforward.

A significant shortcoming of the ZA is the fact that after shell crossing (“pancake formation”), matter continues to flow throughout the pancake without ever turning around, washing out structures at small scales. This can be fixed phenomenologically by adding some small effective viscosity to Eq. (105), which then becomes the Burgers' equation⁹

$$\frac{\partial \mathbf{u}(\mathbf{x}, \tau)}{\partial \tau} - \frac{\mathcal{H}(\tau)}{2} \mathbf{u}(\mathbf{x}, \tau) + \mathbf{u}(\mathbf{x}, \tau) \cdot \nabla \mathbf{u}(\mathbf{x}, \tau) = \nu \nabla^2 \mathbf{u}(\mathbf{x}, \tau). \quad (107)$$

This is the so-called *adhesion approximation* [278]. This equation has the nice property that for a potential flow it can be reduced to a linear diffusion equation, and therefore solved exactly. Given the initial conditions, this can be used to predict the location of pancakes and clusters, giving good agreement when compared to numerical simulations [381]. More detailed comparisons with numerical simulations for density field statistics show an improvement over the ZA at small scales [683]; however, at weakly non-linear scales the adhesion approximation is essentially equal to the ZA.

⁹ An attempt to see how this equation might arise from the physics of multi-streaming has been given in [109].

The *linear potential approximation* [97,13] assumes that the gravitational potential remains the same as in the linear regime; therefore,

$$\nabla^2 \Phi(\mathbf{x}, \tau) = \frac{3}{2} \Omega_m \mathcal{H}^2(\tau) \delta_1(\mathbf{x}, \tau), \quad (108)$$

where $\delta_1(\mathbf{x}, \tau) = D_1^{(+)}(\tau) \delta_1(\mathbf{x})$ is the linearly extrapolated density field. The idea behind this approximation is that since $\Phi \propto \delta/k^2$, the gravitational potential is dominated by long-wavelength modes more than the density field, and therefore it ought to obey linear PT to a better approximation.

In the *frozen flow approximation* [433], the velocity field is instead assumed to remain linear,

$$\theta(\mathbf{x}, \tau) = -\mathcal{H}(\tau) f(\Omega_m, \Omega_A) \delta_1(\mathbf{x}, \tau), \quad (109)$$

i.e. the velocity field kernels $G_n^{(s)} \equiv 0$ ($n > 1$). In the next sections we will briefly review how these different approximations compare in the weakly non-linear regime [470,471,47,557], see e.g. Table 4.

2.9. Numerical simulations

2.9.1. Introduction

Cosmological dark matter simulations have become a central tool in predicting the evolution of structure in the universe well into the non-linear regime. Current state-of-the-art numerical simulations can follow the dynamics of about 10^9 particles (see e.g. [163]), which although impressive, is still tens of orders of magnitude smaller than the number of dark matter particles expected in a cosmological volume, as mentioned in the introduction.

However, this is not an insurmountable limitation. As we discussed in Section 2.1, in the limit that the number of particles $N \gg 1$, collisionless dark matter obeys the Vlasov equation for the distribution function in phase space, Eq. (12). The task of numerical simulations is to sample this distribution by partitioning phase space into N elementary volumes, “particles” with positions, velocities and (possibly different) masses m_i , $i = 1, \dots, N$, and following the evolution of these test particles due to the action of gravity and the expansion of the universe (technically, these particles obey the equations of the characteristics of the Vlasov equation). The number of particles N fixes the *mass resolution* of the numerical simulation.

Each particle i can be thought of as carrying a “smooth” density profile, which can be viewed as a “cloud” of typical size ϵ_i . The parameter ϵ_i is called the softening length (associated to particle i). In general, $\epsilon_i \propto m_i^{1/3}$. This softening is introduced to suppress interactions between nearby particles in order to reduce N -body relaxation, which is an artifact of the discrete description of the distribution function. It fixes the *spatial resolution* of the simulation. In general, it is chosen to be a small fraction of the (local or global) mean inter-particle separation, but this can vary significantly depending on the type of code used.

In this section, we briefly discuss methods used to solve numerically the Vlasov equation. A complete discussion of N -body methods is beyond the scope of this work, we shall only describe the most common methods closely following [155]; for a comprehensive review see e.g. [63].

The basic steps in an N -body simulation can be summarized as follows:

- (i) implementation of initial conditions ([379,199], see e.g. [64] and references therein for recent developments);

- (ii) calculation of the force by solving the Poisson equation;
- (iii) update of positions and velocities of particles;
- (iv) diagnostics, e.g. tests of energy conservation;
- (v) go back to (ii) until simulation is completed.

In general, step (iii) is performed with time integrators accurate to second order, preferably symplectic (i.e. that preserve phase-space volume). The Leapfrog integrator (e.g. [314]), where velocities and positions are shifted from each other by half a time step, is probably the most common one. The Predictor–Corrector scheme is also popular since it allows easy implementation of individual, varying time step per particle (e.g. [601]). Low-order integrators are used mostly to minimize the storage of variables for a large number of particles whose orbits must be integrated and to reduce the cost of the force calculation. Because of the chaotic nature of gravitational dynamics, it is not feasible to follow very accurately individual particle orbits but only to properly recover the properties of bound objects in a statistical sense.

All the methods that we describe in what follows mainly differ in the calculation of the force applied to each particle or, in other words, in how the Poisson equation is solved.

2.9.2. Direct summation

Also known as particle–particle (PP) method (e.g. [1]), it consists in evaluating the force on each particle by summing directly the influence exerted on it by all neighbors. This method is robust but very CPU consuming: scaling as $\mathcal{O}(N^2)$, it allows a small number of particles, typically $N \sim 10^3$ – 10^5 . It was revived recently by the development of special hardware dedicated to the computation of the Newtonian force (e.g. [427]), mostly used for stellar dynamics calculations (but see e.g. [243] for a cosmological application).

2.9.3. The tree algorithm

The tree code is the most natural improvement of the PP method. It uses the fact that the influence of remote structures on each particle can be computed by performing a multipole expansion on clusters containing many particles. With appropriate selection of the clusters, the expansion can be truncated at low order. Therefore, the list of interactions on each particle is much shorter than in the PP method, of order $\sim \log N$, resulting in a $\mathcal{O}(N \log N)$ code. The practical implementation of the tree code consists in decomposing hierarchically the system on a tree structure, which can be, for example, a mutually nearest neighbor binary tree (e.g. [8]), or a space-balanced Oct tree in which each branch is a cubical portion of space (e.g. [22,309,89]). Then a criterion is applied to see whether or not a given cluster of particles has to be broken into smaller pieces (or equivalently, if it is necessary to walk down the tree).

Various schemes exist (e.g. [545]), the simplest one for the Oct tree [22] consisting in subdividing the cells until the condition $s/r \leq \theta$ is fulfilled, where s is the size of the cell, r is the distance of the cell center of mass to the particle and θ is a tunable parameter of order unity.

The tree data structure has many advantages: (i) the CPU spent per time step does not depend significantly on the degree of clustering of the system; (ii) implementation of individual time steps per particle is fairly easy and this can speed up the simulation significantly; (iii) the use of individual masses per particle allows “zooming” in a particular region, for example a cluster, a galaxy halo or a void: the location of interest is sampled accurately with high-resolution particles (with small mass),

while tidal effects are modeled by low-resolution particles of mass increasing with distance from the high-resolution region; (iv) implementation on parallel architectures with distributed memory is relatively straightforward (e.g. [546,193,601]). However, tree codes are rather demanding in memory (25–35 words per particles, e.g. [163]) and accurate handling of periodic boundaries (e.g. [310]) is costly.

Typically, simulations using the tree code can involve up to $\sim 10^7$ – 10^8 particles if done on parallel supercomputers. They have high spatial resolution, of order $\epsilon \sim \lambda/(10\text{--}20)$, where λ is the mean inter-particle distance.

2.9.4. The PM algorithm

In the particle–mesh (PM) method (e.g. see [314,191,454,86]), the mass of each particle is interpolated on a fixed grid of size N_g (with N_g^3 sites) to compute the density. The Poisson equation is solved on the grid, generally by using a fast Fourier transform, then forces are interpolated back on the particles. Implementing a PM code is thus rather simple, even on parallel architectures. Scaling as $\mathcal{O}(N, N_g^3 \log N_g)$, PM simulations have generally the advantage being low CPU consumers and require reasonable amount of memory. Thus, a large number of particles can be used, $N \sim 10^7$ – 10^9 , and typically $N_g = N^{1/3}$ or $2N^{1/3}$. The main advantage and weakness of the PM approach is its low spatial resolution. Indeed, the softening parameter is fixed by the size of the grid, $\epsilon \sim L/N_g$, where L is the size of the box: large softening length reduces the effects of N -body relaxation and allows good phase-space sampling, but considerably narrows the available dynamic scale range. To achieve a spatial resolution comparable to that of a tree code while keeping the advantage of the PM code, very large values of N_g and N would be needed, implying a tremendous cost both in memory and in CPU.

2.9.5. Hybrid methods

To increase spatial resolution of the PM approach, several improvements have been suggested.

The most popular one is the P³M code (PP+PM) where the PM force is supplemented with a short-range contribution obtained by direct summation of individual interactions between nearby particles (e.g. [314,199]). Implementation of this code on a parallel supercomputer (T3E) produced a very large cosmological simulation with 10^9 particles in a “Hubble” volume of size $L=2000h^{-1}$ Mpc [420]. The main caveat of the P³M approach is that as the system evolves to a more clustered state, the time spent in calculation of PP interactions becomes increasingly significant. To reduce the slowing-down due to PP interactions, it was proposed to use a hierarchy of adaptive meshes in regions of high particle density [162], giving birth to a very efficient N -body code, the adaptive P³M (AP³M).

Instead of direct PP summations to correct the PM force for short-range interactions, it is possible to use a tree algorithm in high-density regions [695] or in all PM cells [12] similarly as in the P³M code. Both these methods are potentially faster than their P³M competitor.

In the same spirit as in AP³M, but without the PP part, another alternative is to use adaptive mesh refinement (AMR): the PM mesh is increased locally when required with a hierarchy of nested rectangular sub-grids (e.g. [675,6,341,264]). The forces can be computed at each level of the hierarchy by a Fourier transform with appropriate boundary conditions. In fact, the sub-grids need not be rectangular if one uses Oct tree structures, which is theoretically even more efficient. In this

adaptive refinement tree (ART) method [386], the Poisson equation is solved by relaxation methods (e.g. [314,532]).

Finally, it is worth mentioning a Lagrangian approach, which consists in using a mesh with fixed size like in the PM code, but moving with the flow so that resolution increases in high-density regions and decreases elsewhere [269,516]. However, this potentially powerful method presents some difficulties, e.g. mesh distortions may induce severe force anisotropies.

3. Random cosmic fields and their statistical description

In this section we succinctly recall current ideas about the physical origin of stochasticity in cosmic fields in different cosmological scenarios. We then present the statistical tools that are commonly used to describe random cosmic fields such as power spectra, probability distribution functions, moments and cumulants, and give some mathematical properties of interest.

3.1. The need for a statistical approach

As we shall review in detail in the following sections, the current explanation of the large-scale structure of the universe is that the present distribution of matter on cosmological scales results from the growth of primordial, small, seed fluctuations on an otherwise homogeneous universe amplified by gravitational instability. Tests of cosmological theories which characterize these primordial seeds are not deterministic in nature but rather statistical, for the following reasons. First, we do not have direct observational access to primordial fluctuations (which would provide definite initial conditions for the deterministic evolution equations). In addition, the time scale for cosmological evolution is so much longer than that over which we can make observations, that it is not possible to follow the evolution of single systems. In other words, what we observe through our past light cone is different objects at different times of their evolution; therefore, testing the evolution of structure must be done statistically.

The observable universe is thus modeled as a stochastic realization of a statistical ensemble of possibilities. The goal is to make statistical predictions, which in turn depend on the statistical properties of the primordial perturbations leading to the formation of large-scale structures. Among the two classes of models that have emerged to explain the large-scale structure of the universe, the physical origin of stochasticity can be quite different and thus give rise to very different predictions.

The most widely considered models, based on the inflationary paradigm [279], generically give birth to adiabatic¹⁰ Gaussian initial fluctuations, at least in the simplest single-field models [602,304, 280,20]. In this case the origin of stochasticity lies on quantum fluctuations generated in the early universe; we will consider this case in more detail below. However, one should keep in mind that inflation is not necessarily the only mechanism that leads to Gaussian, or almost Gaussian, initial conditions. For instance, topological defects based on the non-linear σ -model in the large N -limit

¹⁰ As opposed to *isocurvature* fluctuations which is a set of individual perturbations such that the total fluctuation amplitude vanishes. In the adiabatic case, the total amplitude does not vanish and this leads to perturbations in the spatial curvature.

would also give Gaussian initial conditions [655,333]. And in general the central limit theorem ensures that such initial conditions are likely to happen in very broad classes of models.

The second class of models that have been developed for structure formation are based on topological defects, of which cosmic strings have been studied in most detail. In this case, the origin of stochasticity lies on thermal fluctuations of a field that undergoes a phase transition as the universe cools, and is likely to obey non-Gaussian properties. Note, however, that these two classes of models do not necessarily exclude each other. For instance, formation of cosmic strings is encountered in specific models of inflation [65,66,352]. There are also models inspired by duality properties of superstring theories, in which an inflationary phase can be encountered but structure formation is caused by the quantum fluctuations of the axion field¹¹ [668,159,111] rather than the inflation field. With such a mechanism the initial metric fluctuations will not obey Gaussian statistics.

3.1.1. Physical origin of fluctuations from inflation

In models of inflation the stochastic properties of the fields originate from quantum fluctuations of a scalar field, the inflation. It is beyond the scope of this review to describe inflationary models in any detail. We, instead, refer the reader to recent reviews for a complete discussion [399,400,415]. However, it is worth recalling that in such models (at least for the simplest single-field models within the slow-roll approximation) all fluctuations originate from scalar adiabatic perturbations. During the inflationary phase the energy density of the universe is dominated by the density stored in the inflation field. This field has quantum fluctuations that can be decomposed in Fourier modes using the creation and annihilation operators $a_{\mathbf{k}}^\dagger$ and $a_{\mathbf{k}}$ for a wave mode \mathbf{k} :

$$\delta\varphi = \int d^3\mathbf{k} [a_{\mathbf{k}}\psi_{\mathbf{k}}(t)\exp(i\mathbf{k}\cdot\mathbf{x}) + a_{\mathbf{k}}^\dagger\psi_{\mathbf{k}}^*(t)\exp(-i\mathbf{k}\cdot\mathbf{x})] . \quad (110)$$

The operators obey the standard commutation relation

$$[a_{\mathbf{k}}, a_{-\mathbf{k}'}^\dagger] = \delta_{\mathbf{D}}(\mathbf{k} + \mathbf{k}') , \quad (111)$$

and the mode functions $\psi_{\mathbf{k}}(t)$ are obtained from the Klein–Gordon equation for φ in an expanding universe. We give here its expression for a de-Sitter metric (i.e. when the spatial sections are flat and H is constant),

$$\psi_{\mathbf{k}}(t) = \frac{H}{(2k)^{1/2}k} \left(i + \frac{k}{aH} \right) \exp \left[\frac{i\mathbf{k}}{aH} \right] , \quad (112)$$

where a and H are, respectively, the expansion factor and the Hubble constant that are determined by the overall content of the universe through the Friedmann equations, Eqs. (4) and (5).

When the modes exit the Hubble radius, $k/(aH) \ll 1$, one can see from Eq. (112) that the dominant mode reads

$$\varphi_{\mathbf{k}} \approx \frac{iH}{\sqrt{2}k^{3/2}}(a_{\mathbf{k}} + a_{-\mathbf{k}}^\dagger), \quad \delta\varphi = \int d^3\mathbf{k} \varphi_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{x}} . \quad (113)$$

Thus these modes are all proportional to $a_{\mathbf{k}} + a_{-\mathbf{k}}^\dagger$. One important consequence of this is that the quantum nature of the fluctuations has disappeared [281,375,376]: any combinations of $\varphi_{\mathbf{k}}$ commute

¹¹ However, this generally leads to isocurvature fluctuations rather than adiabatic.

with each other. The field φ can then be seen as a classic stochastic field where *ensemble averages identify with vacuum expectation values*,

$$\langle \dots \rangle \equiv \langle 0 | \dots | 0 \rangle . \quad (114)$$

After the inflationary phase the modes re-enter the Hubble radius. They leave imprints of their energy fluctuations in the gravitational potential, the statistical properties of which can therefore be deduced from Eqs. (111) and (113). All subsequent stochasticity that appears in the cosmic fields can thus be expressed in terms of the random variable $\varphi_{\mathbf{k}}$.

3.1.2. Physical origin of fluctuations from topological defects

In models of structure formation with topological defects, stochasticity originates from thermal fluctuations. One important difficulty in this case is that topological defects generally behave as active seeds, and except in some special cases (see for instance [194]), the dynamical evolution of these seeds is non-linear and non-local, hence requiring heavy numerical calculation for their description. This is in particular true for cosmic strings that form a network whose evolution is extremely complex (see for instance [90]). Therefore, in this case it is not possible to write down in general how the stochasticity in cosmic fields relates to more fundamental processes. See [674] for a review of the physics of topological defects. Current observations of multiple acoustic peaks in the power spectrum of microwave background anisotropies severely constrain significant contributions to perturbations from active seeds [476,282,397].

3.2. Correlation functions and power spectra

From now on, we consider a cosmic scalar field whose statistical properties we want to describe. This field can either be the cosmic density field, $\delta(\mathbf{x})$, the cosmic gravitational potential, the velocity divergence field, or any other field of interest.

3.2.1. Statistical homogeneity and isotropy

A random field is called *statistically homogeneous*¹² if all the joint multipoint *probability distribution functions* $p(\delta_1, \delta_2, \dots)$ or its *moments*, ensemble averages of local density products, remain the same under translation of the coordinates $\mathbf{x}_1, \mathbf{x}_2, \dots$ in space (here $\delta_i \equiv \delta(\mathbf{x}_i)$). Thus the probabilities depend only on the relative positions. A stochastic field is called *statistically isotropic* if $p(\delta_1, \delta_2, \dots)$ is invariant under spatial rotations. We will assume that cosmic fields are statistically homogeneous and isotropic, as predicted by most cosmological theories. The validity of this assumption can and should be tested against the observational data. Examples of primordial fields which do not obey statistical homogeneity and isotropy are fluctuations in compact hyperbolic spaces (see e.g. [82]). Furthermore, redshift distortions in galaxy redshift surveys introduce significant deviations from statistical isotropy and homogeneity in the redshift-space density field, as will be reviewed in Section 7.

¹² This is in contrast with a *homogeneous* field, which takes the same value everywhere in space.

3.2.2. Two-point correlation function and power spectrum

The two-point correlation function is defined as the joint ensemble average of the density at two different locations,

$$\xi(r) = \langle \delta(\mathbf{x})\delta(\mathbf{x} + \mathbf{r}) \rangle, \quad (115)$$

which depends only on the norm of \mathbf{r} due to statistical homogeneity and isotropy. The density contrast $\delta(\mathbf{x})$ is usually written in terms of its Fourier components,

$$\delta(\mathbf{x}) = \int d^3\mathbf{k} \delta(\mathbf{k}) \exp(i\mathbf{k} \cdot \mathbf{x}). \quad (116)$$

The quantities $\delta(\mathbf{k})$ are then complex random variables. As $\delta(\mathbf{x})$ is real, it follows that

$$\delta(\mathbf{k}) = \delta^*(-\mathbf{k}). \quad (117)$$

The density field is, therefore, determined entirely by the statistical properties of the random variable $\delta(\mathbf{k})$. We can compute the correlators in Fourier space,

$$\langle \delta(\mathbf{k})\delta(\mathbf{k}') \rangle = \int \frac{d^3\mathbf{x}}{(2\pi)^3} \frac{d^3\mathbf{r}}{(2\pi)^3} \langle \delta(\mathbf{x})\delta(\mathbf{x} + \mathbf{r}) \rangle \exp[-i(\mathbf{k} + \mathbf{k}') \cdot \mathbf{x} - i\mathbf{k}' \cdot \mathbf{r}] \quad (118)$$

which gives,

$$\begin{aligned} \langle \delta(\mathbf{k})\delta(\mathbf{k}') \rangle &= \int \frac{d^3\mathbf{x}}{(2\pi)^3} \frac{d^3\mathbf{r}}{(2\pi)^3} \xi(r) \exp[-i(\mathbf{k} + \mathbf{k}') \cdot \mathbf{x} - i\mathbf{k}' \cdot \mathbf{r}] \\ &= \delta_D(\mathbf{k} + \mathbf{k}') \int \frac{d^3\mathbf{r}}{(2\pi)^3} \xi(r) \exp(i\mathbf{k} \cdot \mathbf{r}) \equiv \delta_D(\mathbf{k} + \mathbf{k}')P(k), \end{aligned} \quad (119)$$

where $P(k)$ is by definition the density *power spectrum*. The inverse relation between two-point correlation function and power spectrum thus reads

$$\xi(r) = \int d^3\mathbf{k} P(k) \exp(i\mathbf{k} \cdot \mathbf{r}). \quad (120)$$

There are basically two conventions in the literature regarding the definition of the power spectrum, which differ by a factor of $(2\pi)^3$. In this review we use the convention in Eqs. (36), (116) and (119) which lead to Eq. (120). Another popular choice is to reverse the role of $(2\pi)^3$ factors in the Fourier transforms, i.e. $\delta(\mathbf{k}) \equiv \int d^3\mathbf{r} \exp(-i\mathbf{k} \cdot \mathbf{r})\delta(\mathbf{r})$, and then modify Eq. (119) to read $\langle \delta(\mathbf{k})\delta(\mathbf{k}') \rangle \equiv (2\pi)^3 \delta_D(\mathbf{k} + \mathbf{k}')P(k)$, which leads to $k^3 P(k)/(2\pi^2)$ being the contribution per logarithmic wavenumber to the variance, rather than $4\pi k^3 P(k)$ as in our case.

3.2.3. The Wick theorem for Gaussian fields

The power spectrum is a well-defined quantity for almost all homogeneous random fields. This concept becomes, however, extremely fruitful when one considers a *Gaussian* field. It means that any joint distribution of local densities is Gaussian distributed. Any ensemble average of product of variables can then be obtained by product of ensemble averages of pairs. We write explicitly this

property for the Fourier modes as it will be used extensively in this work:

$$\langle \delta(\mathbf{k}_1) \cdots \delta(\mathbf{k}_{2p+1}) \rangle = 0, \quad (121)$$

$$\langle \delta(\mathbf{k}_1) \cdots \delta(\mathbf{k}_{2p}) \rangle = \sum_{\text{all pair associations } p \text{ pairs } (i,j)} \prod \langle \delta(\mathbf{k}_i) \delta(\mathbf{k}_j) \rangle. \quad (122)$$

This is the *Wick theorem*, a fundamental theorem for classic and quantum field theories.

The statistical properties of the random variables $\delta(\mathbf{k})$ are then entirely determined by the shape and normalization of $P(k)$. A specific cosmological model will eventually be determined e.g. by the power spectrum in the linear regime, by Ω_m and Ω_Λ only as long as one is only interested in the dark matter behavior.¹³

As mentioned in the previous section, in the case of an inflationary scenario the initial energy fluctuations are expected to be distributed as a Gaussian random field [602,304,280,20]. This is a consequence of the commutation rules given by Eq. (111) for the creation and annihilation operators for a free quantum field. They imply that

$$[(a_{\mathbf{k}} + a_{-\mathbf{k}}^\dagger), (a_{\mathbf{k}'} + a_{-\mathbf{k}'}^\dagger)] = \delta_D(\mathbf{k} + \mathbf{k}'). \quad (123)$$

As a consequence of this, the relations in Eqs. (121) and (122) are verified for $\varphi_{\mathbf{k}}$ for all modes that exit the Hubble radius, which long afterwards come back in as classical stochastic perturbations. These properties obviously apply also to any quantities linearly related to $\varphi_{\mathbf{k}}$.

3.2.4. Higher-order correlators: diagrammatics

In general, it is possible to define higher-order correlation functions. They are defined as the *connected* part (denoted with subscript c) of the joint ensemble average of the density in an arbitrarily number of locations. They can be formally written as

$$\xi_N(\mathbf{x}_1, \dots, \mathbf{x}_N) = \langle \delta(\mathbf{x}_1), \dots, \delta(\mathbf{x}_N) \rangle_c \quad (124)$$

$$\equiv \langle \delta(\mathbf{x}_1), \dots, \delta(\mathbf{x}_N) \rangle - \sum_{\mathcal{S} \in \mathcal{P}(\{\mathbf{x}_1, \dots, \mathbf{x}_N\})} \prod_{s_i \in \mathcal{S}} \xi_{\#s_i}(\mathbf{x}_{s_i(1)}, \dots, \mathbf{x}_{s_i(\#s_i)}), \quad (125)$$

where the sum is made over the proper partitions (any partition except the set itself) of $\{\mathbf{x}_1, \dots, \mathbf{x}_N\}$ and s_i is thus a subset of $\{\mathbf{x}_1, \dots, \mathbf{x}_N\}$ contained in partition \mathcal{S} . When the average of $\delta(\mathbf{x})$ is defined as zero, only partitions that contain no singlets contribute.

The decomposition in connected and non-connected parts can be easily visualized. It means that any ensemble average can be decomposed in a product of connected parts. They are defined, for instance, in Fig. 1. The tree-point moment is “written” in Fig. 2 and the four-point moment in Fig. 3.

In case of a Gaussian field, all connected correlation functions are zero except ξ_2 . This is a consequence of Wick’s theorem. As a result, the only non-zero connected part is the two-point correlation function. An important consequence is that the statistical properties of any field, not necessarily linear, built from a Gaussian field δ can be written in terms of combinations of two-point functions of δ . Note that in a diagrammatic representation, the connected moments of any of such field is represented by a *connected* graph. This is illustrated in Fig. 4 for the field $\delta = \phi^2$: the

¹³ Note that there are now emerging models with a non-standard vacuum equation of state, the so-called quintessence models [536,707], in which the vacuum energy is that of a non-static scalar field. In this case, the detailed behavior of the large-scale structure growth will depend on the dynamical evolution of the quintessence field.

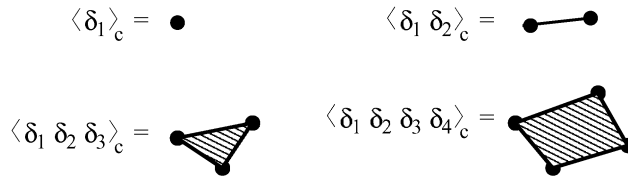


Fig. 1. Representation of the connected part of the moments.

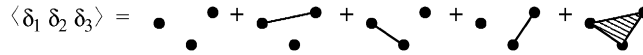


Fig. 2. Writing of the three-point moment in terms of connected parts.

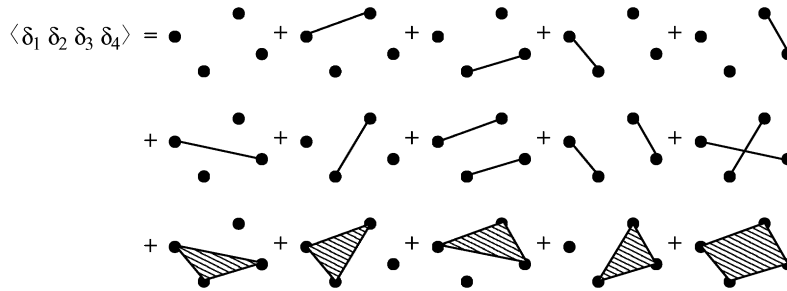


Fig. 3. Same as Fig. 2 for the four-point moment.

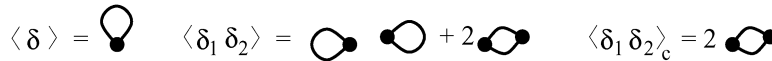


Fig. 4. Disconnected and connected part of the two-point function of the field δ assuming it is given by $\delta = \phi^2$ with ϕ Gaussian.

connected part of the two-point function of this field is obtained by all the diagrams that explicitly join the two points. The other ones contribute to the moments, but not to its connected part.

The connected part has the important property that it vanishes when one or more points are separated by infinite separation. In addition, it provides a useful way of characterizing the statistical properties since unlike unconnected correlation functions, each connected correlation provides independent information.

These definitions can be extended to Fourier space. Because of homogeneity of space $\langle \delta(\mathbf{k}_1) \cdots \delta(\mathbf{k}_N) \rangle_c$ is always proportional to $\delta_D(\mathbf{k}_1 + \cdots + \mathbf{k}_N)$. Then we can define $P_N(\mathbf{k}_1, \dots, \mathbf{k}_N)$ with

$$\langle \delta(\mathbf{k}_1) \cdots \delta(\mathbf{k}_N) \rangle_c = \delta_D(\mathbf{k}_1 + \cdots + \mathbf{k}_N) P_N(\mathbf{k}_1, \dots, \mathbf{k}_N). \tag{126}$$

One particular case that will be discussed in the following is for $n = 3$, the bispectrum, which is usually denoted by $B(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)$.

3.2.5. Probabilities and correlation functions

Correlation functions are directly related to the multi-point probability function, in fact they can be defined from them. Here we illustrate this for the case of the density field, as these results are frequently used in the literature. The physical interpretation of the two-point correlation function is that it measures the excess over random probability that two particles at volume elements dV_1 and dV_2 are separated by distance $x_{12} \equiv |\mathbf{x}_1 - \mathbf{x}_2|$,

$$dP_{12} = n^2 [1 + \xi(x_{12})] dV_1 dV_2 , \quad (127)$$

where n is the mean density. If there is no clustering (random distribution), $\xi = 0$ and the probability of having a pair of particles is just given by the mean density squared, independent of distance. Since the probability of having a particle in dV_1 is $n dV_1$, the conditional probability that there is a particle at dV_2 given that there is one at dV_1 is

$$dP(2|1) = n[1 + \xi(x_{12})] dV_2 . \quad (128)$$

The nature of clustering is clear from this expression; if objects are clustered ($\xi(x_{12}) > 0$), then the conditional probability is enhanced, whereas if objects are anticorrelated ($\xi(x_{12}) < 0$) the conditional probability is suppressed over the random distribution case, as expected. Similar to Eq. (127), for the three-point case the probability of having three objects is given by

$$dP_{123} = n^3 [1 + \xi(x_{12}) + \xi(x_{23}) + \xi(x_{31}) + \xi_3(x_{12}, x_{23}, x_{31})] dV_1 dV_2 dV_3 , \quad (129)$$

where ξ_3 denotes the three-point (connected) correlation function. If the density field were Gaussian, $\xi_3 = 0$, and all probabilities are determined by $\xi(r)$ alone. Analogous results hold for higher-order correlations (e.g. see [508]).

3.3. Moments, cumulants and their generating functions

3.3.1. Moments and cumulants

One particular case for Eq. (125) is when all points are at the same location. Because of statistical homogeneity $\xi_p(\mathbf{x}, \dots, \mathbf{x})$ is independent of the position \mathbf{x} and it reduces to the cumulants of the one-point density probability distribution functions, $\langle \delta^p \rangle_c$. Relation (125) tells us also how the cumulants are related to the moments $\langle \delta^p \rangle$. For convenience we write here the first few terms

$$\begin{aligned} \langle \delta \rangle_c &= \langle \delta \rangle , \\ \langle \delta^2 \rangle_c &= \sigma^2 = \langle \delta^2 \rangle - \langle \delta \rangle_c^2 , \\ \langle \delta^3 \rangle_c &= \langle \delta^3 \rangle - 3\langle \delta^2 \rangle_c \langle \delta \rangle_c - \langle \delta \rangle_c^3 , \\ \langle \delta^4 \rangle_c &= \langle \delta^4 \rangle - 4\langle \delta^3 \rangle_c \langle \delta \rangle_c - 3\langle \delta^2 \rangle_c^2 - 6\langle \delta^2 \rangle_c \langle \delta \rangle_c^2 - \langle \delta \rangle_c^4 , \\ \langle \delta^5 \rangle_c &= \langle \delta^5 \rangle - 5\langle \delta^4 \rangle_c \langle \delta \rangle_c - 10\langle \delta^3 \rangle_c \langle \delta^2 \rangle_c - 10\langle \delta^3 \rangle_c \langle \delta \rangle_c^2 - 15\langle \delta^2 \rangle_c^2 \langle \delta \rangle_c \\ &\quad - 10\langle \delta^2 \rangle_c \langle \delta \rangle_c^3 - \langle \delta \rangle_c^5 . \end{aligned} \quad (130)$$

In most cases $\langle \delta \rangle = 0$ and the above equations simplify considerably. In the following we usually denote σ^2 the local second-order cumulant. The Wick theorem then implies that in case of a Gaussian field, σ^2 is the only non-vanishing cumulant.

It is important to note that the local PDF is essentially characterized by its *cumulants* which constitute a set of independent quantities. This is important since in most of applications that follow the higher order, cumulants are small compared to their associated moments. Finally, let us note that a useful mathematical property of cumulants is that $\langle (b\delta)^n \rangle_c = b^n \langle \delta^n \rangle_c$ and $\langle (b + \delta)^n \rangle_c = \langle \delta^n \rangle_c$, where b is an ordinary number.

3.3.2. Smoothing

The density distribution is usually smoothed with a filter W_R of a given size, R , commonly a top-hat or a Gaussian window. Indeed, this is required by the discrete nature of galaxy catalogs and N -body experiments used to simulate them. Moreover, we shall see later that the scale-free nature of gravitational clustering implies some remarkable properties about the scaling behavior of the smoothed density distribution. The quantities of interest are then the moments $\langle \delta_R^p \rangle$ and the cumulants $\langle \delta_R^p \rangle_c$ of the smoothed density field

$$\delta_R(\mathbf{x}) = \int W_R(\mathbf{x}' - \mathbf{x}) \delta(\mathbf{x}') d^3 \mathbf{x}' . \quad (131)$$

Note that for the top-hat window,

$$\langle \delta_R^p \rangle_c = \int_{v_R} \zeta_p(\mathbf{x}_1, \dots, \mathbf{x}_p) \frac{d^{\mathcal{D}} \mathbf{x}_1 \cdots d^{\mathcal{D}} \mathbf{x}_p}{v_R^p} \quad (132)$$

(where $\mathcal{D}=2$ or 3 is the dimension of the field) is nothing but the average of the N -point correlation function over the corresponding cell of volume v_R .

For a smooth field, equations in Section 3.3.1 are valid for δ as well as δ_R . Some corrections are required if δ is a sum of Dirac delta functions as in real galaxy catalogs. We shall come back to this in Section 6.

In the remaining of this section, we shall omit the subscript R which stands for smoothing, but it will be implicitly assumed.

3.3.3. Generating functions

It is convenient to define a function from which all moments can be generated, namely the *moment generating function* defined by

$$\mathcal{M}(t) \equiv \sum_{p=0}^{\infty} \frac{\langle \delta^p \rangle}{p!} t^p = \int_{-\infty}^{+\infty} p(\delta) e^{t\delta} d\delta = \langle \exp(t\delta) \rangle . \quad (133)$$

The moments can obviously be obtained by subsequent derivatives of this function at the origin $t = 0$. A cumulant generating function can similarly be defined by

$$\mathcal{C}(t) \equiv \sum_{p=2}^{\infty} \frac{\langle \delta^p \rangle_c}{p!} t^p . \quad (134)$$

A fundamental result is that the cumulant generating function is given by the logarithm of the moment generation function (see e.g. Appendix D in [67] for a proof)

$$\mathcal{M}(t) = \exp[\mathcal{C}(t)] . \quad (135)$$

In case of a Gaussian PDF, this is straightforward to check since $\langle \exp(t\delta) \rangle = \exp(\sigma^2 t^2/2)$.

3.4. Probability distribution functions

The probability distribution function (PDF) of the local density can be obtained from the cumulant generating function by inverting Eq. (133).¹⁴ This inverse relation involves the inverse Laplace transform, and can formally be written in terms of an integral in the complex plane (see [16] and Appendix E for a detailed account of this relation):

$$P(\delta) = \int_{-i\infty}^{i\infty} \frac{dt}{2\pi i} \exp[t\delta + \mathcal{C}(t)] . \quad (136)$$

For a Gaussian distribution, the change of variable $t \rightarrow it$ gives the familiar Gaussian integral.

This can be easily generalized to multidimensional PDFs. We then have

$$P(\delta_1, \dots, \delta_p) = \int_{-i\infty}^{i\infty} \frac{dt_1}{2\pi i} \cdots \int_{-i\infty}^{i\infty} \frac{dt_p}{2\pi i} \exp \left[\left(\sum_{q=1}^p t_q \delta_q \right) + \mathcal{C}(t_1, \dots, t_p) \right] , \quad (137)$$

with

$$\mathcal{C}(t_1, \dots, t_p) = \sum_{q_1, \dots, q_p} \langle \delta_1^{q_1} \cdots \delta_p^{q_p} \rangle_c \frac{t_1^{q_1} \cdots t_p^{q_p}}{q_1! \cdots q_p!} . \quad (138)$$

3.5. Weakly non-Gaussian distributions: Edgeworth expansion

Throughout this review we will be often dealing with fields that depart only weakly from a Gaussian distribution. To be more specific, they depart in such a way that

$$\langle \delta^p \rangle_c \sim \sigma^{2p-2} \quad (139)$$

when σ is small.¹⁵ It is then natural to define the coefficient S_p as

$$S_p = \frac{\langle \delta^p \rangle_c}{\sigma^{2p-2}} . \quad (140)$$

(Similar definitions will be introduced subsequently for the other fields.) Introducing the S_p generating function (sometimes also called the cumulant generating function) with

$$\varphi(y) = \sum_{p=2}^{\infty} S_p \frac{(-1)^{p-1}}{p!} y^p = -\sigma^2 \mathcal{C}(-y/\sigma^2) , \quad (141)$$

¹⁴ However, it may happen that the moment or cumulant generating function is not defined because of the lack of convergence of the series in Eq. (133). In this case the PDF is not uniquely defined by its moments. In particular, this is the case for the log-normal distribution. There are indeed other PDFs that have the same moments [312].

¹⁵ This is a consequence of Gaussian initial conditions and the fact that non-linearities in the equations of motion are quadratic, see Section 4.

we get from Eq. (136)

$$P(\delta) d\delta = \frac{d\delta}{2\pi i \sigma^2} \int_{-i\infty}^{+i\infty} dy \exp \left[-\frac{\varphi(y)}{\sigma^2} + \frac{\delta y}{\sigma^2} \right]. \quad (142)$$

Then a number of approximations and truncations can be applied to this expression to decompose the local PDF. This leads to the Edgeworth form of the Gram–Charlier series [609] applied to statistics of weakly non-linear fields. This expansion was derived initially in [405,406] and later proposed in cosmological contexts [552,49,356].

The Edgeworth expansion can be derived from Eq. (142) of the density PDF assuming that the density contrast δ is of the order of σ and small. The relevant values of y are then also of the order of σ and are thus expected to be small. It is then legitimate to expand the function $\varphi(y)$:

$$\varphi(y) \approx -\frac{1}{2}y^2 + \frac{S_3}{3!}y^3 - \frac{S_4}{4!}y^4 + \frac{S_5}{5!}y^5 \pm \dots. \quad (143)$$

To calculate the density PDF, we substitute expansion (143) into the integral in Eq. (142). Then we make a further expansion of the *non-Gaussian* part of the factor $\exp[-\varphi(y)/\sigma^2]$ with respect to both y and σ assuming they are of the same order.

Finally, collecting the terms of the same order in σ we obtain the so-called Edgeworth form of the Gram–Charlier series for density PDF,

$$P(\delta) d\delta = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp(-v^2/2) \left[1 + \sigma \frac{S_3}{6} H_3(v) + \sigma^2 \left(\frac{S_4}{24} H_4(v) + \frac{S_3^2}{72} H_6(v) \right) \right. \\ \left. + \sigma^3 \left(\frac{S_5}{120} H_5(v) + \frac{S_4 S_3}{144} H_7(v) + \frac{S_3^3}{1296} H_9(v) \right) + \dots \right] d\delta, \quad (144)$$

where $v = \delta/\sigma$ and $H_n(v)$ are the Hermite polynomials

$$H_n(v) \equiv (-1)^n \exp(v^2/2) \frac{d^n}{dv^n} \exp(-v^2/2) \\ = v^n - \frac{n(n-1)}{1!} \frac{v^{n-2}}{2} + \frac{n(n-1)(n-2)(n-3)}{2!} \frac{v^{n-4}}{2^2} - \dots, \quad (145)$$

thus

$$H_3(v) = v^3 - 3v, \quad (146)$$

$$H_4(v) = v^4 - 6v^2 + 3, \quad (147)$$

$$H_5(v) = v^5 - 10v^3 + 15v, \quad (148)$$

⋮

This is a universal form for any slightly non-Gaussian field, i.e. when σ is small and S_p are finite. Note that the parameters S_p might vary weakly with σ affecting expansion (144) beyond the third-order term (see [49]).

With such an approach, it is possible to get an approximate form of the density PDF from a few known low-order cumulants. This method is irreplaceable when only a few cumulants have been

derived from first principles. However, it is important to note that this expansion is valid only in the *slightly* non-Gaussian regime. The validity domain of form (144) is limited to finite values of δ/σ , typically $\delta/\sigma \lesssim 0.5$.

A well-known problem with the Edgeworth expansion is that it does not give a positive-definite PDF, in particular this manifests itself in the tails of the distribution. To improve this behavior, an Edgeworth-like expansion about the Gamma PDF (which has exponential tails) has been explored in [258]. To bypass the positivity problem, it was proposed to apply the Edgeworth expansion to the logarithm of the density instead of the density itself [148]. With this change of variable, motivated by dynamics [136], the approximation works well even into the non-linear regime for $\sigma^2 \lesssim 10$ [148,656].

Extensions of Eq. (144) have been written for joint PDFs [406,409]. Note that it can be done only when the cross-correlation matrix between the variables is regular (see [56] for details).

4. From dynamics to statistics: N -point results

A general approach to go from dynamics to statistics would be to solve the Vlasov equation from initial conditions for the phase-space density function $f(\mathbf{x}, \mathbf{p})$ given by a stochastic process such as inflation. Correlation functions in configuration space reviewed in Section 3 can be trivially extended to phase space, and the Vlasov equation yields equations of motion for these phase-space correlation functions. The result is a set of coupled non-linear integro-differential equations, the so-called BBGKY hierarchy,¹⁶ in which the one-point density is related to the two-point phase-space correlation function, the two-point depends on the three-point, and so forth. However, as mentioned in Section 2, if we restrict ourselves to the single stream regime, study of the Vlasov equation reduces to studying the evolution of the density and velocity fields given by the continuity, Euler and Poisson equations. Therefore, all we have to consider in this case is the correlation functions of density and velocity fields.

In this section, we review how the results discussed in Section 2 about the time evolution of density and velocity fields can be used to understand the evolution of their statistical properties, characterized by correlation functions as summarized in the previous section. Most of the calculations will be done assuming Gaussian initial conditions; in this case the main focus is in quantitative understanding of the emergence of non-Gaussianity due to non-linear evolution. In Section 4.4 we discuss results derived from non-Gaussian initial conditions. In Section 5 we present, with similar structure, analogous results for one-point statistics, with emphasis on the evolution of local moments and PDFs.

4.1. The weakly non-linear regime: “tree-level” PT

4.1.1. Emergence of non-Gaussianity

If the cosmic fields are Gaussian, their power spectrum $P(k, \tau)$,

$$\langle \tilde{\delta}(\mathbf{k}, \tau) \tilde{\delta}(\mathbf{k}', \tau) \rangle_c = \delta_D(\mathbf{k} + \mathbf{k}') P(k, \tau) \quad (149)$$

¹⁶ After N.N. Bogoliubov, M. Born, H.S. Green, J.G. Kirkwood and J. Yvon, who independently obtained the set of equations between 1935 and 1962. Rigorously, this route from the Vlasov equation to the BBGKY equations is restricted to the so-called “fluid limit” in which the number of particles is effectively infinite and there are no relaxation effects.

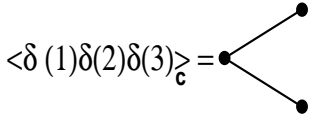


Fig. 5. Tree diagrams for the three-point function or bispectrum.

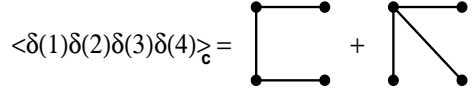


Fig. 6. Tree diagrams for the four-point function or trispectrum.

(or, equivalently, their two-point correlation function) completely describes the statistical properties. However, as we saw in Section 2, the dynamics of gravitational instability is non-linear, and therefore non-linear evolution inevitably leads to the development of non-Gaussian features.

The statistical characterization of non-Gaussian fields is, in general, a non-trivial subject. As we discussed in the previous section, the problem is that in principle all N -point correlation functions are needed to specify the statistical properties of cosmic fields. In fact, for general non-Gaussian fields, it is not clear that correlation functions (either in real or Fourier space) are the best set of quantities that describes the statistics in the most useful way.

The situation is somewhat different for gravitational clustering from Gaussian initial conditions. Here it is possible to calculate in a model-independent way precisely how the non-Gaussian features arise, and what is the most natural statistical description. In particular, since the non-linearities in the equations of motion are quadratic, gravitational instability generates connected higher-order correlation functions that scale as $\xi_N \propto \xi_2^{N-1}$ at large scales, where $\xi_2 \ll 1$ and PT applies [232]. This scaling can be naturally represented by connected tree diagrams, where each link represents the two-point function (or power spectrum in Fourier space), since for N points ($N - 1$) links are necessary to connect them in a tree-like fashion.

As a consequence of this scaling, the so-called *hierarchical amplitudes* Q_N defined by

$$Q_N \equiv \frac{\xi_N}{\sum_{\text{labelings}} \prod_{\text{edges } ij}^{N-1} \xi_2(r_{ij})}, \tag{150}$$

where the denominator is given by all the topological distinct tree diagrams (the different N^{N-2} ways of drawing $N - 1$ links that connect N points), are a very useful set of statistical quantities to describe the properties of cosmic fields. In particular, they are independent of the amplitude of the two-point function, and for scale-free initial conditions they are independent of overall scale. As we shall see, the usefulness of these statistics is not just restricted to the weakly non-linear regime (large scales); in fact, there are reasons to expect that in the opposite regime, at small scales where $\xi_2 \gg 1$, the scaling $\xi_N \propto \xi_2^{N-1}$ is recovered. In this sense, the hierarchical amplitudes Q_N (and their one-point cousins, the S_p parameters) are the most natural set of statistics to describe the non-Gaussianity that results from gravitational clustering.

Figs. 5 and 6 show the tree diagrams that describe the three- and four-point function induced by gravity. As we already said, $N - 1$ links (representing ξ_2) are needed to describe the connected N -point function, and furthermore, the number of lines coming out of a given vertex is the order in PT that gives rise to such a diagram. For example, the diagram in Fig. 5 requires linear and second-order PT, representing $\langle \delta_2(1)\delta_1(2)\delta_1(3) \rangle_c$ (as in Section 2, subscripts describe the order in PT). On the other hand, the diagrams in Fig. 6 require up to third order in PT. The first term represents $\langle \delta_1(1)\delta_2(2)\delta_2(3)\delta_1(4) \rangle_c$ whereas the second describes $\langle \delta_1(1)\delta_3(2)\delta_1(3)\delta_1(4) \rangle_c$.

$$\langle \delta(1)\delta(2) \rangle_c = \text{---} + [\text{---} + \text{---}]$$

Fig. 7. Diagrams for the two-point function or power spectrum up to one loop. See Eqs. (165) and (166) for one-loop diagram amplitudes.

$$\langle \delta(1)\delta(2)\delta(3) \rangle_c = \text{---} + \text{---} + \text{---} + \text{---}$$

Fig. 8. Diagrams for the three-point function or bispectrum up to one loop.

In general, a consistent calculation of the connected p -point function induced by gravity to leading order (“tree-level”) requires from first to $(p - 1)$ th order in PT [232]. At large scales, where $\xi_2 \ll 1$, tree-level PT leads to hierarchical amplitudes Q_N which are independent of ξ_2 . As $\xi_2 \rightarrow 1$, there are corrections to tree-level PT which describe the ξ_2 dependence of the Q_N amplitudes. These are naturally described in terms of diagrams as well, in particular, the next to leading order contributions (“one-loop” corrections) require from first to $(p + 1)$ th order in PT [557]. These are represented by one-loop diagrams, i.e. connected diagrams where there is one closed loop. The additional link over a tree diagram required to form a closed loop leads to $Q_N \propto \xi_2$.

Figs. 7 and 8 show the one-loop diagrams for the power spectrum and bispectrum. The one-loop corrections to the power spectrum (the two terms in square brackets in Fig. 7) describe the non-linear corrections to the linear evolution, that is, the effects of mode coupling and the onset of non-linear structure growth. Recall that each line in a diagram represents the power spectrum $P^{(0)}(k)$ (or two-point function) of the *linear* density field. As a result, the one-loop power spectrum scales $P^{(1)}(k) \propto P^{(0)}(k)^2$.

Are all these diagrams really necessary? In essence, what the diagrammatic representation does is to order the contributions of the same order irrespective of the statistical quantity being considered. For example, it is not consistent to consider the evolution of the power spectrum in second-order PT (second term in Fig. 7) since there is a contribution of the *same* order coming from third-order PT (third term in Fig. 7). Instead, one should consider the evolution of the power spectrum to “one-loop” PT (which includes the two contributions of the same order, the terms in square brackets in Fig. 7). A similar situation happens with the connected four-point function induced by gravity; it is inconsistent to calculate it in second-order PT (first term in Fig. 6), rather a consistent calculation of the four-point function to leading order requires “tree-level” PT (which also involves third-order PT, i.e. the second term in Fig. 6).

We will now review results on the evolution of different statistical quantities in tree-level PT.

4.1.2. Power spectrum evolution in linear PT

The simplest (trivial) application of PT is the leading order contribution to the evolution of the power spectrum. Since we are dealing with the two-point function in Fourier space ($N = 2$), only linear theory is required, that is, the connected part is just given by a single line joining the two points.

In this review we are concerned about time evolution of the cosmic fields during the matter domination epoch. In this case, as we discussed in Section 2, diffusion effects are negligible and the evolution can be cast in terms of perfect fluid equations that describe conservation of mass and momentum. In this case, the evolution of the density field is given by a simple time-dependent scaling of the “linear” power spectrum

$$P(k, \tau) = [D_1^{(+)}(\tau)]^2 P_L(k) , \quad (151)$$

where $D_1^{(+)}(\tau)$ is the growing part of the linear growth factor. One must note, however, that the “linear” power spectrum specified by $P_L(k)$ ¹⁷ derives from the linear evolution of density fluctuations through the radiation domination era and the resulting decoupling of matter from radiation. This evolution must be followed by using general relativistic Boltzmann numerical codes [499,76,416,578], although analytic techniques can be used to understand quantitatively the results [320,321]. The end result is that

$$P_L(k) = k^{n_p} T^2(k) , \quad (152)$$

where n_p is the primordial spectral index ($n_p = 1$ denotes the canonical scale-invariant spectrum [300,706,499]),¹⁸ $T(k)$ is the transfer function that describes the evolution of the density field perturbations through decoupling ($T(0) \equiv 1$). It depends on cosmological parameters in a complicated way, although in simple cases (where the baryonic content is negligible) it can be approximated by a fitting function that depends on the shape parameter $\Gamma \equiv \Omega_m h$ [76,21]. For the adiabatic CDM scenario, $T^2(k) \rightarrow \ln^2(k)/k^4$ as $k \rightarrow \infty$, due to the suppression of fluctuations growth during the radiation dominated era, see e.g. [197] for a review.

4.1.3. The bispectrum induced by gravity

We now focus on the non-linear evolution of the three-point cumulant of the density field, the bispectrum $B(\mathbf{k}_1, \mathbf{k}_2, \tau)$, defined by (compare with Eq. (149))

$$\langle \tilde{\delta}(\mathbf{k}_1, \tau) \tilde{\delta}(\mathbf{k}_2, \tau) \tilde{\delta}(\mathbf{k}_3, \tau) \rangle_c = \delta_D(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) B(\mathbf{k}_1, \mathbf{k}_2, \tau) . \quad (153)$$

As we discussed already, it is convenient to define the reduced bispectrum Q as follows [229,232]:

$$\tilde{Q} \equiv \frac{B(\mathbf{k}_1, \mathbf{k}_2, \tau)}{P(k_1, \tau)P(k_2, \tau) + P(k_2, \tau)P(k_3, \tau) + P(k_3, \tau)P(k_1, \tau)} , \quad (154)$$

which has the desirable property that it is scale and time independent to lowest order (tree level) in non-linear PT,

$$\tilde{Q}^{(0)} = \frac{2F_2(\mathbf{k}_1, \mathbf{k}_2)P(k_1, \tau)P(k_2, \tau) + \text{cyc.}}{P(k_1, \tau)P(k_2, \tau) + P(k_2, \tau)P(k_3, \tau) + P(k_3, \tau)P(k_1, \tau)} , \quad (155)$$

¹⁷ We denote the linear power spectrum interchangeably by $P_L(k)$ or by $P^{(0)}(k)$.

¹⁸ This corresponds to fluctuations in the gravitational potential at the Hubble radius scale that have the same amplitude for all modes, i.e. the gravitational potential has a power spectrum $P_\varphi \sim k^{-3}$, as predicted by inflationary models, see Eq. (113).

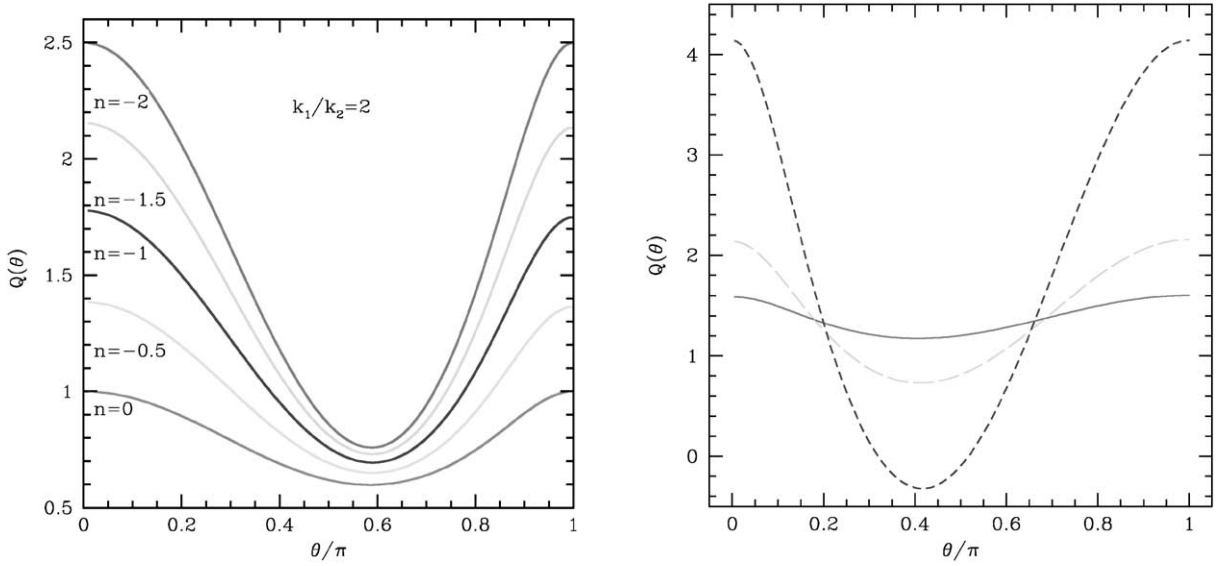


Fig. 9. The tree-level reduced bispectrum $\tilde{Q}^{(0)}$ for triangle configurations given by $k_1/k_2 = 2$ as a function of the angle θ ($\hat{k}_1 \cdot \hat{k}_2 = \cos \theta$). The different curves correspond to spectral indices $n = -2, -1.5, -1, -0.5, 0$ (from top to bottom).

Fig. 10. The tree-level three-point amplitude in real space $Q^{(0)}$ for triangle configurations given by $r_{12}/r_{23} = 2$ as a function of the angle θ ($\hat{r}_{12} \cdot \hat{r}_{23} = \cos \theta$). The different curves correspond to spectral indices $n = -2, -1.5, -1$ (from top to bottom at $\theta = 0.4\pi$).

where $F_2(\mathbf{k}_1, \mathbf{k}_2)$ denotes the second-order kernel obtained from the equations of motion, as in Section 2.4.2. Recall that this kernel is very insensitive to cosmological parameters [see Eq. (71)], as a consequence of this, the tree-level reduced bispectrum $\tilde{Q}^{(0)}$ is almost independent of cosmology [236,313]. In addition, from Eq. (155) it follows that $\tilde{Q}^{(0)}$ is independent of time and normalization [232]. Furthermore, for scale-free initial conditions, $P_L(k) \propto k^n$, $\tilde{Q}^{(0)}$ is also independent of overall scale. For the particular case of equilateral configurations ($k_1 = k_2 = k_3$ and $\hat{k}_i \cdot \hat{k}_j = -0.5$ for all pairs), $\tilde{Q}^{(0)}$ is independent of spectral index as well, $\tilde{Q}_{\text{eq}}^{(0)} = \frac{4}{7}$. In general, for scale-free initial power spectra, $\tilde{Q}^{(0)}$ depends on configuration shape through, e.g. the ratio k_1/k_2 and the angle θ defined by $\hat{k}_1 \cdot \hat{k}_2 = \cos \theta$. In fact, since bias between the galaxies and the underlying density field is known to change this shape dependence [235], measurements of the reduced bispectrum Q in galaxy surveys could provide a measure of bias which is insensitive to other cosmological parameters [236], unlike the usual determination from peculiar velocities which has a degeneracy with the density parameter Ω_m . We will review these applications in Section 8.

Fig. 9 shows $\tilde{Q}^{(0)}$ for the triangle configuration given by $k_1/k_2 = 2$ as a function of the angle θ between these wave vectors ($\cos \theta \equiv \hat{\mathbf{k}}_1 \cdot \hat{\mathbf{k}}_2$) for different spectral indices. The shape or configuration dependence of $\tilde{Q}^{(0)}$ comes from the second-order perturbation theory kernel $F_2^{(s)}$ (see Eqs. (155) and (170)) and can be understood in physical terms as follows. From the recursion relations given in Section 2, we can write

$$F_2^{(s)}(\mathbf{k}_1, \mathbf{k}_2) = \frac{5}{14}[\alpha(\mathbf{k}_2, \mathbf{k}_1) + \alpha(\mathbf{k}_1, \mathbf{k}_2)] + \frac{2}{7}\beta(\mathbf{k}_1, \mathbf{k}_2), \quad (156)$$

with α and β defined in Eq. (39). The terms in square brackets contribute a constant term, independent of configuration, coming from the $\theta \times \delta$ term in the equations of motion, plus terms which depend on configuration and describe gradients of the density field in the direction of the flow (i.e. the term $\mathbf{u} \cdot \nabla \delta$ in the continuity equation). Similarly, the last term in Eq. (156) contributes configuration-dependent terms which come from gradients of the velocity divergence in the direction of the flow (due to the term $(\mathbf{u} \cdot \nabla)\mathbf{u}$ in Euler's equation). Therefore, the configuration dependence of the bispectrum reflects the anisotropy of structures and flows generated by gravitational instability. The enhancement of correlations for collinear wave vectors ($\theta = 0, \pi$) in Fig. 9, reflects the fact that gravitational instability generates density and velocity divergence gradients which are mostly parallel to the flow [559]. The dependence on the spectrum is also easy to understand: models with more large-scale power (smaller spectral indices n) give rise to anisotropic structures and flows with larger coherence length, which upon ensemble averaging lead to a more anisotropic bispectrum.

4.1.4. The three-point correlation function

The three-point function ξ_3 can be found straightforwardly by Fourier transformation of the bispectrum, leading to

$$\begin{aligned} \xi_3(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = & \left[\frac{10}{7} \zeta(x_{13}) \zeta(x_{23}) + \nabla \zeta(x_{13}) \cdot \nabla^{-1} \zeta(x_{23}) \right. \\ & \left. + \nabla \zeta(x_{23}) \cdot \nabla^{-1} \zeta(x_{13}) + \frac{4}{7} (\nabla_a \nabla_b^{-1} \zeta(x_{13})) (\nabla_a \nabla_b^{-1} \zeta(x_{23})) \right] + \text{cyc.} , \end{aligned} \quad (157)$$

where the inverse gradient is defined by the Fourier representation

$$\nabla^{-1} \zeta(x) \equiv -i \int d^3 \mathbf{k} \exp(i \mathbf{k} \cdot \mathbf{x}) \frac{\mathbf{k}}{k^2} P(k) . \quad (158)$$

For scale-free initial conditions, $P(k) \propto k^n$, $\zeta(x) \propto x^{-(n+3)}$ (with $n < 0$ for convergence), and thus

$$\begin{aligned} \xi_3(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = & \left[\frac{10}{7} + \frac{n+3}{n} (\hat{x}_{13} \cdot \hat{x}_{23}) \left(\frac{x_{23}}{x_{13}} + \frac{x_{13}}{x_{23}} \right) \right. \\ & \left. + \frac{4}{7} \left[\frac{3 - 2(n+3) + (n+3)^2 (\hat{x}_{13} \cdot \hat{x}_{23})^2}{n^2} \right] \right] \zeta(x_{13}) \zeta(x_{23}) + \text{cyc.} \end{aligned} \quad (159)$$

Similar to Fourier space, we can define the three-point amplitude in real space Q ,¹⁹

$$Q = \frac{\xi_3(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)}{\zeta(x_{12}) \zeta(x_{23}) + \zeta(x_{23}) \zeta(x_{31}) + \zeta(x_{31}) \zeta(x_{12})} , \quad (160)$$

which is shown in Fig. 10 for spectral indices $n = -2, -1.5, -1$ (solid, dashed and short-dashed, respectively). Note that in real space the three-point amplitude Q has a stronger shape dependence for spectra with more power on small scales (larger spectral index n), unlike the case of Fourier space. This is because scales are weighted differently. Since $\zeta(x)$ is actually equivalent to $k^3 P(k)$ rather than $P(k)$, using $\zeta(x)/x^3$ to define Q in real space rather than $\zeta(x)$ leads to a similar behavior with spectral index than in Fourier space.

¹⁹ In this case, however, one must be careful not to use such a statistic for scales near the zero-crossing of $\zeta(r)$ [100].

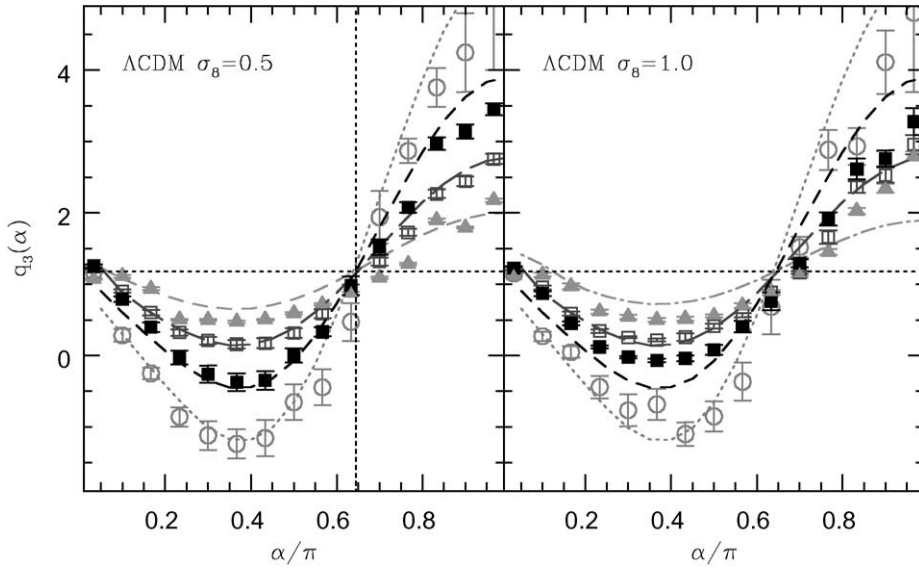


Fig. 11. Tree-level PT predictions of the three-point amplitude $Q^{(0)}$ in the Λ CDM model for triangle configurations given by $r_{12}/r_{23} = 1$ as a function of the angle α ($\hat{r}_{12} \cdot \hat{r}_{23} = \cos \alpha$). The different curves correspond to different triangle sides $r_{12} = 6, 12, 18, 24$ Mpc/h (from top to bottom at $\theta = 0.4\pi$). Symbols with error bars correspond to measurements in numerical simulations at $\sigma_8 = 0.5$ (left panel) and $\sigma_8 = 1.0$ (right panel). From [23].

Note that for scale-free initial conditions, the three-point amplitude for equilateral triangles reduces to the following simple expression as a function of spectral index n :

$$Q_{\text{eq}} = \frac{18n^2 + 19n - 3}{7n^2}. \quad (161)$$

Fig. 11 shows a comparison of the tree-level PT prediction for Q_3 in Λ CDM models (lines) with the fully non-linear values of Q_3 measured in N -body simulations (symbols with error bars). Even on the earlier outputs ($\sigma_8 = 0.5$, left panel) corrections to the tree-level results become important at scales $r_{12} < 12$ Mpc/h. At larger scales there is an excellent agreement with tree-level PT. This seems in contradiction with claims in [346], but note that for the later outputs ($\sigma_8 = 1.0$, right panel) non-linear corrections can be significant at very large scales $r_{12} < 18$ Mpc/h so that for precision measurements one needs to take into account the loop corrections (see [23] for more details).

4.2. The transition to the non-linear regime: “loop corrections”

4.2.1. One-loop PT and previrialization

In the previous section we discussed the leading order contribution to correlations functions, and found that these are given by tree-level PT, resulting in the linear evolution of the power spectrum and in hierarchical amplitudes Q_N independent of the amplitude of fluctuations. Higher-order corrections to tree-level PT (organized in terms of “loop” diagrams) can in principle be calculated, but what new physics do they describe? Essentially, one-loop PT describes the first effects of mode–mode coupling in the evolution of the power spectrum, and the dependence of the hierarchical amplitudes

Q_N on ξ_2 . It also gives a quantitative estimate of where tree-level PT breaks down, and leads to a physical understanding of the transition to the non-linear regime.

One of the main lessons learned from one-loop PT is the fact that non-linear growth of density and velocity fields can be *slower* than in linear PT, in contrast with e.g. the spherical collapse model where non-linear growth is always faster than linear. This effect is due to tidal effects which lead to non-radial motions and thus less effective collapse of perturbations. This was conjectured as a possibility and termed “previrialization” [171]; numerical simulations however showed evidence in favor [677,510] and against [207] this idea. The first quantitative calculation of the evolution of power spectra beyond linear theory for a wide class of initial conditions and comparison with numerical simulations was done in [613], where it was shown that one-loop corrections to the linear power spectrum can be either negative or positive depending on whether the initial spectral index was larger or smaller than $n \approx -1$. Subsequent work confirmed these predictions in greater detail [428,408,558]; in particular, the connection between one-loop corrections to the power spectrum and previous work on previrialization was first emphasized in [408]. In fact, a detailed investigation shows that one-loop PT predicts the change of behavior to occur at $n \approx -1.4$ [558], and divergences appear for $n \gtrsim -1$ which must be cutoff at some small scale in order to produce finite results. We shall come back to this problem below.

In addition, one-loop corrections to the bispectrum show a very similar behavior with initial spectral index [559,560]. For $n \lesssim -1.4$ one-loop corrections increase the configuration dependence of Q , whereas in the opposite case they tend to flatten it out. These results for scale-free initial conditions are relevant for understanding other spectra. Indeed, calculations for CDM spectra [27,334,560] showed that the non-linear power spectrum is smaller than the linear one close to the non-linear scale, where the effective spectral index is $n \gtrsim -1$. Furthermore, these results give insight into the evolution of CDM type of initial spectra: transfer of power happens from large to small scales because more positive spectral indices evolve slower than negative ones. In fact, as a result, non-linear evolution drives the non-linear power spectrum closer to the critical index $n \approx -1$ [558,14].

4.2.2. The one-loop power spectrum

As mentioned above, one-loop corrections to power spectrum (or equivalently to the two-point correlation function) have been extensively studied in the literature [353,678,354,135,613,428,334,27,408,558].²⁰ We now briefly review these results.

We can write the power spectrum up to one-loop corrections as

$$P(k, \tau) = P^{(0)}(k, \tau) + P^{(1)}(k, \tau) + \dots, \quad (162)$$

where the superscript (n) denotes an n -loop contribution, the tree-level (0-loop) contribution is just the linear spectrum,

$$P^{(0)}(k, \tau) = [D_1^{(+)}]^2 P_L(k), \quad (163)$$

and the one-loop contribution consists of two terms (see Fig. 7),

$$P^{(1)}(k, \tau) = P_{22}(k, \tau) + P_{13}(k, \tau), \quad (164)$$

²⁰ Multi-loop corrections to the power spectrum were considered in [237], including the full contributions up to two loops and the most important terms at large k in 3- and 4-loop order.

Table 5
Contributions to the one-loop power spectrum as a function of spectral index n

n	$P_{13}/(\pi A^2 a^4)$	$P_{22}/(\pi A^2 a^4)$
1	$-\frac{122}{315}k^3k_c^2$	$\frac{18}{49}k^4k_c$
0	$-\frac{244}{315}k^2k_c$	$\frac{29\pi^2}{196}k^3$
-1	$\frac{128}{225}k - \frac{4}{3}k \ln \frac{k_c}{\epsilon} - \frac{176}{315}k \ln \frac{k}{k_c}$	$\frac{80}{147}k + \frac{4}{3}k \ln \frac{k}{\epsilon}$
-2	$\frac{5\pi^2}{28k} - \frac{4}{3\epsilon}$	$\frac{75\pi^2}{196k} + \frac{4}{3\epsilon}$

with

$$P_{22}(k, \tau) \equiv 2 \int [F_2^{(s)}(\mathbf{k} - \mathbf{q}, \mathbf{q})]^2 P_L(|\mathbf{k} - \mathbf{q}|, \tau) P_L(q, \tau) d^3\mathbf{q}, \quad (165)$$

$$P_{13}(k, \tau) \equiv 6 \int F_3^{(s)}(\mathbf{k}, \mathbf{q}, -\mathbf{q}) P_L(k, \tau) P_L(q, \tau) d^3\mathbf{q}. \quad (166)$$

Here P_{ij} denotes the amplitude given by a connected diagram representing the contribution from $\langle \delta_i \delta_j \rangle_c$ to the power spectrum. We have assumed Gaussian initial conditions, for which P_{ij} vanishes if $i+j$ is odd. Note the different structure in the two contributions, Eq. (165) is positive definite and describes the effects of mode coupling between waves with wave vectors $\mathbf{k} - \mathbf{q}$ and \mathbf{q} , i.e. if $P_L(k) = 0$ for $k > k_c$, then $P_{22}(k) = 0$ only when $k > 2k_c$. On the other hand, Eq. (166) is in general negative (leading to the effects of previrialization mentioned above) and does not describe mode coupling, i.e. $P_{13}(k)$ is proportional to $P_L(k)$. This term can be interpreted as the one-loop correction to the propagator in Eq. (87) [569], i.e. the non-linear correction to the standard $a(\tau)$ linear growth.

The structure of these contributions can be illustrated by their calculation for scale-free initial conditions, where the linearly extrapolated power spectrum is $P_L(k) = Aa^2k^n$, shown in Table 5. The linear power spectrum is cutoff at low wavenumbers (infrared) and high wavenumbers (ultraviolet) to control divergences that appear in the calculation; that is, $P_L(k) = 0$ for $k < \epsilon$ and $k > k_c$. These results assume $k \gg \epsilon$ and $k \ll k_c$, otherwise there are additional terms [428,558].

The general structure of divergences is that for $n \leq -1$ there are infrared divergences that are caused by terms of the kind $\int P(q)/q^2 d^3q$; these are cancelled when the partial contributions are added. In fact, it is possible to show that this cancellation still holds for leading infrared divergences to arbitrary number of loops [336]. It was shown in [557] that this cancellation is general, infrared divergences arise due to the rms velocity field (whose large-scale limit variance is $\int P(q)/q^2 d^3q$), but since a homogeneous flow cannot affect equal-time correlation functions because of Galilean invariance of the equations of motion, these terms must cancel at the end.

Ultraviolet divergences are more harmful. We see from Table 5 that as $n \geq -1$ the P_{13} contribution becomes ultraviolet divergent (and when $n \geq 1$ for P_{22} as well), but in this case there is no cancellation. Thus, one-loop corrections to the power spectrum are meaningless at face value for scale-free initial conditions with $n \geq -1$. Furthermore, one-loop corrections to the bispectrum are also divergent for scale-free initial spectra as $n \rightarrow -1$. Of course, it is possible that these divergences

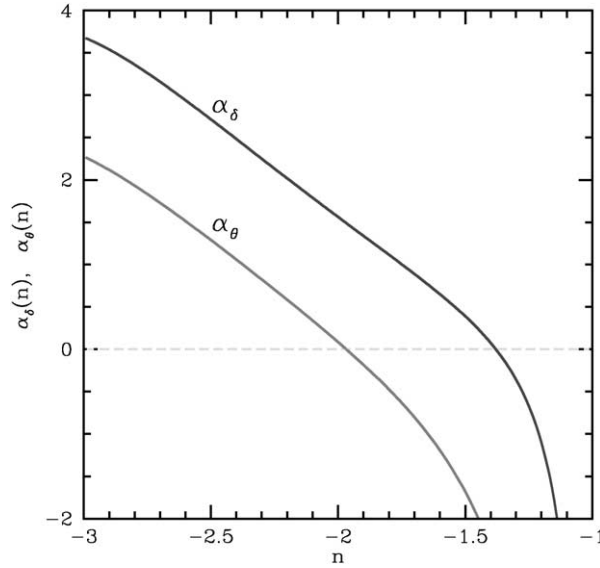


Fig. 12. One-loop corrections to the power spectrum of the density field as a function of spectral index [see Eq. (169)]. Also shown is the one-loop corrections to the velocity divergence power spectrum, $\alpha_\theta(n)$. Note that non-linear effects can slow down the growth of the velocity power spectrum for a broader class of initial conditions than in the case of the density field.

are cancelled by higher-order terms, but to date this has not been investigated. This seems a rather academic problem since no linear power spectrum relevant in cosmology is scale free, and for CDM-type spectra there are no divergences. On the other hand, understanding this problem may shed light on aspects of gravitational clustering in the transition to the non-linear regime.

To characterize the degree of non-linear evolution when including one-loop corrections, it is convenient to define a physical scale from the linear power spectrum, the *non-linear scale* R_0 , as the scale where the smoothed linear variance is unity:

$$\sigma_\ell^2(R_0) = \int d^3\mathbf{k} P_L(k, \tau) W^2(kR_0) \equiv 1. \quad (167)$$

For scale-free initial conditions and a Gaussian filter, $W(x) = \exp(-x^2/2)$, Eq. (167) gives $R_0^{n+3} = 2\pi A a^2 \Gamma[(n+3)/2]$. This is related to the non-linear scale defined from the power spectrum, $\Delta(k_{\text{nl}}) = 4\pi k_{\text{nl}}^3 P(k_{\text{nl}}) = 1$ by

$$k_{\text{nl}} R_0 = \Gamma[(n+5)/2]. \quad (168)$$

Fig. 12 displays the one-loop correction to the power spectrum in terms of the function $\alpha_\delta(n)$ defined by

$$\Delta(k) \equiv \frac{2(kR_0)^{n+3}}{\Gamma[(n+3)/2]} [1 + \alpha_\delta(n)(kR_0)^{n+3}], \quad (169)$$

which measures the strength of one-loop corrections (and similarly for the velocity divergence spectrum replacing α_δ by α_θ). This function has been calculated using the technique of dimensional regularization in [558] (see Appendix D for a brief discussion of this). From Fig. 12 we see that

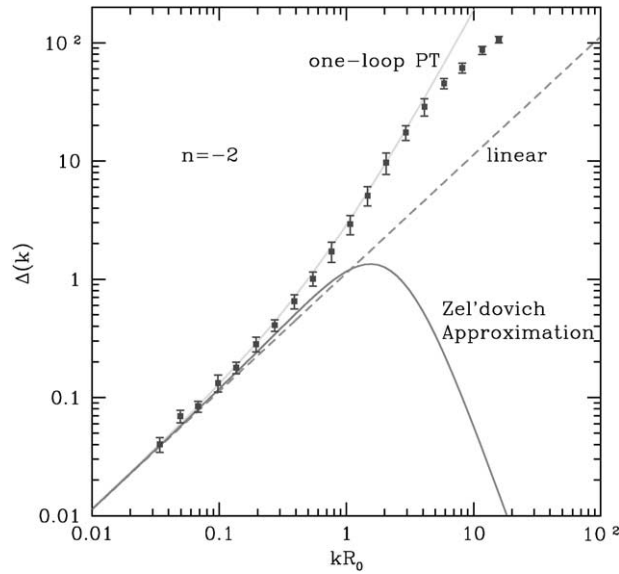


Fig. 13. The power spectrum for $n=-2$ scale-free initial conditions. Symbols denote measurements in numerical simulations from [560]. Lines denote linear PT, one-loop PT [Eq. (169)] and the Zel'dovich approximation results [Eq. (181)], as labeled.

loop corrections are significant with α_δ close to unity or larger for spectral indices $n \lesssim -1.7$. For $n_c \approx -1.4$ one-loop corrections to the power spectrum vanish (and for the bispectrum as well [559]). For this “critical” index, tree-level PT should be an excellent approximation. One should keep in mind, however, that the value of the critical index can change when higher-order corrections are taken into account, particularly, given the proximity of n_c to $n = -1$ where ultraviolet divergences drive $\alpha \rightarrow -\infty$. On the other hand, recent numerical results agree very well with $n_c \approx -1.4$, at least for redshifts $z \sim 3$ evolved from CDM-like initial spectra [702].

Fig. 12 also shows the one-loop correction coefficient α_θ for the velocity divergence spectrum. We see that generally velocities grow much slower than the density field when non-linear contributions are taken into account. For $n \gtrsim -1.9$ one-loop PT predicts that velocities grow slower than in linear PT. Although this has not been investigated in detail against numerical simulations, the general trend makes sense: tidal effects lead to increasingly non-radial motions as n increases, thus the velocity divergence should grow increasingly slower than in the linear case.

Fig. 13 compares the results of one-loop corrections for $n = -2$ against numerical simulations, whereas the top left panel in Fig. 14 shows results for $n = -1.5$. In both cases we see very good agreement even into considerably non-linear scales where $\Delta(k) \sim 10\text{--}100$, providing a substantial improvement over linear PT. Also note the general trend, in agreement with numerical simulations, that non-linear corrections are significantly larger for $n = -2$ than for $n = -1.5$.

4.2.3. The one-loop bispectrum

The loop expansion for the bispectrum, $B = B^{(0)} + B^{(1)} + \dots$, is given by the tree-level part $B^{(0)}$ in terms a single diagram from second-order PT (see Fig. 5) plus its permutations over external

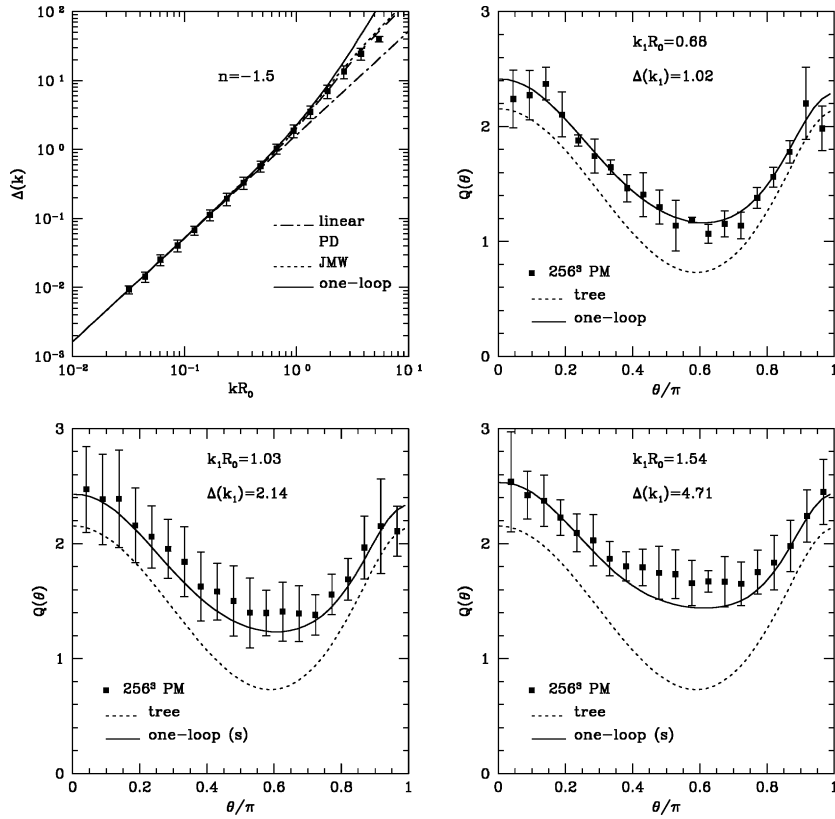


Fig. 14. The left top panel shows the non-linear power spectrum as a function of scale for $n = -1.5$ scale-free initial conditions. Symbols denote measurements in numerical simulations, whereas lines show the linear, and the fitting formulas of [335,494] and one-loop perturbative results, as indicated. The other three panels show the reduced bispectrum Q for triangle configurations with $k_1/k_2 = 2$, as a function of the angle θ between \mathbf{k}_1 and \mathbf{k}_2 , in numerical simulations and for tree-level and one-loop PT. The panels correspond to stages of non-linear evolution characterized by $\Delta(k_1)$. Taken from [560].

momenta (recall that $\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3 \equiv 0$):

$$\begin{aligned}
 B^{(0)} \equiv & 2P_L(k_1)P_L(k_2)F_2^{(s)}(\mathbf{k}_1, \mathbf{k}_2) + 2P_L(k_2)P_L(k_3)F_2^{(s)}(\mathbf{k}_2, \mathbf{k}_3) \\
 & + 2P_L(k_3)P_L(k_1)F_2^{(s)}(\mathbf{k}_3, \mathbf{k}_1) .
 \end{aligned} \tag{170}$$

The one-loop contribution consists of four distinct diagrams involving up to fourth-order solutions [559,560],

$$B^{(1)} \equiv B_{222} + B_{321}^I + B_{321}^{II} + B_{411} , \tag{171}$$

where

$$\begin{aligned}
 B_{222} \equiv & 8 \int d^3\mathbf{q} P_L(q, \tau) F_2^{(s)}(-\mathbf{q}, \mathbf{q} + \mathbf{k}_1) P_L(|\mathbf{q} + \mathbf{k}_1|, \tau) \\
 & \times F_2^{(s)}(-\mathbf{q} - \mathbf{k}_1, \mathbf{q} - \mathbf{k}_2) P_L(|\mathbf{q} - \mathbf{k}_2|, \tau) F_2^{(s)}(\mathbf{k}_2 - \mathbf{q}, \mathbf{q}) ,
 \end{aligned} \tag{172}$$

$$B_{321}^I \equiv 6P_L(k_3, \tau) \int d^3\mathbf{q} P_L(q, \tau) F_3^{(s)}(-\mathbf{q}, \mathbf{q} - \mathbf{k}_2, -\mathbf{k}_3) P_L(|\mathbf{q} - \mathbf{k}_2|, \tau) \\ \times F_2^{(s)}(\mathbf{q}, \mathbf{k}_2 - \mathbf{q}) + \text{permutations} , \quad (173)$$

$$B_{321}^{II} \equiv 6P_L(k_2, \tau) P_L(k_3, \tau) F_2^{(s)}(\mathbf{k}_2, \mathbf{k}_3) \int d^3\mathbf{q} P_L(q, \tau) F_3^{(s)}(\mathbf{k}_3, \mathbf{q}, -\mathbf{q}) \\ + \text{permutations} , \quad (174)$$

$$B_{411} \equiv 12P_L(k_2, \tau) P_L(k_3, \tau) \int d^3\mathbf{q} P_L(q, \tau) F_4^{(s)}(\mathbf{q}, -\mathbf{q}, -\mathbf{k}_2, -\mathbf{k}_3) \\ + \text{permutations} . \quad (175)$$

For the reduced bispectrum \tilde{Q} [see Eq. (154)], the loop expansion yields

$$\tilde{Q} \equiv \frac{B^{(0)} + B^{(1)} + \dots}{\Sigma^{(0)} + \Sigma^{(1)} + \dots} , \quad (176)$$

where $\Sigma^{(0)} \equiv P_L(k_1)P_L(k_2) + P_L(k_2)P_L(k_3) + P_L(k_3)P_L(k_1)$, and its one-loop correction $\Sigma^{(1)} \equiv P^{(0)}(k_1)P^{(1)}(k_2) + \text{permutations}$ (recall $P^{(0)} \equiv P_L$). For large scales, it is possible to expand $\tilde{Q} \equiv \tilde{Q}^{(0)} + \tilde{Q}^{(1)} + \dots$, which gives

$$\tilde{Q}^{(0)} \equiv \frac{B^{(0)}}{\Sigma^{(0)}} , \quad \tilde{Q}^{(1)} \equiv \frac{B^{(1)} - \tilde{Q}^{(0)}\Sigma^{(1)}}{\Sigma^{(0)}} . \quad (177)$$

Note that $\tilde{Q}^{(1)}$ depends on the normalization of the linear power spectrum, and its amplitude increases with time evolution. For initial power-law spectra $P_L(k) = Aa^2k^n$ with $n = -2$, the calculation using dimensional regularization (see Appendix D) yields a closed form; otherwise, the result can be expressed in terms of hypergeometric functions of two variables [559] or computed by direct numerical integration [560].

Fig. 14 shows the predictions of one-loop PT compared to N -body simulations for scale-free initial conditions with $n = -1.5$. In the top right panel, we see that the predictions of Eq. (177) agree very well with simulations at the non-linear scale. In the bottom panels, where $\Delta > 1$, we have used Eq. (176) instead of Eq. (177). At these scales Eq. (176) *saturates*, that is, the one-loop quantities $B^{(1)}$ and $\Sigma^{(1)}$ dominate over the corresponding tree-level values and further time evolution does not change much the amplitude Q , because $B^{(1)}$ and $\Sigma^{(1)}$ have the same scale and, by self-similarity, time dependence. At even more non-linear scales, simulations show that the configuration dependence of the bispectrum is completely washed out [560].

Using the one-loop power spectrum for $n = -2$ given in Table 5, $P^{(1)}(k) = A^2a^455\pi^3/(98k)$, $\tilde{Q}^{(1)}$ follows from Eq. (177). The calculation can be done analytically [559]; for conciseness we reproduce here only the result for equilateral configurations,

$$\tilde{Q}_{\text{eq}} = \frac{4}{7} + \frac{1\,426\,697}{3\,863\,552} \pi^{3/2} k R_0 = 0.57[1 + 3.6 k R_0] \quad (n = -2) \quad (178)$$

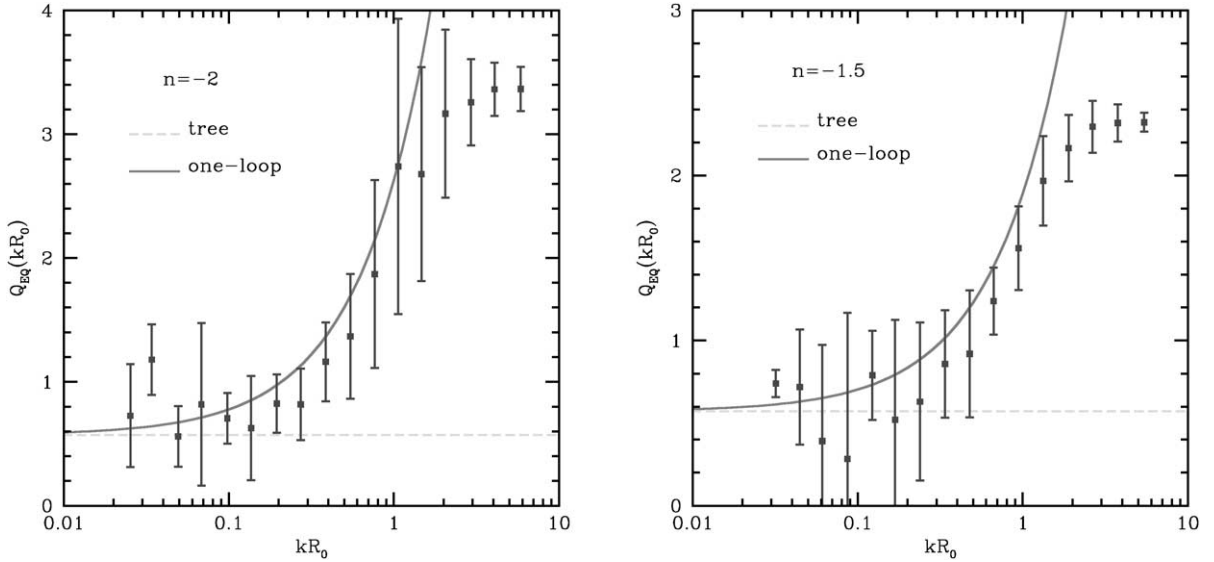


Fig. 15. One-loop bispectrum predictions for equilateral configurations for scale-free spectra with $n = -2$, Eq. (178), and $n = -1.5$, Eq. (179), against N -body simulations measurements from [560]. Error bars come from different output times, assuming self-similarity, see Section 4.5.1. This might not be well obeyed for $n = -2$, due to the importance of finite-volume effects for such a steep spectrum, particularly at late times, see [418] and discussion in Section 6.12.1.

and for $n = -1.5$ we have from numerical integration [560]

$$\tilde{Q}_{\text{eq}} = \frac{4}{7} + 1.32(kR_0)^{3/2} = 0.57[1 + 2.316(kR_0)^{3/2}] \quad (n = -1.5). \quad (179)$$

Fig. 15 compares these results against N -body simulations. We see that despite the strong corrections, with one-loop coefficients larger than unity, one-loop predictions are accurate even at $kR_0 = 1$. As we pointed out before, many of the scale-free results carry over to the CDM case taking into account the effective spectral index. Fig. 16 illustrates the fact that one-loop corrections can increase quite significantly the configuration dependence of the bispectrum at weakly non-linear scales (left panel) when the spectral index is $n < -2$, in agreement with numerical simulations. On the other extreme, in the highly non-linear regime (right panel), the bispectrum becomes effectively independent of triangle shape, with amplitude that approximately matches that of colinear amplitudes in tree-level PT.

Based on results from N -body simulations, it has been pointed out in [234] (see also [240]) that for $n = -1$ non-linear evolution tends to “wash out” the configuration dependence of the bispectrum present at the largest scales (and given by tree-level perturbation theory), giving rise to the so-called hierarchical form $Q \approx \text{const}$ in the strongly non-linear regime (see Section 4.5.5). One-loop perturbation theory must predict this feature in order to be a good description of the transition to the non-linear regime. In fact, numerical integration [559] of the one-loop bispectrum for different spectral indices from $n = -2$ to -1 shows that there is a change in behavior of the non-linear evolution: for $n \lesssim -1.4$ the one-loop corrections *enhance* the configuration dependence of the bispectrum, whereas for $n \gtrsim -1.4$, they tend to cancel it, in qualitative agreement with numerical simulations. Note that this “critical index” $n_c \approx -1.4$ is the same spectral index at which one-loop corrections

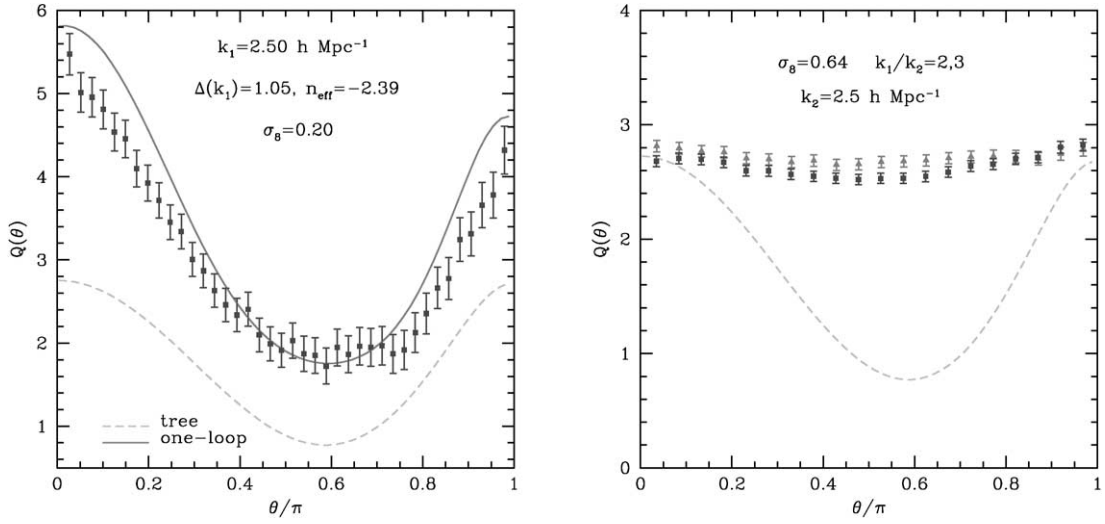


Fig. 16. The left panel shows the one-loop bispectrum predictions for CDM model at scales approaching the non-linear regime, for $k_1/k_2 = 2$ and $\Delta \approx 1$ (left) against numerical simulations [560]. The right panel shows the saturation of \hat{Q} at small scales in the highly non-linear regime, for two different ratios for $k_1/k_2 = 2, 3$ and $\Delta \gtrsim 100$ [563]. Dashed lines in both panels correspond to tree-level PT results.

to the power spectrum vanish, marking the transition between faster and slower than linear growth of the variance of density fluctuations.

4.3. The power spectrum in the Zel'dovich approximation

The ZA [705] is one of the rare cases in which exact (non-perturbative) results can be obtained. However, given the drastic approximation to the dynamics, these exact results for the evolution of clustering statistics are of limited interest due to their restricted regime of validity. The reason behind this is that in the ZA when different streams cross they pass each other without interacting, because the evolution of fluid elements is local. As a result, high-density regions become washed out. Nonetheless, the ZA often provides useful insights into non-linear behavior.

For Gaussian initial conditions, the full non-linear power spectrum in the ZA can be obtained as follows [77,430,556,220,642]. Changing from Eulerian to Lagrangian coordinates, the Fourier transform of the density field is $\delta(\mathbf{k}) = \int d^3\mathbf{q} \exp[i\mathbf{k} \cdot (\mathbf{q} + \Psi)]$, where $\Psi(\mathbf{q})$ is the displacement field. The power spectrum is thus

$$P(k) = \int d^3q \exp(i\mathbf{k} \cdot \mathbf{q}) \langle \exp(i\mathbf{k} \cdot \Delta\Psi) \rangle, \quad (180)$$

where $\Delta\Psi \equiv \Psi(\mathbf{q}_1) - \Psi(\mathbf{q}_2)$ and $\mathbf{q} = \mathbf{q}_1 - \mathbf{q}_2$. For Gaussian initial conditions the ZA displacement is a Gaussian random field, so Eq. (180) can be evaluated in terms of the two-point correlator of $\Psi(\mathbf{q})$. An analytic result for the power spectrum in the ZA has been obtained in [642] for scale-free

initial conditions with $-3 \leq n \leq -1$. For $n = -2$ it is

$$\Delta(k) = \frac{k/k_{\text{nl}}}{[1 + (\pi^2/64)(k/k_{\text{nl}})^2]^2} \left(1 + \frac{3\pi^2}{64} \frac{k/k_{\text{nl}}}{\sqrt{1 + (\pi^2/64)(k/k_{\text{nl}})^2}} \right), \quad (181)$$

where the non-linear wavenumber obeys $\Delta_L(k_{\text{nl}}) = 1$. This result is shown in Fig. 13 (note that in the figure we use R_0 to characterize the non-linear scale, $k_{\text{nl}}R_0 = \Gamma[(n+5)/2]$), together with the prediction of one-loop PT, linear theory and measurements in N -body simulations (symbols with error bars). Clearly the lack of power at small scales due to shell crossing makes the ZA prediction a poor description of the non-linear power spectrum. Attempts have been made in the literature to truncate the small-scale power in the initial conditions and then use ZA [138], this improves the cross-correlation coefficient between ZA and N -body simulation density fields [138,106,455] but it does not bring the power spectrum into agreement [106,455]. Similar results for the effect of shell crossing on the power spectrum hold for 2LPT and 3LPT, see e.g. [106,455,367].

4.4. Non-Gaussian initial conditions

4.4.1. General results

So far we have discussed results for Gaussian initial conditions. When the initial conditions are not Gaussian, higher-order correlation functions are non-zero from the beginning and their evolution beyond linear PT is non-trivial [238]. Here we present a brief summary of the general results for the power spectrum and bispectrum, in the next section we discuss the application to the χ^2 model, for which correlation functions beyond linear perturbation theory have been derived [565]. This belongs to the class of dimensional scaling models, in which the hierarchy of initial correlation functions obey $\xi_N \sim \xi_2^{N/2}$. Another dimensional scaling model that has been studied is the non-linear σ -model [333]. In addition, hierarchical scaling models, where $\xi_N \sim \xi_2^{N-1}$ as generated by gravity from Gaussian initial conditions, have been studied in [414,670]. Most quantitative studies of non-Gaussian initial conditions, however, have been done using one-point statistics rather than correlation functions, we review them in Section 5.6.

It is worth emphasizing that the arguments developed in this section (and in Section 5.6) are valid only if the history of density fluctuations can be well separated into two periods, (i) imprint of non-Gaussian initial fluctuations at very early times, where $\sigma_1 \ll 1$, and then (ii) growth of these fluctuations due to gravitational instability. This is a good approximation for most physically motivated non-Gaussian models.

Let us consider the evolution of the power spectrum and bispectrum from arbitrary non-Gaussian initial conditions.²¹ The first non-trivial correction to the linear evolution of the power spectrum involves second-order PT since $\langle \delta^2 \rangle = \langle (\delta_1 + \delta_2 + \dots)^2 \rangle \approx \langle \delta_1^2 \rangle + 2\langle \delta_1 \delta_2 \rangle + \dots$; the second term which vanishes for the Gaussian case (since $\langle \delta_1 \delta_2 \rangle \sim \langle \delta_1^3 \rangle$) leads instead to²²

$$P(k) = P^I(k) + 2 \int d^3 \mathbf{q} F_2(\mathbf{k} + \mathbf{q}, -\mathbf{q}) B^I(\mathbf{k}, \mathbf{q}), \quad (182)$$

²¹ See [672] for a recent study of the trispectrum for non-Gaussian initial conditions.

²² See Section 5.6 for additional explanation of the new contributions that appear due to primordial non-Gaussianity.

which depends on the initial bispectrum B^I , and similarly for the non-linear evolution of the bispectrum

$$B_{123}^{(0)} = B_{123}^I + B_{123}^G + \int d^3\mathbf{q} F_2(\mathbf{k}_1 + \mathbf{k}_2 - \mathbf{q}, \mathbf{q}) P_4^I(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_1 + \mathbf{k}_2 - \mathbf{q}, \mathbf{q}), \quad (183)$$

where B_{123}^I denotes the contribution of the initial bispectrum, scaled to the present time using linear PT, $B_{123}^I(\tau) \propto [D_1^{(+)}(\tau)]^3$, B_{123}^G represents the usual gravitationally induced bispectrum, Eq. (155), and the last term represents the contribution coming from the initial trispectrum linearly evolved to the present, P_4^I given by

$$\langle \delta^I(\mathbf{k}_1) \delta^I(\mathbf{k}_2) \delta^I(\mathbf{k}_3) \delta^I(\mathbf{k}_4) \rangle_c \equiv \delta_D(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3 + \mathbf{k}_4) P_4^I(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4). \quad (184)$$

Clearly, the complicated term in Eq. (183) is the last one, which involves a convolution of the initial trispectrum with the second-order PT kernel $F_2(\mathbf{k}_1, \mathbf{k}_2)$. Note that only the first term scales as $[D_1^{(+)}(\tau)]^3$, the last two terms have the same scaling with time, $[D_1^{(+)}(\tau)]^4$, and therefore dominate at late times. The structure of these contributions is best illustrated by considering a specific model, as we now do.

4.4.2. χ^2 initial conditions

An example that shows how different the bispectrum can be in models with non-Gaussian initial conditions is the chi-squared model [513,514]. There are, in fact, a number of inflationary models in the literature that motivate χ^2 initial conditions [380,7,405,512]. It is also possible that this particular model may be a good representation of the general behavior of dimensional scaling models, and thus provide valuable insight. In this case, the density field after inflation is proportional to the square of a Gaussian scalar field $\phi(\mathbf{x})$, $\rho(\mathbf{x}) \propto \phi(\mathbf{x})^2$. The initial correlations are most easily calculated in real space [514],

$$\xi_2^I = 2 \frac{\xi_\phi^2(r)}{\sigma_\phi^4}, \quad (185)$$

$$\xi_3^I = 2^{3/2} \sqrt{\xi_2^I(r_{12}) \xi_2^I(r_{23}) \xi_2^I(r_{31})}, \quad (186)$$

$$\begin{aligned} \xi_4^I = 4 \left[\sqrt{\xi_2^I(r_{12}) \xi_2^I(r_{23}) \xi_2^I(r_{34}) \xi_2^I(r_{41})} + \sqrt{\xi_2^I(r_{12}) \xi_2^I(r_{24}) \xi_2^I(r_{43}) \xi_2^I(r_{31})} \right. \\ \left. + \sqrt{\xi_2^I(r_{13}) \xi_2^I(r_{32}) \xi_2^I(r_{24}) \xi_2^I(r_{41})} \right], \quad (187) \end{aligned}$$

where $r_{ij} \equiv |\mathbf{r}_i - \mathbf{r}_j|$. However, non-linear corrections are more difficult to calculate in real space [238], so we turn to Fourier space. The initial density power spectrum and bispectrum read (a similar expression holds for the trispectrum, see [565])

$$P^I(k) = 2 \int d^3\mathbf{q} P_\phi(q) P_\phi(|\mathbf{k} - \mathbf{q}|), \quad (188)$$

$$B^I(k_1, k_2, k_3) = 12 \int d^3\mathbf{q} P_\phi(q) P_\phi(|\mathbf{k}_1 - \mathbf{q}|) P_\phi(|\mathbf{k}_2 + \mathbf{q}|), \quad (189)$$

where $P_\phi(k)$ denotes the power spectrum of the ϕ field. For scale-free spectra, $P_\phi(k) \propto k^{n_\phi}$, $P^I(k) \propto k^{2n_\phi+3}$, with amplitude calculable in terms of gamma functions, similarly, the bispectrum can be expressed in terms of hypergeometric functions [565]. To calculate the hierarchical amplitude to tree level we also need the next to leading order evolution of the power spectrum, that is Eq. (182), which depends on the initial bispectrum, Eq. (189). A simple analytic result is obtained for the particular case, $P_\phi(k) = Ak^{-2}$, not too far from the “canonical” $n_\phi = -2.4$ (e.g. giving $n = -1.8$ [513,514]), then [565]

$$P^I(k) = \frac{2\pi^3 A^2}{k} + \frac{96\pi^4 A^3}{7}, \quad B^I(k_1, k_2, k_3) = \frac{12\pi^3 A^3}{k_1 k_2 k_3}. \quad (190)$$

Defining the non-linear scale k_{nl} from the linear power spectrum as usual, $4\pi k_{\text{nl}}^3 P_L(k_{\text{nl}}) = \Delta_L(k_{\text{nl}}) = 1$, it follows that

$$\Delta(k) = \left(\frac{k}{k_{\text{nl}}}\right)^2 \left(1 + \frac{24}{7\sqrt{2}\pi} \frac{k}{k_{\text{nl}}}\right). \quad (191)$$

Then the tree-level hierarchical amplitude reads [565]

$$\tilde{Q}_{123} = \frac{4\sqrt{2}}{\pi} \frac{k_{\text{nl}}}{k_1 + k_2 + k_3} - \frac{192}{7\pi^2} \frac{k_1 k_2 + k_2 k_3 + k_3 k_1}{(k_1 + k_2 + k_3)^2} + \tilde{Q}_{123}^G + \tilde{Q}_{123}(P_4), \quad (192)$$

where \tilde{Q}_{123}^G denotes the hierarchical amplitude obtained from Gaussian initial conditions and $\tilde{Q}_{123}(P_4)$ denotes the contribution from the last term in Eq. (183) which is difficult to calculate analytically. In particular, for equilateral configurations $\tilde{Q}_{\text{eq}}^I = (4\sqrt{2}/3\pi)(k_{\text{nl}}/k)$. On the other hand, for Gaussian initial conditions, $\tilde{Q}_{\text{eq}}^G = \frac{4}{7}$ independent of spectral index; similarly, there is a contribution from non-Gaussian initial conditions that is scale independent, $\delta\tilde{Q}_{\text{eq}} = -64/7\pi^2$. Since $\tilde{Q}_{123}(P_4)$ is also independent of scale, it turns out that the signature of this type of non-Gaussian initial conditions is that \tilde{Q}_{123} shows a strong scale dependence at large scales as $k/k_{\text{nl}} \rightarrow 0$. This is not just a peculiar property of this particular model, but rather of any non-Gaussian initial conditions with dimensional scaling.²³ Note also that \tilde{Q}^I shows, in some sense, the opposite configuration dependence from \tilde{Q}^G , for triangles where $k_1/k_2 = 2$ as in Fig. 9, $\tilde{Q}^I(\theta)$ is an increasing function of θ , as expected from the scale dependence, in particular $\tilde{Q}^I(\pi)/\tilde{Q}^I(0) = \frac{3}{2}$.

Fig. 17 shows the results of using 2LPT (see Section 2.7) evolved from χ^2 initial conditions [565]. The auxiliary Gaussian field ϕ was chosen to have a spectral index $n_\phi = -2.4$, leading to $n = -1.8$ as proposed in [513]. The amplitude of the power spectrum has been chosen to give $k_{\text{nl}} \equiv 0.33 h/\text{Mpc}$. The dashed lines in Fig. 17 (left panel) show the predictions of the first term in Eq. (192) for the reduced bispectrum at $k_1 = 0.068 h/\text{Mpc}$, $k_2 = 2k_1$, as a function of angle θ between \mathbf{k}_1 and \mathbf{k}_2 . This corresponds to $n = -1$, however, it approximately matches the numerical results (triangles, $n = -1.8$). The latter show less dependence on angle, as expected because the scale dependence in the $n = -1.8$ case ($\tilde{Q}^I \propto k^{-0.6}$) is weaker than for $n = -1$ ($\tilde{Q}^I \propto k^{-1}$). The right panel in

²³ See Section 5.6 for a more detailed discussion of this point and its generalizations.

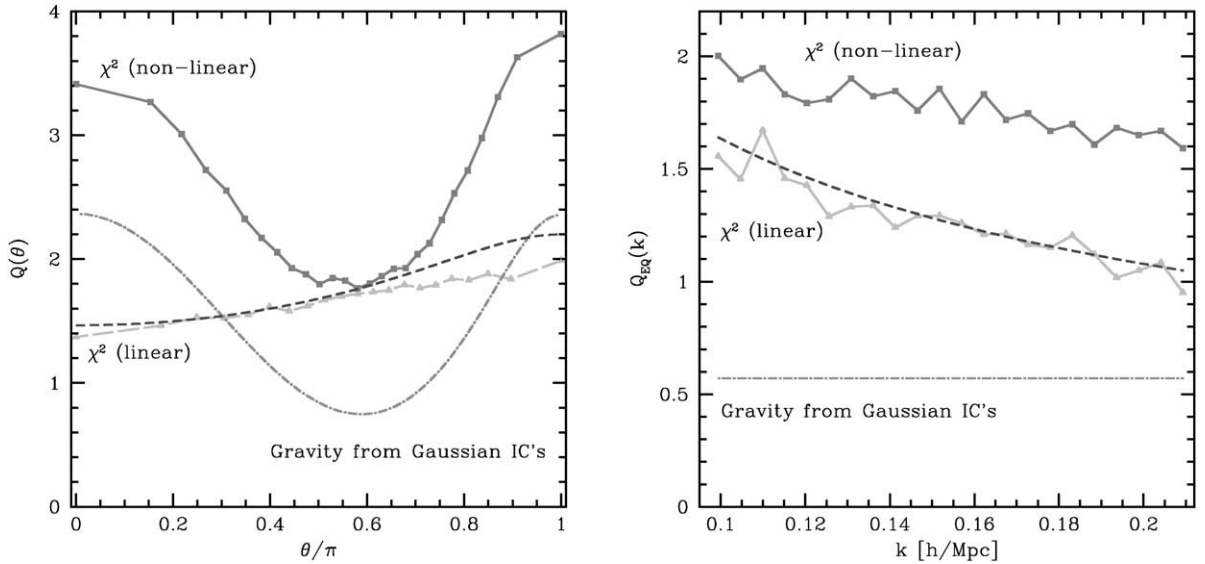


Fig. 17. The reduced bispectrum \tilde{Q} for triangles with sides $k_1 = 0.068 h/\text{Mpc}$ and $k_2 = 2k_1$ as a function of the angle θ between \mathbf{k}_1 and \mathbf{k}_2 (left panel). Right panel shows \tilde{Q} for equilateral triangles as a function of scale k . Triangles denote linear extrapolation from χ^2 initial conditions, whereas square symbols show the result of non-linear evolution. Dot-dashed lines show the predictions of non-linear PT from Gaussian initial conditions with the same initial power spectrum as the χ^2 model.

Fig. 17 shows equilateral configurations as a function of scale for χ^2 initial conditions (triangles) and $\tilde{Q}_{\text{eq}}^1(k) = 0.8(k/k_{\text{nl}})^{-0.6}$ (dashed lines), where the proportionality constant was chosen to fit the numerical result, this is slightly larger than the prediction in the first term of Eq. (192) for $n = -1$ equilateral configurations, and closer to the real-space result $Q_{\text{eq}}(x) = 0.94(x/x_{\text{nl}})^{0.6}$.

The behavior of the χ^2 bispectrum is notoriously different from that generated by gravity from Gaussian initial conditions for identical power spectrum (dot-dashed lines in Fig. 17) [225]. The structures generated by squaring a Gaussian field roughly correspond to the underlying Gaussian high peaks which are mostly spherical, thus the reduced bispectrum is approximately flat. In fact, the increase of \tilde{Q}^1 as $\theta \rightarrow \pi$ seen in Fig. 17 is basically due to the scale dependence of \tilde{Q}^1 , i.e. as $\theta \rightarrow \pi$, the side k_3 decreases and thus \tilde{Q}^1 increases.

As shown in Eq. (192), non-linear corrections to the bispectrum are significant at the scales of interest, so linear extrapolation of the initial bispectrum is insufficient to make comparison with current observations. The square symbols in left panel of Fig. 17 show the reduced bispectrum after non-linear corrections are included. As a result, the familiar dependence of \tilde{Q}_{123} on the triangle shape due to the dynamics of large-scale structures is recovered, and the scale dependence shown by \tilde{Q}^1 is now reduced (right panel in Fig. 17). However, the differences between the Gaussian and χ^2 case are very obvious: the χ^2 evolved bispectrum has an amplitude about 2–4 times larger than that of an initially Gaussian field with the same power spectrum. Furthermore, the χ^2 case shows residual scale dependence that reflects the dimensional scaling of the initial conditions. These signatures can be used to test this model against observations [225,567,211], as we shall discuss in Section 8.

4.5. The strongly non-linear regime

In this section we consider the behavior of the density and velocity fields in the strongly non-linear regime, with emphasis on the connections with PT. Only a limited number of relevant results are known in this regime, due to the complexity of solving the Vlasov equation for the phase-space density distribution. These results, based on simple arguments of symmetry and stability, lead however to valuable insight into the behavior of correlations at small scales.

4.5.1. The self-similar solution

The existence of self-similar solutions relies on two assumptions within the framework of collisionless dark matter clustering:

- (1) There are no characteristic time scales, this requires $\Omega_m = 1$ where the expansion factor scales as a power law, $a \sim t^{2/3}$.
- (2) There are no characteristic length scales. This implies scale-free initial conditions, e.g. Gaussian with initial spectrum $P_1(k) \sim k^n$.

Since gravity is scale free, there are no scales involved in the solution of the coupled Vlasov and Poisson equations. As a result of this, the Vlasov equation admits self-similar solutions with [171]

$$f(\mathbf{x}, \mathbf{p}, t) = t^{-3-3\alpha} \hat{f}(\mathbf{x}/t^\alpha, \mathbf{p}/t^{\beta+1/3}), \quad (193)$$

where $\beta = \alpha + \frac{1}{3}$ and t is the cosmic time. Integration over momentum leads to correlation functions that are only functions of the self-similarity variables $\mathbf{s}_i \equiv \mathbf{x}_i/t^\alpha$, in particular the two-point correlation function reads

$$\xi(\mathbf{x}, t) = f_2\left(\frac{x}{t^\alpha}\right), \quad (194)$$

and similarly for higher-order correlation functions, e.g. $\xi_3(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, t) = f_3(\mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3)$. Note that this solution holds in all regimes, from large to small scales. Using the large-scale behavior expected from linear PT, it is then possible to compute the index α , requiring that $\xi_L(\mathbf{x}, a) \sim a^2 x^{-(n+3)}$ be a function only of the self-similarity variable $xt^{-\alpha}$ leads to

$$\alpha = \frac{4}{3(n+3)}. \quad (195)$$

Note that the self-similar scaling of correlation functions can also be obtained from the fluid equations of motion [558], as expected since only symmetry arguments (which have nothing to do with shell crossing) are involved.²⁴ Self-similarity reduces the dimensionality of the equations of motion; it is possible to achieve further reduction by considering symmetric initial conditions, e.g. planar, cylindrical or spherical. In these cases, exact self-similar solutions can be found by direct numerical integration, see e.g. [214,60]. Although this provides useful insight about the non-linear behavior of isolated perturbations, it does not address the evolution of correlation functions. Detailed results for correlation functions in the non-linear regime can however be obtained by combining the self-similar solution with stable clustering arguments, as we now discuss.

²⁴ For $n = -2$, where finite-volume effects become very important, self-similarity has been difficult to obtain in numerical simulations. However, even in this case current results show that self-similarity is obeyed [338].

4.5.2. Stable clustering

Stable clustering asserts that at small scales, high-density regions decouple from the Hubble expansion and their physical size is stable, i.e. it does not change with time [171]. This implies that the relative motion of particles within gravitationally bound structures should compensate on average the Hubble expansion. Following this idea, general relations can be obtained for the behavior of the two-point correlation function from the continuity equation alone. Indeed, from Eq. (16) it follows that

$$\begin{aligned} \frac{\partial \xi_{12}}{\partial \tau} &= \frac{\partial}{\partial \tau} \langle (1 + \delta(\mathbf{x}_1))(1 + \delta(\mathbf{x}_2)) \rangle \\ &= \langle -\nabla_1 [(1 + \delta(\mathbf{x}_1))\mathbf{u}(\mathbf{x}_1)](1 + \delta(\mathbf{x}_2)) \rangle \\ &\quad - \langle (1 + \delta(\mathbf{x}_1))\nabla_2 [(1 + \delta(\mathbf{x}_2))\mathbf{u}(\mathbf{x}_2)] \rangle . \end{aligned} \quad (196)$$

Pulling out the derivatives using statistical homogeneity, we arrive at the pair conservation equation [171]

$$\frac{\partial \xi_{12}}{\partial \tau} + \nabla_{12} \cdot [\mathbf{u}_{12}(1 + \xi_{12})] = 0 , \quad (197)$$

where the pairwise velocity is defined as

$$\mathbf{u}_{12} \equiv \frac{\langle (1 + \delta(\mathbf{x}_1))(1 + \delta(\mathbf{x}_2))(\mathbf{u}(\mathbf{x}_1) - \mathbf{u}(\mathbf{x}_2)) \rangle}{\langle (1 + \delta(\mathbf{x}_1))(1 + \delta(\mathbf{x}_2)) \rangle} . \quad (198)$$

In the non-linear regime, $\xi \gg 1$, stable clustering implies that the pairwise velocity exactly cancels the Hubble flow, $\mathbf{u}_{12} = -\mathcal{H}\mathbf{x}_{12}$. Under this assumption, Eq. (197) can be readily solved to yield

$$\xi(x, \tau) \approx 1 + \xi(x, \tau) = a^3(\tau)f_2(ax) , \quad (199)$$

which means that the probability of having a neighbor at a fixed physical separation, Eq. (128), becomes independent of time. Eq. (197) can be rewritten as

$$-\frac{u_{12}(x)}{\mathcal{H}x} = \frac{1}{3(1 + \xi(x))} \frac{\partial \xi_{\text{av}}(x)}{\partial \ln a} , \quad (200)$$

which shows that the pairwise velocity is intimately related to the behavior of the two-point function. Here we defined the average two-point function as

$$\xi_{\text{av}}(x) = \frac{3}{x^3} \int_0^x x'^2 dx' \xi(x') \quad (201)$$

and u_{12} is the norm of \mathbf{u}_{12} that can only be along the $\mathbf{x}_2 - \mathbf{x}_1$ direction.

From Eq. (200) it follows that if the time evolution is modeled as following linear PT, then the rhs becomes $2f\xi_{\text{av}}/3$. As $\xi_{\text{av}} \gtrsim 1$, ξ_{av} grows faster than linear theory and thus pairwise velocities overcompensate the Hubble flow; this leads to the well-known “shoulder” (a sudden increase of slope) in the two-point correlation function [271]. These regimes are illustrated in Fig. 18.²⁵ From Eq. (200) it is also clear that a way to model the evolution of the two-point correlation function

²⁵ See [244] for a recent study of the time dependence of the pairwise velocity in the non-linear regime due to merging.

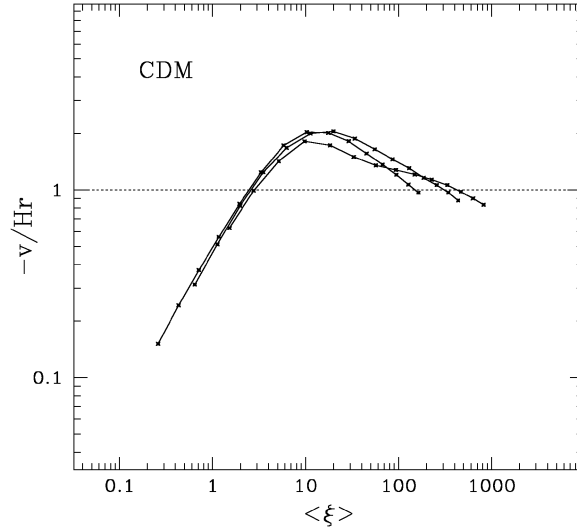


Fig. 18. The ratio of the mean pair (peculiar) velocity to the Hubble velocity, $-u/Hx$, as a function of the mean correlation function ξ_{av} for a CDM model. The pair conservation equation is used to solve for $-u/Hx$ using the evolution of $\xi_{\text{av}}(a, x)$. The three curves are for $a=0.3, 0.6, 0.8$. They would coincide for a scale-free spectrum. They seem to approach the stable clustering value $-u/Hx = 1$ for $\xi_{\text{av}} > 200$. Taken from [337].

is by modeling the dependence of pairwise velocities on ξ_{av} [289,479,358,213,112]. The analysis of high-resolution N -body simulations [358] run by the Virgo Consortium [342] show that the slope of $\xi_2(r)$ indeed exhibits a “shoulder” in the form of an inflection point $d^2\xi_2(r)/dr^2 = 0$ at separation r_* close to the correlation length r_0 where $\xi_2(r_0) = 1$. This property has been recently corroborated for different initial power-spectrum shapes [260]. The equality between r_* and r_0 is related to the fact that loop corrections become important close to the non-linear scale in CDM models at $z = 0$, giving rise to a change in slope. For models where the spectral index at the non-linear scale is very negative (such as CDM models at high redshift, $z \sim 3$, see e.g. [702]), loop corrections can be very large (see Fig. 12), and the non-linear scale r_0 can be much smaller than that where loop corrections become important (related to r_*).

A similar approach can be used to obtain the behavior of higher-order correlation functions under additional stable clustering conditions [508,337]. The starting point is again the continuity equation, Eq. (16), and for the three-point case we have

$$\frac{\partial h_{123}}{\partial \tau} = -\langle \nabla_1 \cdot (A_{123} \mathbf{u}_1) + \nabla_2 \cdot (A_{123} \mathbf{u}_2) + \nabla_3 \cdot (A_{123} \mathbf{u}_3) \rangle, \quad (202)$$

where $A_{123} \equiv (1 + \delta(\mathbf{x}_1))(1 + \delta(\mathbf{x}_2))(1 + \delta(\mathbf{x}_3))$ and $h_{123} \equiv \langle A_{123} \rangle = 1 + \xi_{12} + \xi_{23} + \xi_{31} + \xi_{123}$. Analogous calculations to the two-point case show that

$$\frac{\partial h_{123}}{\partial \tau} + \nabla_{12} \cdot (\mathbf{w}_{12,3} h_{123}) + \nabla_{23} \cdot (\mathbf{w}_{23,1} h_{123}) = 0, \quad (203)$$

where

$$\mathbf{w}_{12,3} \equiv \frac{\langle A_{123}(\mathbf{u}_1 - \mathbf{u}_2) \rangle}{h_{123}}, \quad (204)$$

and similarly for $\mathbf{w}_{23,1}$. Note that these three-body weighted pairwise velocities are actually three-point quantities [337] since a third object is involved, so they are different from Eq. (198). However, in the same spirit as in the two-point case, if we assume that stable clustering leads to $\mathbf{w}_{ij,k} = -\mathcal{H} \mathbf{x}_{ij}$ independent of the position of object k , it follows that the solution of Eq. (203) is

$$\xi_3(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) \approx h_{123} = a^6(\tau) f_3(a \mathbf{x}_1, a \mathbf{x}_2, a \mathbf{x}_3), \quad (205)$$

and thus the probability of having two neighbors at a fixed physical separation $a\mathbf{x}_{12}$ and $a\mathbf{x}_{23}$ from a given object at \mathbf{x}_2 becomes independent of time [e.g. see Eqs. (128) and (129)]. Similar results hold for higher-order N -point correlation functions ξ_N [508], and imply that ξ_N/ξ_2^{N-1} as a function of physical separation become independent of time in the highly non-linear regime ($1 \ll \xi_2 \ll \dots \ll \xi_N$). Note, however, that the additional stability conditions such as $\mathbf{w}_{12,3} \approx -\mathcal{H} \mathbf{x}_{12}$ have not been so far tested against numerical simulations.

4.5.3. Scale invariance

The joint use of stable clustering arguments and the self-similar solution leads to scale-invariant correlation functions in the non-linear regime, with precise predictions for the power-law indices. Eqs. (194) and (199) impose that $f_2(x)$ follows a power law in x ,

$$\xi(x) \sim x^{-\gamma} \quad (206)$$

and matching the time dependences it follows that

$$\gamma = \frac{6}{3\alpha + 2} = \frac{3(n+3)}{(n+5)}. \quad (207)$$

Thus, self-similarity plus stable clustering fixes the full time and spatial dependence of the two-point correlation function in the non-linear regime in terms of the initial conditions [171].

A simple generalization of this argument is to assume that in the non-linear regime $\mathbf{u}_{12} = -h\mathcal{H} \mathbf{x}_{12}$, where h is some constant, not necessarily unity. In this case, Eq. (199) becomes $\xi(x, \tau) = a^{3h}(\tau) f(a^h x)$, and this leads to $\gamma = 3h(n+3)/[2 + h(n+3)]$ [485,697]. Interestingly, if $h(n+3)$ is a constant independent of spectral index n , then the slope of the two-point correlation function becomes independent of initial conditions.²⁶ Current scale-free simulations do not see evidence for a spectral index dependence of the asymptotic value of pairwise velocities and are in reasonable agreement with stable clustering [150,337,164], although the dynamic range in the highly non-linear regime is still somewhat limited. For a different point of view see [486].

The behavior of the higher-order correlation functions can similarly be constrained. Since stable clustering implies that $Q_N \sim \xi_N/\xi_2^{N-1}$ is independent of time, adding self-similarity leads to Q_N being

²⁶ A more detailed analysis of the BBGKY hierarchy shows that, in the absence of self-similarity, power-law solutions for the two-point function in the non-linear regime exist, but their relation to the initial spectral index depends on h , the scaling of ξ_3 in terms of ξ_2 and the skewness of the velocity distribution. Furthermore, perturbations away from self-similarity may not be stable [542,697,698].

independent of overall scale as well; this leads to a scaling relation for higher-order correlations that can be formulated in general as

$$\xi_N(\lambda \mathbf{x}_1, \dots, \lambda \mathbf{x}_N) = \lambda^{-(N-1)\gamma} \xi_N(\mathbf{x}_1, \dots, \mathbf{x}_N), \quad (208)$$

where γ is the index of the two-point function, Eq. (207). As a result, self-similarity plus stable clustering does not fix completely the behavior of the three-point and higher-order correlation functions. Although Q_N does not depend on the overall scale, it does in principle depend on the configuration of the N points, i.e. it can depend on ratios such as x_{12}/x_{23} . This is the same as in tree-level PT, where Q_3 depends on the triangle shape (Figs. 9 and 10).

We should, at this point, reconsider the results in this section from the point of view of the dynamics of gravitational instability. The equations of motion for the two- and three-point correlation functions, Eqs. (197) and (203), which express conservation of pairs and triplets, were obtained from the equation of continuity alone. These are rigorous results. The validity of self-similarity is also rigorous for scale-free initial conditions in a $\Omega_m = 1$ universe. On the other hand, the conditions of stable clustering are only a (physically motivated) ansatz, and they replace what might be obtained by solving the remaining piece of the dynamics, i.e. momentum conservation, in the highly non-linear regime. Note, however, that the conditions of stable clustering can only be part of the story for higher-order correlation functions since these do not explain why e.g. Q_3 tends to become constant independent of triangle configuration in the non-linear regime.

4.5.4. The non-linear evolution of two-point statistics

Self-similarity gives a powerful constraint on the space and time evolution of correlation functions, by requiring that these depend only on the self-similarity variables. However, different initial spectra can lead to very different functions of the self-similarity variables. Hamilton et al. [289] suggested a useful way of thinking about the non-linear evolution of the two-point correlation function, by which the evolution from different initial spectra can all be described by the same (approximately) universal formula, obtained empirically by fitting to numerical simulations.

The starting point is conservation of pairs, Eq. (197), which implies

$$\frac{\partial[x^3(1 + \xi_{\text{av}})]}{\partial \tau} + u_{12} \frac{\partial[x^3(1 + \xi_{\text{av}})]}{\partial x} = 0. \quad (209)$$

Thus, a sphere of radius x such that $x^3(1 + \xi_{\text{av}}) \equiv x_L^3$ is independent of time will contain the same number of neighbors throughout non-linear evolution. At early times, when fluctuations are small, $x_L \approx x$; as clustering develops and becomes non-linear, x becomes smaller than x_L . This motivated the ansatz that the non-linear average two-point correlation function at scale x should be a function of the linear one at scale x_L [289]:

$$\xi_{\text{av}}(x, \tau) = \mathcal{F}_{\text{map}}[\xi_{\text{av}L}(x_L, \tau)], \quad (210)$$

where the mapping \mathcal{F}_{map} was assumed to be universal, i.e. independent of initial conditions. Using more recent numerical simulations Jain et al. [335] showed that there is a dependence of \mathcal{F}_{map} on spectral index (particularly as $n < -1$); in addition Peacock and Dodds [493] extended the mapping above to the power spectrum and arbitrary Ω_m and Ω_Λ . In this case, the non-linear power

spectrum at scale k is assumed to be a function of the linear power spectrum at scale k_L such that $k = [1 + \Delta(k)]^{1/3} k_L$, where $\Delta(k) \equiv 4\pi k^3 P(k)$,

$$\Delta(k, \tau) = \mathcal{F}_{n, \Omega_m, \Omega_\Lambda}[\Delta(k_L, \tau)], \quad (211)$$

where it is emphasized that the mapping depends on spectral index and cosmological parameters. Several groups have reported improved fitting formulae that take into account these extra dependences [335,30,494]. In the most often used version, the fitting function \mathcal{F}_{map} contains five free functions of the spectral index n which interpolate between $\mathcal{F}_{\text{map}}(x) \approx x$ in the linear regime and $\mathcal{F}_{\text{map}} \approx x^{3/2}$ in the non-linear regime where stable clustering is assumed to hold [494]:

$$\mathcal{F}_{\text{map}}(x) = x \left[\frac{1 + B\beta x + [Ax]^{\alpha\beta}}{1 + [(Ax)^\alpha g^3(\Omega)/(Vx^{1/2})]^\beta} \right]^{1/\beta}, \quad (212)$$

where $A = 0.482(1 + n/3)^{-0.947}$, $B = 0.226(1 + n/3)^{-1.778}$, $\alpha = 3.310(1 + n/3)^{-0.244}$, $\beta = 0.862(1 + n/3)^{-0.287}$, $V = 11.55(1 + n/3)^{-0.423}$, and the linear growth factor has been written as $D_1 = ag(\Omega)$ with $g(\Omega) = \frac{5}{2}\Omega_m/[\Omega_m^{4/7} - \Omega_\Lambda + (1 + \Omega_m/2)(1 + \Omega_\Lambda/70)]$ [114]. For models which are not scale free, such as CDM models, the spectral index is taken as $n(k_L) \equiv [d \ln P / d \ln k](k = k_L/2)$ [494]. Extensions of this approach to models with massive neutrinos are considered in [417]; for a description of the non-linear evolution of the bispectrum along these lines see [568].

The ansatz that the non-linear power spectrum at a given scale is a function of the linear power at larger scales is a reasonable first guess, but this cannot be expected to hold in detail. First, as we described in Section 4.2.2, mode coupling leads to a transfer of power from large to small scales (in CDM spectra with decreasing spectral index as a function of scale) and the resulting small-scale power has a contribution from a range of scales in the linear power spectrum. In addition, the mapping above is only based on the pair conservation equation, and thus only takes into account mass conservation. The conditions of validity of the HKLM mapping have been explored in [479], where it is shown that *if* the scaled pairwise velocity $u_{12}/(\mathcal{H}x_{12})$ is *only a function* of the average correlation function, $u_{12}/(\mathcal{H}x_{12}) = H(\xi_{\text{av}})$, then conservation of pairs implies

$$\xi_{\text{av}L}(x_L) = \exp \left[\frac{2}{3} \int^{\xi_{\text{av}}(x)} \frac{ds}{H(s)(1+s)} \right], \quad (213)$$

where x_L and x are related as in the HKLM mapping. In linear PT, $H = 2\xi_{\text{av}}/3$, and if stable clustering holds $H = 1$. In general, however, H cannot be strictly a function of ξ_{av} alone (e.g. due to mode coupling in the weakly non-linear regime). A recent numerical model for the evolution of the pairwise velocity is given in [112], which is used to model the non-linear evolution of the average correlation function.

4.5.5. The hierarchical models

The absence of solutions of the equations of motion in the non-linear regime has motivated the search for consistent relations between correlation functions inspired by observations of galaxy clustering and the symmetries of dynamics, i.e. the self-similar solution. The most common example is the so-called *hierarchical model* for the connected p -point correlation function [275,231] which

naturally obeys the scaling law (208):

$$\xi_N(\mathbf{x}_1, \dots, \mathbf{x}_N) = \sum_{a=1}^{t_N} Q_{N,a} \sum_{\text{labelings edges}} \prod_{N-1} \xi_{AB}. \quad (214)$$

The product is over $N - 1$ edges that link N objects (vertices) A, B, \dots , with a two-point correlation function ξ_{XY} assigned to each edge. These configurations can be associated with “tree” graphs, called N trees. Topologically distinct N trees, denoted by a , in general have different amplitudes, denoted by $Q_{N,a}$, but those configurations which differ only by permutations of the labels $1, \dots, N$ (and therefore correspond to the same topology) have the same amplitude. There are t_N distinct N trees ($t_3 = 1$, $t_4 = 2$, etc., see [232,85]) and a total of N^{N-2} labeled trees.

In summary, the *hierarchical model* represents the connected N -point functions as sums of products of $(N - 1)$ two-point functions, introducing at each level only as many extra parameters $Q_{N,a}$ as there are distinct topologies. In a *degenerate hierarchical model*, the amplitudes $Q_{N,a}$ are furthermore independent of scale and configuration. In this case, $Q_{N,a} = Q_N$, and the hierarchical amplitudes $S_N \simeq N^{N-2} Q_N$. In the general case, it can be expected that the amplitudes Q_N depend on overall scale and configuration. For example, for Gaussian initial conditions, in the *weakly* non-linear regime, $\sigma^2 \ll 1$, perturbation theory predicts a clustering pattern that is hierarchical but not degenerate.

It is important to note that if the degenerate hierarchical holds in the non-linear regime, the Q_N 's should obey positivity constraints. By requiring that the fluctuations of the number density of neighbors should be positive, it follows that [508]

$$Q_3 \geq \frac{1}{3}. \quad (215)$$

This constraint was latter generalized through Schwarz inequalities in [231] to get

$$(2M)^{2M-2} Q_{2M} (2N)^{2N-2} Q_{2N} \geq [(M + N)^{M+N-2} Q_{M+N}]^2, \quad (216)$$

where M and N are integers or odd half-integers. Similar constraints²⁷ have been derived in [57]:

$$(N + 2)^N Q_{N+2} N^{N-2} Q_N \geq [(N + 1)^{N+1} Q_{N+1}]^2. \quad (218)$$

There is no proof, not even indications, that any model fulfilling these constraints is mathematically valid. This is a serious limitation for building such models.

Using the BBGKY hierarchy obtained from the Vlasov equation and assuming a hierarchical form similar to Eq. (214) for the *phase-space N -point distribution function* in the stable clustering limit,

²⁷ A more physically motivated constraint can be derived by imposing that cluster points be more correlated than field points [287,288]. It leads to

$$Q_p \geq \frac{1}{2} \left(\frac{p-1}{p} \right)^{p-3} Q_{p-1} \geq \dots \geq \frac{p!}{2^{p-1} p^{p-2}}, \quad (217)$$

which appear more stringent than the constraints above. These constraints are saturated in the model of Eq. (220) with $Q = \frac{1}{2}$.

Fry [228,231] obtained ($N \geq 3$)

$$Q_N = Q_{N,a} = \frac{1}{2} \left(\frac{N}{N-1} \right) \left(\frac{4Q_3}{N} \right)^{N-2}; \quad (219)$$

in this case, different tree diagrams all have the same amplitude, i.e. the clustering pattern is degenerate. On the other hand, Hamilton [286], correcting an unjustified symmetry assumption in [228,231], instead found

$$Q_{N,\text{snake}} = Q_3^{N-2}, \quad Q_{N,\text{star}} = 0, \quad (220)$$

where “star” graphs correspond to those tree graphs in which one vertex is connected to the other $(N-1)$ vertices, the rest being “snake” graphs (if $Q_3 = \frac{1}{2}$ this corresponds to the Rayleigh–Lévy random walk fractal described in [508]). Summed over the snake graphs, (220) yields

$$Q_N = \frac{N!}{2} \left(\frac{Q_3}{N} \right)^{N-2}. \quad (221)$$

Unfortunately, as emphasized in [286], these results are *not* physically meaningful solutions to the BBGKY hierarchy, but rather a direct consequence of the assumed factorization in *phase space*. As a result, this approach leads to unphysical predictions such as that cluster–cluster correlations are equal to galaxy–galaxy correlations to all orders. It remains to be seen whether physically relevant solutions to the BBGKY hierarchy which satisfy Eq. (214) really do exist. Despite these shortcomings, the results in Eqs. (219) and (220) are often quoted in the literature as physically relevant solutions to the BBGKY hierarchy!

Another phenomenological assumption on the parameters $Q_{N,a}$, which has the virtue of being closer to the mathematical structure found in PT, is provided by the *tree hierarchical model* [41,473,57]. In this case, the parameters $Q_{N,a}$ are obtained by the product of weights v_i associated to each of the vertex appearing in the tree structure,

$$Q_{N,a} = \prod_i v_i^{d_i(a)}. \quad (222)$$

In this expression the product is made over all vertices appearing in configuration a , v_i is the weight of the vertex connected to i lines and $d_i(a)$ is the number of such vertices. The parameter $Q_{N,a}$ is, therefore, completely specified by the star diagram amplitudes. This pattern is analogous to what emerges from PT at large scales, although the parameters $Q_{N,a}$ are here usually taken to be constant, independent of scale and configuration. But even in the absence of this latter hypothesis, the genuine tree structure²⁸ of the tree hierarchical model turned out to be very useful for phenomenological investigations (see [57] and Section 7.1).

4.5.6. Hyperextended perturbation theory

More direct connections with PT results have been proposed to build models of non-linear clustering. One is known as the “hyperextended perturbation theory” (HEPT [563]).²⁹ Its construction is based on the observation that colinear configurations play a special role in gravitational clustering, which become apparent in the discussion on the bispectrum loop corrections (see Section 4.2.3). They correspond to matter flowing parallel to density gradients, thus enhancing clustering at small

²⁸ In the sense that any part of the diagram can be computed irrespective of the global configuration.

²⁹ A more phenomenological model, extended perturbation theory (EPT), is presented in Section 5.13.

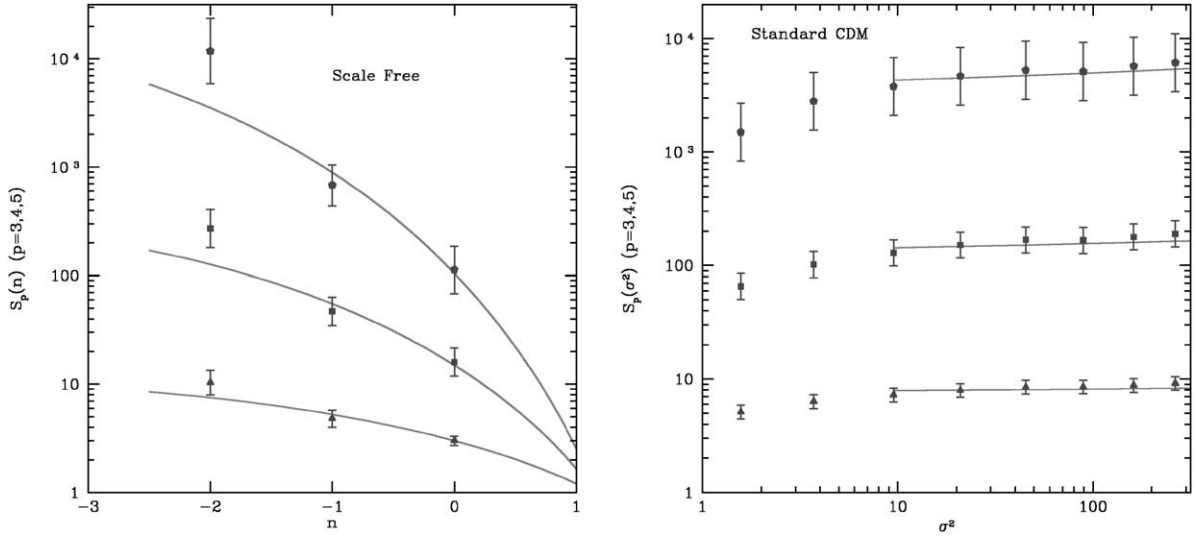


Fig. 19. HEPT compared to N -body simulations for scale-free initial conditions (left) and CDM (right).

scales until eventually giving rise to bound objects that support themselves by velocity dispersion (virialization). *HEPT conjectures that the “effective” Q_N clustering amplitudes in the strongly non-linear regime are the same as the weakly non-linear (tree-level PT) colinear amplitudes*, as shown in Fig. 16 to hold well for three-point correlations.

Note that by effective amplitudes Q_N^{eff} the overall magnitude of Q_N is understood: it is possible that Q_N , for $N > 3$, although independent of overall scale, is a function of configuration. To calculate the resulting S_N parameters, it is further assumed that $S_N \simeq N^{N-2} Q_N^{\text{eff}}$, that is, the S_N are given by the typical configuration amplitude Q_N^{eff} times the total number of labeled trees, N^{N-2} , neglecting a small correction due to smoothing [85]. The resulting non-linear S_N amplitudes follow from tree-level PT [563]:

$$S_3^{\text{sat}}(n) = 3Q_3^{\text{sat}}(n) = 3 \frac{4 - 2^n}{1 + 2^{n+1}}, \quad (223)$$

$$S_4^{\text{sat}}(n) = 16Q_4^{\text{sat}}(n) = 8 \frac{54 - 27 \times 2^n + 2 \times 3^n + 6^n}{1 + 6 \times 2^n + 3 \times 3^n + 6 \times 6^n}, \quad (224)$$

$$S_5^{\text{sat}}(n) = 125Q_5^{\text{sat}}(n) = \frac{125}{6} \frac{N(n)}{D(n)}, \quad (225)$$

where n is the spectral index, obtained from $(n+3) \equiv -\text{dln } \sigma_L^2(R)/\text{dln } R$, $N = 1536 - 11 \cdot 522^n + 1283^n + 664^n + 646^n - 98^n - 212^n - 24^n$, $D = 1 + 122^n + 123^n + 164^n + 246^n + 248^n + 1212^n + 2424^n$. One can check that these Q_N amplitudes satisfy the above positivity constraints, Eqs. (216) and (218) and even the constraint in Eq. (217) as long as $n \lesssim 0.75$, which is well within the physically interesting range.

The left panel of Fig. 19 shows a comparison of these predictions with the numerical simulation measurements in [150] for scale-free initial conditions with $\Omega_m = 1$. The plotted values correspond to

the measured value of S_p when the non-linear variance $\sigma^2 = 100$. We see that the N -body results are generally in good agreement with the predictions of HEPT, Eqs. (223)–(225), keeping in mind that for $n = -2$ finite-volume corrections to the S_p measured in the simulations are quite large and thus uncertain (see Section 6.12.1). The right panel shows a similar comparison of HEPT with numerical simulations in the non-linear regime for the SCDM model ($\Gamma = 0.5$, $\sigma_8 = 0.34$ [147]). The agreement between the N -body results and the HEPT predictions is excellent in this case. The small change in predicted value of S_p with scale is due to the scale dependence of the linear CDM spectral index.

It is interesting to note that for $n = 0$, HEPT predicts $S_p = (2p - 3)!!$, which agrees exactly with the excursion set model developed in [588] for white-noise Gaussian initial fluctuations. In this case, the one-point PDF yields an inverse Gaussian distribution, which has been shown to agree well in the non-linear regime when compared to numerical simulations [588]. This remarkable agreement between HEPT and the excursion set model deserves further study.

5. From dynamics to statistics: the local cosmic fields

We have seen in Section 4 that the non-linear nature of gravitational dynamics leads, through mode-coupling effects, to the emergence of non-Gaussianity. In the previous section we have explored the behavior of multi-point correlation functions. Here we present statistical properties related to the local density contrast in real space. We first describe the results that have been obtained for the moments of the local density field. In particular, we show how to compute the full cumulant generating function of the one-point density contrast at tree level. Results including loop corrections are given when known. Finally, we present techniques for the computation of the density PDF and various applications of these results. When dealing with smoothed fields, we shall assume that filtering is done with a top-hat window unless specified otherwise.

5.1. The density field third moment: skewness

5.1.1. The unsmoothed case

The first non-trivial moment that emerges due to mode coupling is the third moment of the local density PDF, characterized by the skewness parameter. The computation of the leading order term of $\langle \delta^3 \rangle$ is obtained through the expansion $\langle \delta^3 \rangle = \langle (\delta^{(1)} + \delta^{(2)} + \dots)^3 \rangle$. When the terms that appear in this formula are organized in increasing powers of the local linear density, we have $\langle \delta^3 \rangle = \langle (\delta^{(1)})^3 \rangle + 3\langle (\delta^{(1)})^2 \delta^{(2)} \rangle + \dots$, where the neglected terms are of higher order in PT. The first term of this expansion is identically zero for Gaussian initial conditions. The second term is therefore the leading order, “tree-level” in diagrammatic language (see Section 4.1). We then have³⁰

$$\langle \delta^3 \rangle \approx 3\langle (\delta^{(1)})^2 \delta^{(2)} \rangle \quad (226)$$

$$\begin{aligned} &= 3 a^4 \int d^3 \mathbf{k}_1 \dots \int d^3 \mathbf{k}_4 \langle \delta_1(\mathbf{k}_1) \delta_1(\mathbf{k}_2) \delta_1(\mathbf{k}_3) \delta_1(\mathbf{k}_4) \rangle \\ &\quad \times F_2(\mathbf{k}_2, \mathbf{k}_3) \exp[i(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3 + \mathbf{k}_4) \cdot \mathbf{x}] . \end{aligned} \quad (227)$$

³⁰ For simplicity, calculations in this section are done for the Einstein–de Sitter case, $\Omega_m = 1$.

For Gaussian initial conditions, linear Fourier modes $\delta_1(\mathbf{k})$ can only correlate in pairs [Eq. (122)]. If \mathbf{k}_2 and \mathbf{k}_3 are paired, the integral vanishes [because $\langle \delta \rangle = 0$, see the structure of the kernel F_2 in Eq. (45)]. The other two pairings give identical contributions, and thus

$$\langle \delta^3 \rangle = 6a^4 \int d^3 \mathbf{k}_1 \int d^3 \mathbf{k}_2 P(k_1)P(k_2)F_2(\mathbf{k}_1, \mathbf{k}_2) . \quad (228)$$

Integrating over the angle between \mathbf{k}_1 and \mathbf{k}_2 leads to $\langle \delta^3 \rangle = (\frac{34}{7}) \langle \delta^2 \rangle^2$ [508]. For the reasons discussed in Section 4.1.1, it is convenient to rescale the third moment and define the skewness parameter S_3 (see Section 2):

$$S_3 \equiv \frac{\langle \delta^3 \rangle}{\langle \delta^2 \rangle^2} = \frac{34}{7} + \mathcal{O}(\sigma^2) . \quad (229)$$

The skewness measures the tendency of gravitational clustering to create an asymmetry between underdense and overdense regions (see Fig. 20). Indeed, as clustering proceeds there is an increased probability of having large values of δ (compared to a Gaussian distribution), leading to an enhancement of the high-density tail of the PDF. In addition, as underdense regions expand and most of the volume becomes underdense, the maximum of the PDF shifts to negative values of δ . From Eq. (144) we see that the maximum of the PDF is in fact reached at

$$\delta_{\max} \approx -\frac{S_3}{2} \sigma^2 , \quad (230)$$

to first order in σ . We thus see that the skewness factor S_3 contains very useful information on the shape of the PDF.

5.1.2. The smoothed case

At this stage however the calculation in Eq. (229) is somewhat academic because it applies to the statistical properties of the local, unfiltered, density field. In practice, the fields are always observed at a finite spatial resolution (whether it is in an observational context or in numerical simulations). The effect of filtering, which amounts to convolving the density field with some window function, should be taken into account in the computation of S_3 . The main difficulty lies in the complexity this brings into the computation of the angular integral. To obtain the skewness of the local filtered density, δ_R , one indeed needs to calculate

$$\langle \delta_R^3 \rangle = 3 \langle (\delta_R^{(1)})^2 \delta_R^{(2)} \rangle \quad (231)$$

with

$$\delta_R^{(1)} = a \int d^3 \mathbf{k} \delta(\mathbf{k}) \exp[i\mathbf{k} \cdot \mathbf{x}] W_3(k_1 R) , \quad (232)$$

$$\begin{aligned} \delta_R^{(2)} = a^2 \int d^3 \mathbf{k}_1 \int d^3 \mathbf{k}_2 \delta(\mathbf{k}_1) \delta(\mathbf{k}_2) \exp[i(\mathbf{k}_1 + \mathbf{k}_2) \cdot \mathbf{x}] \\ \times F_2(\mathbf{k}_1, \mathbf{k}_2) W_3(|\mathbf{k}_1 + \mathbf{k}_2| R) , \end{aligned} \quad (233)$$

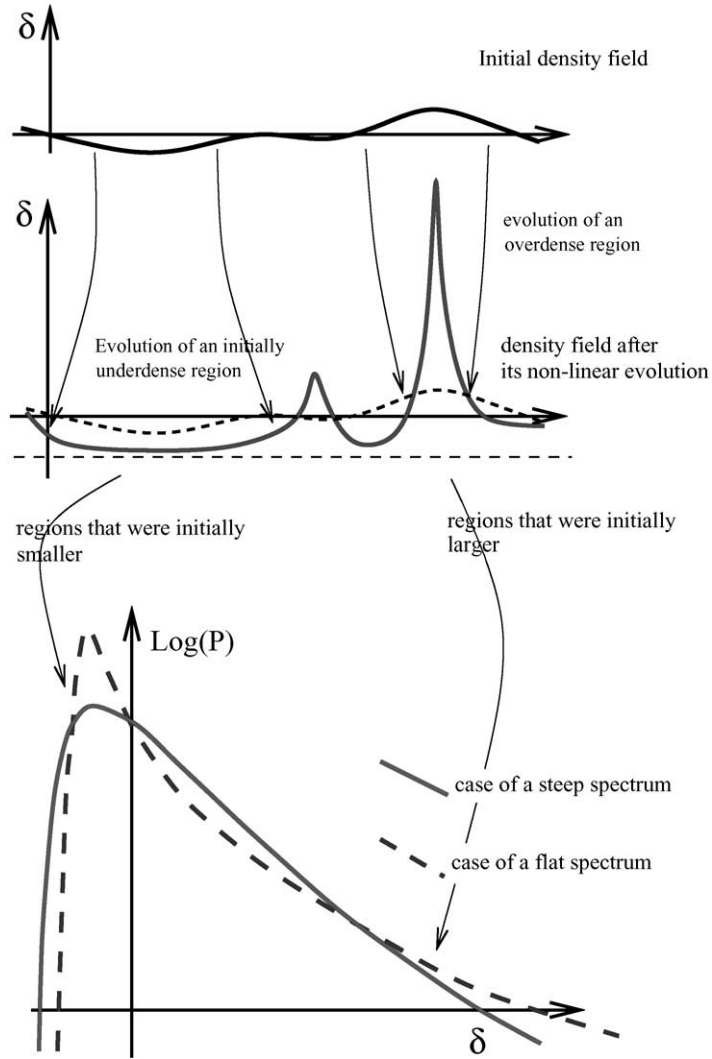


Fig. 20. Skewness is a measure of the asymmetry of the local density distribution function. It appears because underdense regions evolve less rapidly than overdense regions as soon as non-linearities start to play a role. The dependence of skewness on the shape of the power spectrum comes from a mapping between Lagrangian space, in which the initial size of the perturbation is determined, and Eulerian space. For a given filtering scale R , overdense regions come from the collapse of regions that had initially a larger size, whereas underdense regions come from initially smaller regions. As a result, the skewness is expected to be smaller for power spectra with more small-scale fluctuations (steep spectra case, that is when $k^3 P(k)$ is rapidly increasing with k).

where $W_3(k)$ is the 3D filtering function in Fourier space. It leads to the expression for the third moment,

$$\begin{aligned} \langle \delta_R^3 \rangle &= 6a^4 \int d^3 \mathbf{k}_1 \int d^3 \mathbf{k}_2 P(k_1) P(k_2) W_3(k_1 R) W_3(k_2 R) \\ &\quad \times F_2(\mathbf{k}_1, \mathbf{k}_2) W_3(|\mathbf{k}_1 + \mathbf{k}_2| R) \end{aligned} \quad (234)$$

so that the relative angle between \mathbf{k}_1 and \mathbf{k}_2 appears in both F_2 and W_3 . The result depends obviously on the filtering procedure. It turns out that the final result is simple for a top-hat filter in real space. In this case

$$W_3(k) = \sqrt{\frac{3\pi}{2}} \frac{J_{3/2}(k)}{k^{3/2}} = \frac{3}{k^3} [\sin(k) - k \cos(k)]. \quad (235)$$

Following the investigations initiated in [355] for the properties of the top-hat window function³¹ it can be shown (see [46] and Appendix C) that

$$\int \frac{d\Omega_{12}}{4\pi} W_3(|\mathbf{k}_1 + \mathbf{k}_2|) \left[1 - \frac{(\mathbf{k}_1 \cdot \mathbf{k}_2)^2}{k_1^2 k_2^2} \right] = \frac{2}{3} W_3(k_1) W_3(k_2), \quad (236)$$

$$\int \frac{d\Omega_{12}}{4\pi} W_3(|\mathbf{k}_1 + \mathbf{k}_2|) \left[1 + \frac{\mathbf{k}_1 \cdot \mathbf{k}_2}{k_1^2} \right] = W_3(k_1) \left[W_3(k_2) + \frac{1}{3} k_2 W_3'(k_2) \right]. \quad (237)$$

It is easy to see that F_2 can be expressed with the help of the two polynomials involved in the preceding relations. One finally obtains [46]

$$S_3 = \frac{34}{7} + \frac{d \log \sigma^2(R)}{d \log R}. \quad (238)$$

The skewness thus depends on the power spectrum shape (mainly at the filtering scale). For a power-law spectrum, $P(k) \propto k^n$, it follows that $S_3 = \frac{34}{7} - (n + 3)$ [355]. Galaxy surveys indicate that the spectral index n is of the order of $n \approx -1.5$ close to the non-linear scale. Comparisons with numerical simulations have shown that the prediction of Eq. (238) is very accurate, as can be seen in Fig. 27.

5.1.3. Physical interpretation of smoothing

To understand the dependence of the skewness parameter on power spectrum shape it is very instructive to examine in detail the nature of the contributions that appear when the filtering effects are taken into account.

For this purpose let us consider the same problem in Lagrangian space. If one calculates $J^{(2)}$, the second-order expansion of the Jacobian, one obtains [from Eqs. (90) and (94) and assuming $\Omega_m = 1$]

$$J^{(2)} = a^2 \frac{2}{7} \int d^3 \mathbf{k}_1 \int d^3 \mathbf{k}_2 \delta(\mathbf{k}_1) \delta(\mathbf{k}_2) \exp[i(\mathbf{k}_1 + \mathbf{k}_2) \cdot \mathbf{q}] \left[1 - \frac{(\mathbf{k}_1 \cdot \mathbf{k}_2)^2}{k_1^2 k_2^2} \right]. \quad (239)$$

This gives for the density [e.g. Eq. (91)], once the Jacobian (which is a direct estimation of the volume) has been filtered at a given *Lagrangian* scale R ,

$$\begin{aligned} \delta_R^{(2)} &= \int d^3 \mathbf{k}_1 \int d^3 \mathbf{k}_2 a^2 \delta(\mathbf{k}_1) \delta(\mathbf{k}_2) \exp[i(\mathbf{k}_1 + \mathbf{k}_2) \cdot \mathbf{q}] \\ &\times \left[W(k_1 R) W(k_2 R) - \frac{2}{7} W(|\mathbf{k}_1 + \mathbf{k}_2| R) \left(1 - \frac{(\mathbf{k}_1 \cdot \mathbf{k}_2)^2}{k_1^2 k_2^2} \right) \right]. \end{aligned} \quad (240)$$

³¹ These properties have been obtained from the summation theorem of Bessel functions, see e.g. [681]. Such relations hold in any space dimension for top-hat filters.

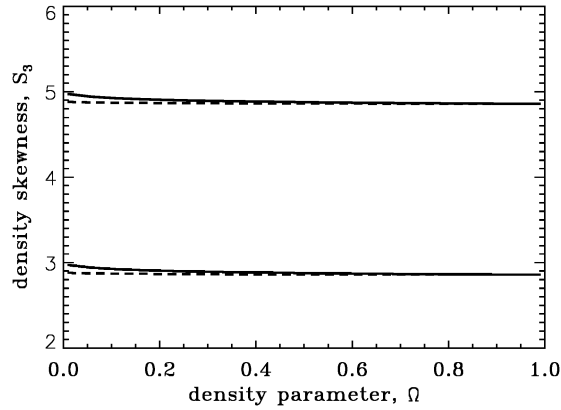


Fig. 21. The skewness S_3 as a function of Ω_m for zero- Ω_A universes (solid lines) and flat universes with $\Omega_m + \Omega_A = 1$ (dashed lines). The upper and lower curves correspond to a power-law spectrum with $n = -3$ and -1 , respectively.

Because smoothing effects are calculated in Lagrangian space (denoted by \mathbf{q}), this expression is different from the Eulerian space filtering result, Eq. (233). In fact, it follows that $S_3^{\text{Lag}} = \frac{34}{7}$ even when filtering effects are taken into account. The mere fact that one does not obtain the same result should not be surprising. In this latter case the filtering has been made at a given *mass* scale. The difference between the two calculations comes from the fact that the larger the mass of a region initially is, the smaller the volume it occupies will be. Filtering at a fixed Eulerian scale therefore mixes different initial mass scales. The asymmetry will then be less than one could have expected because, for a standard hierarchical spectrum, larger mass scales correspond to smaller fluctuations.

5.1.4. Dependence of the skewness on cosmological parameters

As the skewness is induced by gravitational dynamics, it is important to know how much it can depend on cosmological parameters. In general, the parameter S_3 depends on the growth rate of the second-order PT solution, see Section 2.4.3, through

$$S_3 = 3v_2 + \frac{d \log \sigma^2(R)}{d \log R}. \quad (241)$$

Explicit calculations [91] have shown that v_2 can be well approximated by

$$v_2 \approx \frac{4}{3} + \frac{2}{7} \Omega_m^{-2/63}, \quad (242)$$

obtained by expansion about $\Omega_m = 1$ for $\Omega_A = 0$.³² We then have the following result:

$$S_3 = \frac{34}{7} + \frac{6}{7}(\Omega_m^{-0.03} - 1) - (n + 3). \quad (243)$$

A similar result follows when $\Omega_A \neq 0$, see [46,313] and also [223]. In practice, for current applications to data, such a small dependence on cosmological parameters can simply be ignored, as illustrated in Fig. 21. This turns out to be true even when cosmologies with non-standard vacuum equation of state are considered (e.g. quintessence models) [366,259,34].

³² But it is valid for all values of Ω_m of cosmological interest.

5.1.5. The skewness of the local velocity divergence

The skewness of the velocity divergence can obviously be calculated in a similar fashion. However, because of the overall $f(\Omega_m, \Omega_\Lambda)$ factor for the linear growth of velocities, it is natural to expect that the velocity divergence skewness parameter, T_3 , has a significant Ω_m dependence [50]. In general,

$$T_3 \equiv \frac{\langle \theta^3 \rangle}{\langle \theta^2 \rangle^2} = -\frac{1}{f(\Omega_m, \Omega_\Lambda)} \left[3\mu_2 + \frac{d \log \sigma^2(R)}{d \log R} \right]. \quad (244)$$

Taking into account the specific time dependence of μ_2 we get,

$$T_3 = -\frac{1}{f(\Omega_m, \Omega_\Lambda)} \left[2 + \frac{12}{7} \Omega_m^{-1/21} + \frac{d \log \sigma^2(R)}{d \log R} \right], \quad (245)$$

which within a very good accuracy implies that $T_3 \approx -[\frac{26}{7} - (n+3)]/\Omega_m^{0.6}$ for a power-law spectrum. This makes the dimensionless quantity T_3 a very good candidate for the determination of Ω_m independent of galaxy biasing. Attempts to carry out such measurements, however, faced very large systematics in the data [50]. So far no reliable constraints have been drawn from this technique.

5.2. The fourth-order density cumulant: kurtosis

The previous results can be applied to any low-order cumulants of the cosmic field. Fry [232] computed the fourth cumulant of the cosmic density field, but without taking into account the filtering effects. These were included later for top-hat [46] and Gaussian filters [407].

Formally, the fourth-order cumulant of the local density is given by

$$\begin{aligned} \langle \delta^4 \rangle_c &\equiv \langle \delta^4 \rangle - 3\langle \delta^2 \rangle^2 \\ &= 12\langle (\delta^{(1)})^2 (\delta^{(2)})^2 \rangle_c + 4\langle (\delta^{(1)})^3 \delta^{(3)} \rangle_c. \end{aligned} \quad (246)$$

In these equations it is essential to take the connected part only. There are terms that involve loop corrections to the variance that are of the same order in σ but they naturally cancel when the non-connected part of the fourth moment is subtracted out. The consequence is that

$$\langle \delta^4 \rangle_c \sim \langle \delta^2 \rangle^3, \quad (247)$$

and one can define the kurtosis parameter S_4 ,

$$S_4 \equiv \langle \delta^4 \rangle_c / \langle \delta^2 \rangle^3. \quad (248)$$

This equation allows one to compute the leading part of S_4 in the weakly non-linear regime. In general, S_4 can be expressed in terms of the functions D_1 , v_2 and v_3 . This can be obtained by successive applications of the geometrical properties of the top-hat window function (see [46] and Appendix C for details). We have

$$\begin{aligned} S_4 &= 4v_3 + 12v_2^2 + (14v_2 - 2) \frac{d \log[\sigma^2(R_0)]}{d \log R_0} \\ &+ \frac{7}{3} \left(\frac{d \log[\sigma^2(R_0)]}{d \log R_0} \right)^2 + \frac{2}{3} \frac{d^2 \log[\sigma^2(R_0)]}{d \log^2 R_0}. \end{aligned} \quad (249)$$

For a power-law spectrum of index n this leads to

$$S_4 = \frac{60712}{1323} - \frac{62}{3}(n+3) + \frac{7}{3}(n+3)^2. \quad (250)$$

This result is exact for an Einstein–de Sitter universe. It is extremely accurate, within a few percent for all models of cosmological interest. Similar results can be obtained for the velocity divergence.

5.3. Results for Gaussian smoothing filters

So far we have been giving results for a top-hat filter only. The reason is that they can be given in a closed form for any shape of the power spectrum. Another quite natural filter to choose is the Gaussian filter. In this case, however, there are no simple closed forms that are valid for any power spectrum shape. Results are known for power-law spectra only [355,436,407].

The principle of the calculation in this case is to decompose the angular part that enters the window function as a sum of Legendre functions,

$$e^{-\mathbf{p} \cdot \mathbf{q}} = e^{-pq \cos \varphi} = \sum_{m=0}^{\infty} (-1)^m (2m+1) \sqrt{\frac{\pi}{2pq}} I_{m+1/2}(pq) P_m(\cos \varphi), \quad (251)$$

where $I_{m+1/2}(pq)$ are Bessel functions. The integration over φ is made simple by the orthogonality relation between the Legendre polynomials. Finally, each term appearing in the decomposition of the Bessel function

$$I_\nu(z) = \sum_{m=0}^{\infty} \frac{1}{m! \Gamma(\nu+m+1)} \left(\frac{z}{2}\right)^{\nu+2m} \quad (252)$$

can be integrated out for power-law spectra since

$$\int_0^\infty q^\alpha e^{-q^2} dq = \frac{1}{2} \Gamma\left(\frac{\alpha+1}{2}\right), \quad (253)$$

which after resummation leads to hypergeometric functions of the kind ${}_2F_1$. Eventually, the result for S_3 is

$$S_3 = 3 {}_2F_1\left(\frac{n+3}{2}, \frac{n+3}{2}, \frac{3}{2}, \frac{1}{4}\right) - \left(n + \frac{8}{7}\right) {}_2F_1\left(\frac{n+3}{2}, \frac{n+3}{2}, \frac{5}{2}, \frac{1}{4}\right) \quad (254)$$

and similarly the velocity skewness is

$$T_3 = -3 {}_2F_1\left(\frac{n+3}{2}, \frac{n+3}{2}, \frac{3}{2}, \frac{1}{4}\right) + \left(n + \frac{16}{7}\right) {}_2F_1\left(\frac{n+3}{2}, \frac{n+3}{2}, \frac{5}{2}, \frac{1}{4}\right). \quad (255)$$

This result is exact for an Einstein–de Sitter universe but obviously, as for the top-hat filter, S_3 is expected to depend only weakly on cosmological parameters and the dominant dependence of T_3 is that proportional to $1/f(\Omega_m)$. The result for S_3 is shown as a dashed line in Fig. 26.

The kurtosis cannot be calculated in closed form even for power-law spectra (although a semi-analytic formula can be given [407]). However, there exists a simple prescription that allows one to get an approximate expression for the kurtosis. It consists in using the formal expression of the

kurtosis obtained for a top-hat filter but calculated for $n = n_{\text{eff}}$ such that it gives the correct value for the skewness. Such a prescription has been found to give accurate results, about 1% accuracy for $n = -1$ [407].

5.4. The density cumulants hierarchy

In general, the non-linear couplings are going to induce non-zero cumulants at any order. We can define [207]

$$S_p \equiv \langle \delta^p \rangle_c / \langle \delta^2 \rangle^{p-1}, \quad (256)$$

that generalizes the S_3 and S_4 parameters considered in the previous section. All these quantities are finite (and non-zero) at large scales for Gaussian initial conditions and can in principle be computed from PT expansions. However, the direct calculation of S_p becomes extremely difficult with increasing order p due to the complexity of the kernels F_p and G_p . Fortunately, it turns out to be possible to take great advantage of the close relationship between the S_p parameters and the vertices v_p describing the spherical collapse dynamics, as described in Section 2.4.2, to compute the S_p parameters for any p .

In the derivation presented here we adopt a pedestrian approach for building, step by step, the functional shape of the cumulant generating function. A more direct approach has recently been developed in [660,661] in which the generating function of the cumulant is obtained directly, via a saddle-point approximation in the computation of the cumulant generating function which corresponds to its tree-order calculation. This approach avoids technical difficulties encountered in the computation of the Lagrangian space filtering properties and in the Lagrangian–Eulerian mapping and is certainly an interesting complementary view to what we present here.

5.4.1. The unsmoothed density cumulant generating function

The computation of S_p coefficients is based on the property that each of them can be decomposed into a sum of product of “vertices”, at least when filtering effects are not taken into account. As seen before, $S_4 = 12v_2^2 + 4v_3$. This property extends to all orders so that the S_p parameters can be expressed as functions of v_q ’s only ($q = 2, \dots, p - 1$). Note that the vertices v_p defined in Eq. (48) as angular averages of PT kernels correspond to

$$v_p = \langle \delta^{(p)} [\delta^{(1)}]^p \rangle_c / \langle [\delta^{(1)}]^2 \rangle^p. \quad (257)$$

This decomposition of S_p into a sum of product of vertices can be observed easily in a graphical representation. Indeed

$$\langle \delta^p \rangle_c = \sum_{q_i} \langle \delta^{(q_1)} \dots \delta^{(q_p)} \rangle_c, \quad (258)$$

where each δ has been expanded in PT. Each $\delta^{(q)}$ contains a product of q random Gaussian variables $\delta(\mathbf{k})$. Each of these points can be represented by one dot so that when the ensemble average is computed, because of the Wick theorem, dots are connected pairwise. The $\delta^{(q)}$, therefore, can be represented as in Fig. 22 with q outgoing lines.

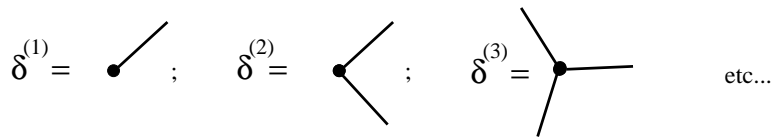


Fig. 22. Diagrammatic representation of $\delta^{(p)}$. Each line stands for a factor $\delta(\mathbf{k})$.

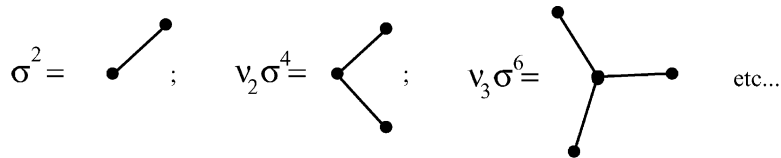


Fig. 23. Computation of the simplest graphs. Each line represents a factor σ^2 . Vertices are obtained from the angular average of the wave vectors leaving v_p .

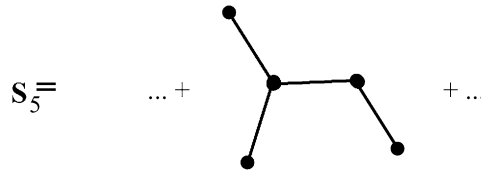


Fig. 24. A graph contributing to S_5 .

Diagrams that contribute to the leading order of S_p are those which contain enough dots so that a connected diagram that minimizes the number of links can be built. The number of links for connecting p points is $p - 1$, we should then have $\sum_i q_i = 2(p - 1)$ so that

$$S_p = \sum_{\text{graphs, } \sum_i q_i = 2(p-1)} \langle \delta^{(q_1)} \dots \delta^{(q_p)} \rangle_c / \langle [\delta^{(1)}]^2 \rangle^{p-1}. \tag{259}$$

An example of such a graph for S_5 is shown in Fig. 24.

It is worth noting that all these diagrams are trees so that the integration over the wave vectors can be made step by step.³³ Then the value of each diagram is obtained by assigning each line to the value of σ^2 and each vertex to v_p depending on the number p of lines it is connected to, see e.g. Fig. 23.

This order-by-order decomposition can actually be replaced by a functional relation at the level of the generating functions. If we define the generating function of S_p as

$$\varphi(y) = \sum_{p=1}^{\infty} -S_p \frac{(-y)^p}{p!} \quad (S_1 = S_2 \equiv 1) \tag{260}$$

and the vertex generating function as

$$\mathcal{G}_\delta(\tau) = \sum_{p=1}^{\infty} v_p \frac{(-\tau)^p}{p!}, \tag{261}$$

³³ This is possible however only when smoothing effects are neglected.

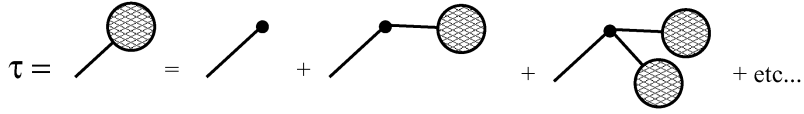


Fig. 25. Graphical representation of Eq. (263), τ is the generating function of graphs with one external line.

it is possible to show that φ and \mathcal{G}_δ are related to each other through the system of equations

$$\varphi(y) = y\mathcal{G}_\delta[\tau(y)] + \frac{1}{2} \tau^2(y) , \tag{262}$$

$$\tau(y) = -y \mathcal{G}'_\delta[\tau(y)] . \tag{263}$$

The demonstration of these equations is not straightforward and is given in Appendix B. To get some insight about these two equations, one can note that τ is the conjugate variable to the one-line vertex (that is v_1 , set to unity at the end of the calculation). As such, it corresponds to the generating function of all graphs with *one* external line. It is then solution of an implicit equation, illustrated in Fig. 25, which corresponds to Eq. (263). Naturally, it involves the vertex generating function. It is to be noted however that in this perspective Eqs. (262) and (263) and the parameter y have no intrinsic physical interpretation. It has been pointed out recently in [660,661] that this system can actually be obtained directly from a saddle-point approximation in the computation of the local density contrast PDF. It gives insights into the physical meaning of the solutions of Eq. (263). We will come back to this point in Section 5.8.

Recall that vertices describe the spherical collapse dynamics (see Section 2.4.2), thus $\mathcal{G}_\delta(\tau)$ corresponds to the density contrast of collapsing structures with spherical symmetry when $(-\tau)$ is its linear density contrast. The first few values of v_p can then be easily computed,

$$v_2 = \frac{34}{21}, \quad v_3 = \frac{682}{189}, \quad v_4 = \frac{446\,440}{43\,659} , \tag{264}$$

which implies,

$$S_3 = 3v_2 = \frac{34}{7} , \tag{265}$$

$$S_4 = 4v_3 + 12v_2^2 = \frac{60\,712}{1323} \approx 45.89 , \tag{266}$$

$$S_5 = 5v_4 + 60v_3v_2 + 60v_2^3 = \frac{200\,575\,880}{305\,613} \approx 656.3 , \tag{267}$$

$$S_6 = 6v_5 + 120v_4v_2 + 90v_3^2 + 720v_3v_2^2 + 360v_2^4 \approx 12\,700$$

$$\vdots \tag{268}$$

At this stage however, the effects of filtering have not been taken into account.

5.4.2. Geometrical properties of smoothing in Lagrangian space

As the examination of the particular case of S_3 has shown, the smoothing effects for a top-hat filter are entirely due to the mapping between Lagrangian and Eulerian spaces. This can be generalized to any order [44].

The Lagrangian space dynamics is jointly described by the displacement field (that plays a role similar to the velocity field) and the Jacobian, whose inverse gives the density. The latter can be expanded with respect to the initial density contrast:

$$J(\mathbf{q}) = 1 + J^{(1)}(\mathbf{q}) + J^{(2)}(\mathbf{q}) + \dots \quad (269)$$

At a given order we will have³⁴

$$J^{(p)}(\mathbf{q}) = a^p \int \frac{d^3\mathbf{k}_1}{(2\pi)^{3/2}} \dots \frac{d^3\mathbf{k}_p}{(2\pi)^{3/2}} J_p(\mathbf{k}_1, \dots, \mathbf{k}_p) \exp[i\mathbf{q} \cdot (\mathbf{k}_1 + \dots + \mathbf{k}_p)] \quad (270)$$

The Jacobian is actually given by the determinant of the deformation tensor, obtained from the first derivative of the displacement field, Ψ , see Eq. (91). The precise relation is

$$J(\mathbf{q}) \equiv \left| \frac{\partial \mathbf{x}}{\partial \mathbf{q}} \right| = 1 + \nabla_{\mathbf{q}} \cdot \Psi + \frac{1}{2} \left[(\nabla_{\mathbf{q}} \cdot \Psi)^2 - \sum_{ij} \Psi_{i,j} \Psi_{j,i} \right] \\ + \frac{1}{6} \left[(\nabla_{\mathbf{q}} \cdot \Psi)^3 - 3 \nabla_{\mathbf{q}} \cdot \Psi \sum_{ij} \Psi_{i,j} \Psi_{j,i} + 2 \sum_{ijk} \Psi_{i,j} \Psi_{j,k} \Psi_{k,i} \right] \quad (271)$$

The equations of motion are closed by the Euler equation, Eq. (90). This shows that the kernels of the Jacobian expansion are built recursively from the function $\beta(\mathbf{k}_1, \mathbf{k}_2) = 1 - (\mathbf{k}_1 \cdot \mathbf{k}_2)^2 / (k_1 k_2)^2$ and

$$\eta(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = 1 - \left(\frac{\mathbf{k}_1 \cdot \mathbf{k}_2}{k_1 k_2} \right)^2 - \left(\frac{\mathbf{k}_2 \cdot \mathbf{k}_3}{k_2 k_3} \right)^2 - \left(\frac{\mathbf{k}_3 \cdot \mathbf{k}_1}{k_3 k_1} \right)^2 \\ + 2 \frac{\mathbf{k}_1 \cdot \mathbf{k}_2 \mathbf{k}_2 \cdot \mathbf{k}_3 \mathbf{k}_3 \cdot \mathbf{k}_1}{k_1^2 k_2^2 k_3^2} \quad (272)$$

We have seen previously that a top-hat filter commutes with β . It can also be shown that

$$\int \frac{d\Omega_1}{4\pi} \frac{d\Omega_2}{4\pi} \frac{d\Omega_3}{4\pi} W(|\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3|R) \eta(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) \\ = \frac{2}{9} W(k_1 R) W(k_2 R) W(k_3 R) \quad (273)$$

Here again, an exact ‘‘commutation property’’ is observed. Successive applications of these geometrical properties³⁵ then imply that [45]

$$j_p \equiv \overline{J_p(\mathbf{k}_1, \dots, \mathbf{k}_p) W(|\mathbf{k}_1 + \dots + \mathbf{k}_p|R)} \quad (274)$$

$$= \overline{J_p(\mathbf{k}_1, \dots, \mathbf{k}_p)} W(k_1 R) \dots W(k_p R), \quad (275)$$

³⁴ We assume $\Omega_m = 1$, but the calculations trivially extend to all cosmologies.

³⁵ This demonstration is incomplete here because the displacement in Lagrangian space is not, in general, potential (see [45] for a more complete demonstration).

where a bar denotes angular-averaged quantities. This is a generalization of the results obtained for parameter S_3 , which has been found to be insensitive to filtering effects in Lagrangian space (for a top-hat filter only).

5.4.3. Lagrangian to Eulerian space mapping: smoothed case

As for the skewness S_3 , a mapping between Lagrangian and Eulerian spaces should permit one to calculate the S_p 's at any order p .

The hierarchy in Eq. (275) gives implicitly the cumulant generating function of the *volume* distribution function for a *fixed mass scale*. One can then make the following remark: the probability that a mass M occupies a volume larger than V is also the probability that a volume V contains a mass lower than M . It suffices for that to consider concentric spheres around a given point \mathbf{x}_0 .³⁶ It is, therefore, possible to relate the real space density PDF to the Lagrangian space one. At this stage however we are only interested in the leading order behavior of the cumulants. We can then notice that, in the small variance limit, the one-point density PDF formally given by Eq. (142), can be calculated by the steepest descent method. The saddle-point position is given by the equation, $d\varphi(y)/dy = \delta$, and in addition we have $d\varphi(y)/dy = \mathcal{G}_\delta(\tau)$, when τ is given implicitly by Eq. (263). The saddle-point position is therefore obtained by a simple change of variable from the linear density τ to the non-linear density contrast δ . It implies that the one-point PDF is roughly given by

$$p(\delta) d\delta \sim \exp\left(-\frac{\tau^2}{2\sigma^2}\right) d\delta \tag{276}$$

with a weakly δ -dependent prefactor. It is important to note that the *leading order* cumulants of this PDF do *not* depend on these prefactors. They are entirely encoded in the τ - δ relation.

As suggested in the previous paragraph, if we now identify $p_E(\delta > \delta_0)$ and $p_L(\delta < \delta_0)$ (one being computed at a fixed real space radius, the other at a fixed mass scale) we obtain a consistency relation

$$-\frac{\tau_E^2}{2\sigma^2(R)} = -\frac{\tau_L^2}{2\sigma^2[(1 + \delta)^{1/3}R]} \tag{277}$$

so that the two have the same leading order cumulants. Here and in the following we use indices L or E for variables that live respectively in Lagrangian or Eulerian space. More precisely, we denote by φ^L the cumulant generating function in Lagrangian space and \mathcal{G}_δ^L the corresponding vertex generating function. In Eulerian space we use the E superscript.³⁷ In the previous equation, the density contrast is a parameter given a priori. The variables τ_E and τ_L depend formally on δ through the saddle-point equations,

$$\delta = \mathcal{G}_\delta^L(\tau_L) = \mathcal{G}_\delta^E(\tau_E), \tag{278}$$

and in Lagrangian space σ is taken at the mass scale corresponding to the density contrast δ (σ is computed a priori in Eulerian space).

³⁶ This statement is however rigorous for centered probabilities only.

³⁷ It is always possible to assume that there exists a function \mathcal{G}_δ^E associated to φ^E , even if there is no associated diagrammatic representation, assuming the same formal functional relation between them.

From these equations we can eliminate τ_L to get an implicit equation between \mathcal{G}_δ^E and τ_E ,

$$\mathcal{G}_\delta^E(\tau_E) = \mathcal{G}_\delta^L \left(\frac{\sigma[(1 + \mathcal{G}_\delta^E(\tau_E))^{1/3} R]}{\sigma(R)} \tau_E \right), \quad (279)$$

where $\mathcal{G}_\delta^L(\tau_L)$ is known and is obtained from spherical collapse dynamics. The cumulant generating function, $\varphi^E(y)$, is then built from $\mathcal{G}_\delta^E(\tau_E)$ the same way as $\varphi^L(y)$ was from $\mathcal{G}_\delta^L(\tau_L)$ [Eqs. (262) and (263)].

Expanding this function around $y = 0$ leads to explicit expressions for the first few values of S_p . They can be written as functions of successive logarithmic derivatives of the variance,

$$\gamma_p \equiv \frac{d^p \log \sigma^2(R)}{d \log^p R}, \quad (280)$$

and read

$$S_3 = \frac{34}{7} + \gamma_1, \quad (281)$$

$$S_4 = \frac{60\,712}{1323} + \frac{62}{3} \gamma_1 + \frac{7}{3} \gamma_1^2 + \frac{2}{3} \gamma_2, \quad (282)$$

$$S_5 = \frac{200\,575\,880}{305\,613} + \frac{1\,847\,200}{3969} \gamma_1 + \frac{6940}{63} \gamma_1^2 + \frac{235}{27} \gamma_1^3 \\ + \frac{1490}{63} \gamma_2 + \frac{50}{9} \gamma_1 \gamma_2 + \frac{10}{27} \gamma_3, \quad (283)$$

$$S_6 = \frac{351\,903\,409\,720}{27810783} + \frac{3\,769\,596\,070}{305613} \gamma_1 + \frac{17\,907\,475}{3969} \gamma_1^2 \\ + \frac{138\,730}{189} \gamma_1^3 + \frac{1210}{27} \gamma_1^4 + \frac{3\,078\,965}{3969} \gamma_2 + \frac{23\,680}{63} \gamma_1 \gamma_2 \\ + \frac{410}{9} \gamma_1^2 \gamma_2 + \frac{35}{9} \gamma_2^2 + \frac{3790}{189} \gamma_3 + \frac{130}{27} \gamma_1 \gamma_3 + \frac{5}{27} \gamma_4 \\ \vdots \quad (284)$$

For a power-law spectrum, these coefficients depend only on spectral index n , through $\gamma_1 = -(n+3)$ and $\gamma_i = 0$ for $i \geq 2$. They are plotted as functions of n in Fig. 26. They all appear to be decreasing functions of n . The above predictions were compared against numerical experiments, as illustrated in Fig. 27 for CDM. The agreement between theory and measurements is close to perfect as long as the variance is below unity. It is quite remarkable to see that the validity domain of PT results does not deteriorate significantly when the cumulant order increases.

5.5. One-loop corrections to one-point moments

We now consider results that include the dependence of S_p parameters on the variance. Due to the complexity of these calculations, only few exact results are known, but there are useful approximate results from the spherical collapse model.

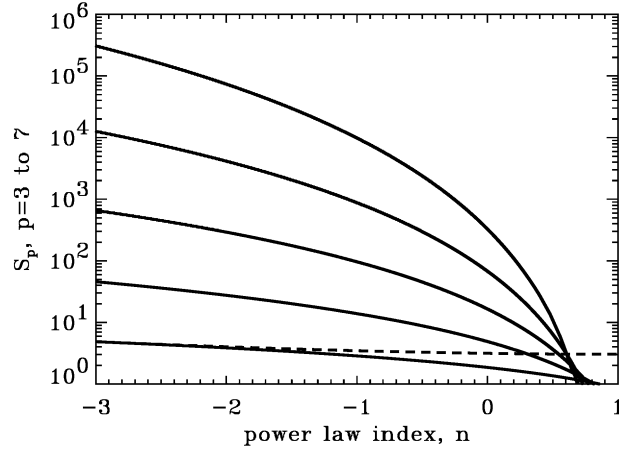


Fig. 26. The predicted S_p parameters for power-law spectra as functions of the spectral index. The results are shown for top-hat filter except for the dashed line which corresponds to the skewness for a Gaussian filter.

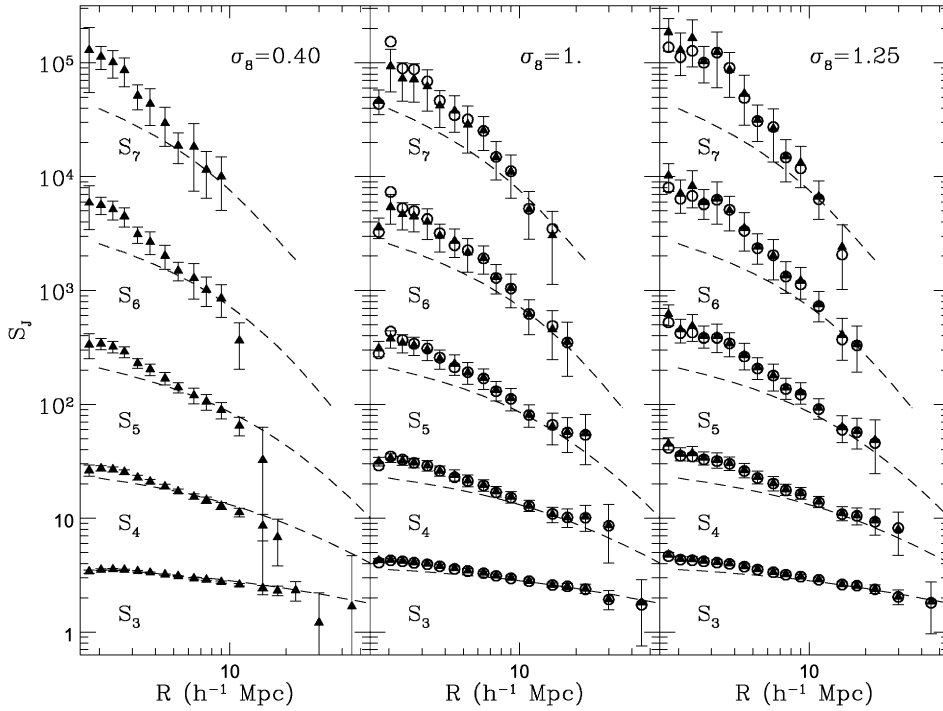


Fig. 27. The S_p parameters for $3 \leq p \leq 7$. Comparisons between theoretical predictions and results from numerical simulations (from [28]) (σ_8 is the linear variance in a sphere of radius $8h^{-1}$ Mpc).

5.5.1. Exact results

To get loop corrections for the one-point density moments, it is necessary to expand both the second moment and the higher-order moments with respect to the linear variance σ_L ,

Table 6
Tree-level and one-loop corrections predicted by various non-linear approximations

Moment expansions	$s_{2,4}$	$S_{3,0}$	$S_{3,2}$	$S_{4,0}$	$S_{4,2}$
FFA, unsmoothed	0.43	3	1	16	15.0
LPA, unsmoothed	0.72	3.40	2.12	21.22	37.12
ZA, unsmoothed	1.27	4	4.69	30.22	98.51
Exact PT, unsmoothed	1.82	4.86	9.80	45.89	—
Exact PT, top-hat smoothing, $n = -2$	0.88	3.86	3.18	27.56	—
Exact PT, Gaussian smoothing, $n = -2$	0.88	4.02	3.83	30.4	—

$$\sigma^2 = \sigma_L^2 + \sum_{n=3}^{\infty} s_{2,n} \sigma_L^n \quad (285)$$

and

$$S_p(\sigma_L) = S_{p,0} + \sum_{n=1}^{\infty} S_{p,n}^L \sigma_L^n . \quad (286)$$

Note that for Gaussian initial conditions, the contributions with n odd vanish. The S_p parameters can also be expanded with respect to the non-linear variance,

$$S_p(\sigma) = S_{p,0} + \sum_{n=1}^{\infty} S_{p,n} \sigma^n , \quad (287)$$

and it is easy to see that $S_{p,2} = S_{p,2}^L$, $S_{p,4} = S_{p,4}^L - S_{p,2}^L s_{2,4}$, etc. for Gaussian initial conditions. Table 6 shows the results of one-loop corrections in various approximations to the dynamics described in Section 2.8 (frozen flow approximation, FFA; linear potential approximation, LPA; and ZA), and exact PT [557]. These results, however, ignore the effects of smoothing which, as is known from tree-level results, are significant.

Taking into account smoothing effects in the exact PT framework has only been done numerically for the case $n = -2$, where the one-loop bispectrum yields a closed form [559]. The resulting one-loop coefficients are shown in Table 6 as well, for top-hat and Gaussian smoothing. When $n \geq -1$, one-loop corrections to S_3 diverge, as for the power spectrum and bispectrum.

5.5.2. The spherical collapse model approximation

Given the complexity of loop calculations, approximate expressions have been looked for. The so-called spherical collapse (SC) model prescription [222] provides a nice and elegant way for getting *approximate* loop corrections for the local cumulants.³⁸

This model consists in assuming that shear contributions in the equations of motion in Lagrangian space can be neglected, which implies that density fluctuations grow locally according to spherical collapse dynamics. In this case, the cumulants can be obtained by a simple non-linear transformation

³⁸ Another prescription, which turns out to be not as accurate, is given in [534].

Table 7

Values for the higher-order perturbative contributions in the SC model for the unsmoothed ($n = -3$) and smoothed ($n = -2, -1, 0$) density fields, for a top-hat filter and a power-law power spectrum. When known exact one-loop results are quoted in brackets. More details can be found in [222]

SC	Unsmoothed		Smoothed	
	$n = -3$	$n = -2$	$n = -1$	$n = 0$
$s_{2,4}$	1.44 [1.82]	0.61 [0.88]	0.40 [∞]	0.79 [∞]
$s_{2,6}$	3.21	0.34	0.05	0.68
$S_{3,0}$	4.86	3.86	2.86	1.86
$S_{3,2}^L$	10.08 [9.80]	3.21 [3.18]	0.59 [∞]	-0.02 [∞]
$S_{3,4}^L$	47.94	3.80	0.07	0.06
$S_{4,0}$	45.89	27.56	13.89	4.89
$S_{4,2}^L$	267.72	63.56	7.39	-0.16
$S_{4,4}^L$	2037.2	138.43	1.99	0.31

of the local Lagrangian density contrast δ ,

$$\delta = (1 + \mathcal{G}_\delta(-\delta_{\text{lin}})) \langle [1 + \mathcal{G}_\delta(-\delta_{\text{lin}})]^{-1} \rangle_L - 1, \quad (288)$$

expressed in terms of the linear density contrast δ_{lin} assumed to obey Gaussian statistics. Note that the ensemble average in Eq. (288) is computed in Lagrangian space.³⁹ Given the fact that the usual ensemble average in Eulerian space is related to the Lagrangian one through $\langle X \rangle_L \equiv \langle (1 + \delta)X \rangle$, the normalization factor $\langle [1 + \mathcal{G}_\delta(\delta_{\text{lin}})]^{-1} \rangle_L$ is required to obey the constraint that $\langle (1 + \delta)^{-1} \rangle_L = \langle 1 \rangle_E = 1$.

For Gaussian initial conditions, the SC model reproduces the tree-level results. Its interest comes from the fact that estimates of loop corrections can be obtained by pursuing relatively simple calculations to the required order. In addition, as we shall see in the next section, it allows a straightforward extension to non-Gaussian initial conditions. The smoothing effects, as shown from calculations exact up to tree level, introduce further complications but can be taken into account by simply changing the vertex generating function \mathcal{G}_δ in Eq. (288) to the one found in Eq. (279). Rigorously, this equation is valid only at tree level: its extension to loop corrections in the SC model can hardly be justified,⁴⁰ but turns out to be a good approximation.

When comparisons are possible, the SC model is seen to provide predictions that are in good agreement with exact PT results (see Table 7), in particular for the S_p parameters. Indeed, for the variance (or cumulants), the SC prescription does not work as well (see e.g. Fig. 28). The reason for this is tidal contributions, which are neglected in the SC approximation and lead to the previrialization effects discussed for the exact PT case in Section 4.2.1. Tidal effects tend to cancel for S_p because of the ratios of cumulants involved. In the SC prescription no divergences are found for $n \geq -1$, thus the interpretation of those remains unresolved.

When tested against numerical simulations, the SC model provides a good account of the departure from tree-level results as illustrated in Fig. 28 for CDM models (see also Fig. 37).

³⁹ Which means that all matter elements are equally weighted, instead of volume elements.

⁴⁰ In the SC model, the kernels in the Jacobian of the mapping from Lagrangian to Eulerian space present no angular dependence, and this is actually incompatible with the commutation property in Eq. (275).

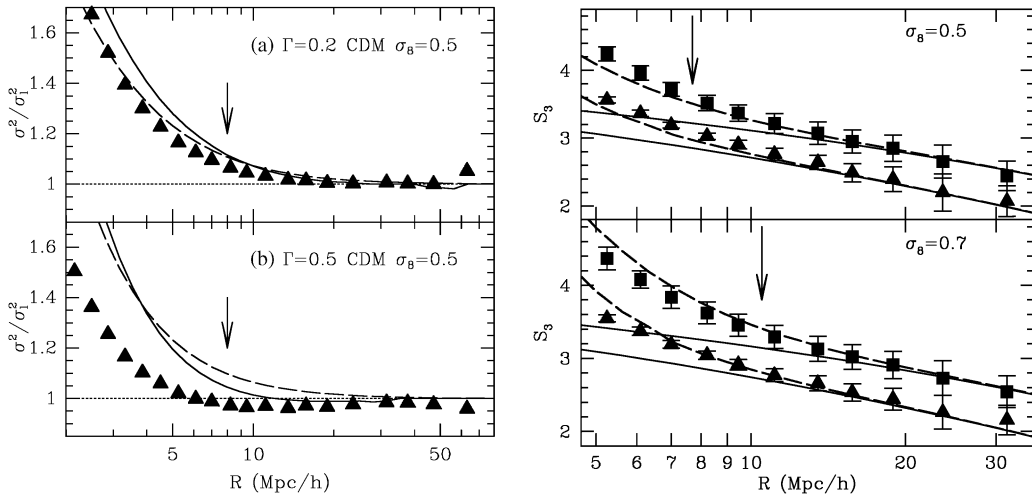


Fig. 28. Non-linear evolution of the variance (left panels) and of the skewness parameter S_3 (right panels) from 10 realizations of flat CDM N -body simulations. Two models are considered, Λ CDM with $\Omega_m + \Omega_\Lambda = 1$ and $\Gamma = 0.2$, and SCDM with $\Omega_m = 1$ and $\Gamma = 0.5$, where Γ is the shape parameter of the power spectrum [201]. In the left panels, symbols show the ratio of the non-linear to the linear variance as a function of radius. The value of Γ is indicated in the panels, while σ_8^2 stands for the linear variance in a sphere of radius $8h^{-1}$ Mpc. The SC model predictions are shown as a short-dashed line while one-loop PT predictions are shown as a solid line. The arrows indicate where $\sigma_l = 0.5$. In the right panels, the output times correspond to $\sigma_8 = 0.5$ (top) and $\sigma_8 = 0.7$ (bottom). Squares and triangles correspond to measurements in $\Gamma = 0.2$ and 0.5 simulations, respectively. Each case is compared to the corresponding PT tree-level predictions (solid lines) and SC model (long-dashed). From [222].

5.6. Evolution from non-Gaussian initial conditions

We now discuss the effects of non-Gaussian initial conditions on the evolution of smoothed moments of the density field. As pointed out in Section 4.4, this is a complicated subject due to the infinite number of possible non-Gaussian initial conditions. For this reason, there are few general results, and only some particular models have been worked out in detail. Early work concentrated on numerical simulation studies [464,684,139] of models with positive and negative primordial skewness and comparison with observations. In addition, a number of studies considered the evolution of higher-order moments from non-Gaussian initial conditions given by cosmic strings [146,9] and texture models [252] using numerical simulations. Recently, measurements of higher-order moments in numerical simulations with χ_N^2 initial conditions with N degrees of freedom were given in [689].

General properties of one-point moments evolved from non-Gaussian initial conditions were considered using PT in [238,333,124,255,195]. To illustrate the main ideas, let us write the PT expression for the first one-point moments:

$$\langle \delta^2 \rangle = \langle \delta_1^2 \rangle + [2\langle \delta_1 \delta_2 \rangle] + \langle \delta_2^2 \rangle + 2\langle \delta_1 \delta_3 \rangle + \mathcal{O}(\sigma^5), \quad (289)$$

$$\langle \delta^3 \rangle = [\langle \delta_1^3 \rangle] + 3\langle \delta_1^2 \delta_2 \rangle + [3\langle \delta_2^2 \delta_1 \rangle + 3\langle \delta_1^2 \delta_3 \rangle] + \mathcal{O}(\sigma^6), \quad (290)$$

$$\langle \delta^4 \rangle = \langle \delta_1^4 \rangle + [4\langle \delta_1^3 \delta_2 \rangle] + 6\langle \delta_1^2 \delta_2^2 \rangle + 4\langle \delta_1^3 \delta_3 \rangle + \mathcal{O}(\sigma^7), \quad (291)$$

where we simply use the PT expansion $\delta = \delta_1 + \delta_2 + \dots$. Square brackets denote terms which scale as odd powers of δ_1 , and thus vanish for Gaussian initial conditions. A first general remark one can make is that these additional terms give to non-Gaussian initial conditions a different *scaling* than for the Gaussian case [238,124]. In addition, the other terms in the skewness have contribution from non-Gaussian initial conditions as well; this does not modify the scaling of these terms but it can significantly change the amplitude. When dealing with non-Gaussian initial conditions, the time dependence and scale dependence must be considered separately. To illustrate this, consider the evolution of the S_p parameters as a function of smoothing scale R and redshift z , assuming for simplicity $\Omega_m = 1$ [so that the growth factor is $a(z) = (1+z)^{-1}$], at largest scales where linear PT applies we have

$$S_p(R, z) \sim (1+z)^{p-2} S_p^1(R). \quad (292)$$

For dimensional scaling models, where the initial conditions satisfy $S_p^1(R) \sim [\sigma_1(R)]^{2-p}$, this implies $S_p(R, z) \sim [\sigma_1(R, z)]^{2-p}$; that is, the S_p parameters scale as inverse powers of the variance *at all times*. Note, however, that Eq. (292) is more general, it implies that irrespective of scaling considerations, *in non-Gaussian models the S_p parameters should be an increasing function of redshift*; this can be used to constrain primordial non-Gaussianity from observations.⁴¹ However, we caution that, as mentioned in Section 4.4, all these arguments are valid if the non-Gaussian fluctuations were generated at early times, and their sources are not active during structure formation.

At what scale does the approximation of linear perturbation theory, Eq. (292), break down? The answer to this question is, of course, significantly model dependent, but it is very important in order to constraint primordial non-Gaussianity. Indeed, we can write the second and third moments from Eqs. (182) and (183):

$$\sigma^2(R) = \sigma_1^2(R) + 2 \int d^3\mathbf{k} W^2(kR) \int d^3\mathbf{q} F_2(\mathbf{k} + \mathbf{q}, -\mathbf{q}) B^1(\mathbf{k}, \mathbf{q}), \quad (293)$$

$$\begin{aligned} \langle \delta^3(R) \rangle &= \langle \delta_1^3(R) \rangle + \langle \delta_G^3(R) \rangle + \int d^3\mathbf{k}_1 \int d^3\mathbf{k}_2 W(k_1R) W(k_2R) W(k_{12}R) \\ &\quad \times \int d^3\mathbf{q} F_2(\mathbf{k}_1 + \mathbf{k}_2 - \mathbf{q}, \mathbf{q}) P_4^1(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_1 + \mathbf{k}_2 - \mathbf{q}, \mathbf{q}), \end{aligned} \quad (294)$$

where $k_{12} \equiv |\mathbf{k}_1 + \mathbf{k}_2|$, B^1 and P_4^1 denote the initial bispectrum and trispectrum, respectively, and the subscript ‘‘G’’ denotes the usual contribution to the third moment due to gravity from Gaussian initial conditions. Therefore, as discussed in Section 4.4 for the bispectrum, corrections to the linear evolution of S_3 depend on the relative magnitude of the initial bispectrum and trispectrum compared to the usual gravitationally induced skewness.

This model dependence can be parametrized in a very useful way under the additional assumption of spherical symmetry. In the spherical collapse model, it is possible to work out entirely the perturbation expansion for one-point moments from non-Gaussian initial conditions, but the solutions are not exact as discussed further below.⁴² Consider non-Gaussian initial conditions with dimensional

⁴¹ Such a method is potentially extremely powerful, as galaxy biasing would tend if anything to actually *decrease* the S_p parameters with z , as bias tends to become larger in the past, see e.g. [635] and discussion in Section 8.

⁴² Some additional results have been recently obtained for the PDF from specific type of non-Gaussian initial conditions, see [662].

scaling. To take into account non-Gaussian terms, one has to rewrite Eq. (286) as

$$S_p(\sigma_L) = \sum_{n=-p+2}^{-1} S_{p,n}^L \sigma_L^n + S_{p,0} + \sum_{n=1}^{\infty} S_{p,n}^L \sigma_L^n, \quad (295)$$

where $\sigma_L = \sigma_I$ is given by linear theory as in Eq. (293). The first non-vanishing perturbative contributions to the variance, skewness and kurtosis read [255]

$$\begin{aligned} s_{2,3} &= [\tfrac{1}{3}S_3^G - 1]B_3, \\ s_{2,4} &= 3 - \tfrac{4}{3}S_3^G - \tfrac{5}{18}(S_3^G)^2 + \tfrac{1}{4}S_4^G + [1 - \tfrac{1}{2}S_3^G - \tfrac{1}{12}(S_3^G)^2 + \tfrac{1}{12}S_4^G]B_4, \\ S_{3,-1}^L &\equiv S_3^{(0)} = B_3, \\ S_{3,0} &= S_3^G - 2[\tfrac{1}{3}S_3^G - 1]B_3^2 + [\tfrac{1}{2}S_3^G - 1]B_4, \\ S_{3,1}^L &= [\tfrac{1}{6}S_3^G - \tfrac{17}{18}(S_3^G)^2 + \tfrac{5}{8}S_4^G]B_3 + [3 - 2S_3^G + \tfrac{1}{3}(S_3^G)^2]B_3^3 \\ &\quad + [-4 + \tfrac{8}{3}S_3^G - \tfrac{1}{6}(S_3^G)^2 - \tfrac{1}{6}S_4^G]B_3B_4 \\ &\quad + [1 - \tfrac{2}{3}S_3^G - \tfrac{1}{12}(S_3^G)^2 + \tfrac{1}{8}S_4^G]B_5, \\ S_{4,-2} &\equiv S_4^{(0)} = B_4 \\ S_{4,-1}^L &= 4S_3^G B_3 + [3 - S_3^G]B_3B_4 + [\tfrac{2}{3}S_3^G - 1]B_5, \\ S_{4,0} &= S_4^G + [3 + 7S_3^G - \tfrac{14}{3}(S_3^G)^2 + \tfrac{3}{2}S_4^G]B_3^2 \\ &\quad + [-1 - \tfrac{10}{3}S_3^G + \tfrac{1}{6}(S_3^G)^2 + \tfrac{5}{4}S_4^G]B_4 \\ &\quad + [6 - 4S_3^G + \tfrac{2}{3}(S_3^G)^2]B_3^2B_4 + [-3 + \tfrac{3}{2}S_3^G + \tfrac{1}{4}(S_3^G)^2 - \tfrac{1}{4}S_4^G]B_4^2 \\ &\quad + [-3 + 3S_3^G - \tfrac{2}{3}(S_3^G)^2]B_3B_5 + [1 - \tfrac{5}{6}S_3^G - \tfrac{1}{18}(S_3^G)^2 + \tfrac{1}{6}S_4^G]B_6. \end{aligned} \quad (296)$$

Here the non-Gaussianity in the initial conditions is characterized via the dimensionless scaling amplitudes

$$B_p \equiv \frac{\langle \delta_I^p \rangle_c}{\sigma_I^p}. \quad (297)$$

For non-Gaussian initial conditions seeded by topological defects such as textures [655,252] or cosmic strings [146,9], B_p is expected to be of order unity.⁴³ For reference, Table 8 lists these results for

⁴³ For cosmic strings, this statement is valid if the scale considered is sufficiently large, $R \gtrsim 1.5(\Omega_m h^2)^{-1}$ Mpc, see [9] for details.

Table 8

Values of the higher-order perturbative contributions in the SC model from non-Gaussian initial conditions with $B_J = 1$ for the unsmoothed ($n = -3$) and smoothed ($n = -2, -1, 0$) density fields for a top-hat window and a power-law spectrum

SC $B_J = 1$	Unsmoothed		Smoothed	
	$n = -3$	$n = -2$	$n = -1$	$n = 0$
$s_{2,3}$	0.62	0.29	-0.05	-0.38
$s_{2,4}$	1.87	0.74	0.44	0.98
$s_{2,5}$	3.36	0.60	-0.05	-1.05
$S_{3,0}$	5.05	4.21	3.38	2.55
$S_{3,1}^L$	7.26	3.91	1.55	0.19
$S_{3,2}^L$	23.53	7.37	1.18	0.20
$S_{4,-1}^L$	19.81	16.14	12.48	8.81
$S_{4,0}$	85.88	52.84	28.31	12.27
$S_{4,1}^L$	332.51	128.51	32.83	2.70

$B_p = 1$ and power-law initial spectra as a function of spectral index n . In this case it is clear that non-linear corrections to the linear result, Eq. (292), can be very important even at large scales. Even more so, χ^2 initial conditions (with spectral index such that it reproduces observations) have $B_3 \approx 2.5$ and $B_4 \approx 10$ [514,689]; therefore, non-linear corrections are particularly strong [255,565].

When compared to exact PT calculations or to measurements in numerical simulations, the SC model is seen to provide quite accurate predictions. This is illustrated in Fig. 29 for the skewness and kurtosis in texture models [255]. These parameters evolve slowly from non-Gaussian initial conditions toward the (Gaussian) gravitational predictions. However, even at present time, a systematic shift can be observed in Fig. 29 between the Gaussian and the non-Gaussian case, well described by the SC predictions taken at appropriate order. The main signature of non-Gaussianity remains at the largest scales, where the S_p parameters show a sharp increase: this is the scaling regime of Eq. (292) where observations can best constrain non-Gaussianity [594,195]. This is explicitly illustrated in Section 8.

5.7. Transients from initial conditions

The standard procedure in numerical simulations is to set up the initial perturbations, assumed to be Gaussian, by using the ZA [705]. This gives a useful prescription to perturb the positions of particles from some initial homogeneous pattern (commonly a grid or a “glass” [688]) and assign them velocities according to the growing mode in linear perturbation theory. In this way, one can generate fluctuations with any desired power spectrum and then numerically evolve them forward in time to the present epoch.

Although the ZA correctly reproduces the linear growing modes of density and velocity perturbations, non-linear correlations are known to be inaccurate when compared to the exact dynamics [274,355,46,116,356], see also Table 7. This implies that it may take a non-negligible amount of time for the exact dynamics to establish the correct statistical properties of density and velocity fields. This transient behavior affects in greater extent statistical quantities which are sensitive to

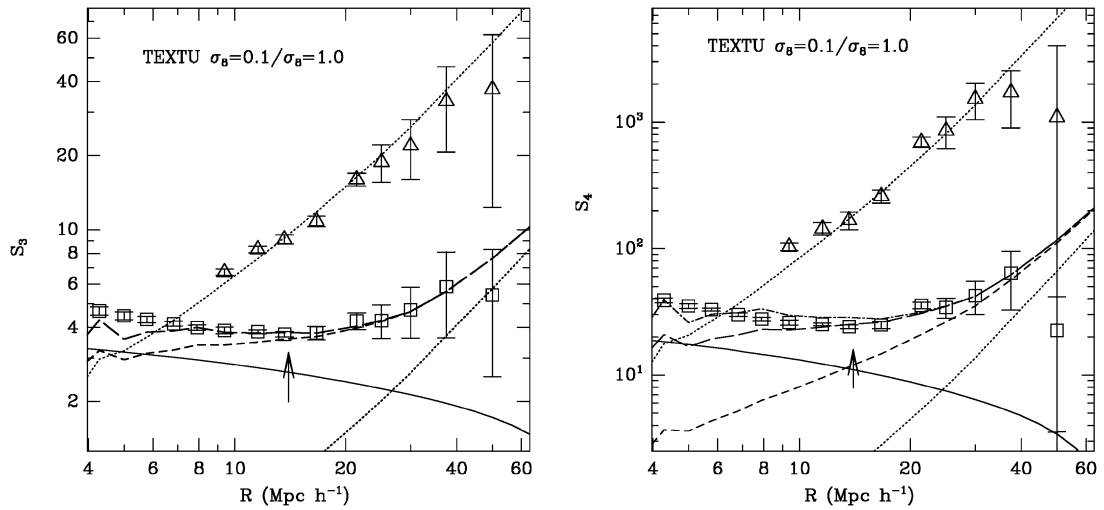


Fig. 29. The skewness and kurtosis, S_3 and S_4 , for texture-like non-Gaussian models. The triangles show the initial conditions ($\sigma_8 = 0.1$), which are fitted well by the dimensional scaling, $S_3 = B_3/\sigma$ and $S_4 = B_4/\sigma^2$ with $B_3 = B_4 \simeq 0.5$, shown as the upper dotted line. Squares show S_3 and S_4 for a later output corresponding to $\sigma_8 = 1.0$. The SC predictions for the $\sigma_8 = 1$ output are shown as short-dashed (including the second-order contribution) and long-dashed line (including the third-order). The continuous line shows the corresponding tree-level PT prediction for Gaussian initial conditions. The lower dotted lines correspond to the linear theory prediction. In the right panel the dot long-dashed line displays the SC prediction including the fourth-order perturbative contribution. From [255].

phase correlations of density and velocity fields; by contrast, the two-point function, variance and power spectrum of density fluctuations at large scales can be described by linear perturbation theory, and are thus unaffected by the incorrect higher-order correlations imposed by the initial conditions.

In Section 2.4.6 we presented the solutions involving the full time dependence from arbitrary initial conditions [561]. Again, we assume $\Omega_m = 1$ for simplicity. The recursion relations for PT kernels including transients result from using the following ansatz in Eq. (86).

$$\Psi_a^{(n)}(\mathbf{k}, z) = \int d^3\mathbf{k}_1 \cdots \int d^3\mathbf{k}_n [\delta_D]_n \mathcal{F}_a^{(n)}(\mathbf{k}_1, \dots, \mathbf{k}_n; z) \delta_1(\mathbf{k}_1) \cdots \delta_1(\mathbf{k}_n), \quad (298)$$

where $a = 1, 2$, $z \equiv \ln a(\tau)$ with $a(\tau)$ the scale factor, and the n th-order solutions for density and velocity fields are components of the vector Ψ_b , i.e. $\Psi_1^{(n)} \equiv \delta_n$, $\Psi_2^{(n)} \equiv \theta_n$. In Eq. (298), $[\delta_D]_n \equiv \delta_D(\mathbf{k} - \mathbf{k}_1 - \cdots - \mathbf{k}_n)$.

The kernels $\mathcal{F}_a^{(n)}$ now depend on time and reduce to the standard ones when transients die out, that is $\mathcal{F}_1^{(n)} \rightarrow F_n$, $\mathcal{F}_2^{(n)} \rightarrow G_n$ when $z \rightarrow \infty$. Also, Eq. (298) incorporates in a convenient way initial conditions, i.e. at $z = 0$, $\mathcal{F}_a^{(n)} = \mathcal{I}_a^{(n)}$, where the kernels $\mathcal{I}_a^{(n)}$ describe the initial correlations imposed at the start of the simulation. For the ZA we have

$$\mathcal{I}_1^{(n)} = F_n^{\text{ZA}}, \quad \mathcal{I}_2^{(n)} = G_n^{\text{ZA}}. \quad (299)$$

Although most existing initial conditions codes use the ZA prescription to set up their initial conditions, there is another prescription to set initial velocities suggested in [199], which avoids the high

initial velocities that result from the use of ZA because of small-scale density fluctuations approaching unity when starting a simulation at low redshifts. This procedure corresponds to recalculating the velocities from the gravitational potential due to the perturbed particle positions, obtained by solving again Poisson equation after particles have been displaced according to the ZA. Linear PT is then applied to the density field to obtain the velocities, which implies instead that the initial velocity field is such that the divergence field $\Theta(x) \equiv \theta(\mathbf{x})/(-f\mathcal{H})$ has the same higher-order correlations as the ZA density perturbations. In this case

$$\mathcal{F}_1^{(n)} = F_n^{ZA}, \quad \mathcal{F}_2^{(n)} = F_n^{ZA}. \quad (300)$$

The recursion relations for $\mathcal{F}_a^{(n)}$, which solve the non-linear dynamics at arbitrary order in PT, can be obtained by substituting Eq. (298) into Eq. (86), which yields [561]

$$\begin{aligned} \mathcal{F}_a^{(n)}(\mathbf{k}_1, \dots, \mathbf{k}_n; z) &= e^{-nz} g_{ab}(z) \mathcal{F}_b^{(n)}(\mathbf{k}_1, \dots, \mathbf{k}_n) \\ &+ \sum_{m=1}^{n-1} \int_0^z ds e^{n(s-z)} g_{ab}(z-s) \gamma_{bcd}(\mathbf{k}^{(m)}, \mathbf{k}^{(n-m)}) \\ &\times \mathcal{F}_c^{(m)}(\mathbf{k}_1, \dots, \mathbf{k}_m; s) \mathcal{F}_d^{(n-m)}(\mathbf{k}_{m+1}, \dots, \mathbf{k}_n; s), \end{aligned} \quad (301)$$

where we have assumed the summation convention over repeated indices, which run between 1 and 2. Eq. (301) reduces to the standard recursion relations for Gaussian initial conditions ($\mathcal{F}_a^{(n)} = 0$ for $n > 1$) when transients are neglected, i.e. the time dependence of $\mathcal{F}_a^{(n)}$ is neglected and the lower limit of integration is replaced by $s = -\infty$. Also, it is easy to check from Eq. (301) that if $\mathcal{F}_a^{(n)} = (F_n, G_n)$, then $\mathcal{F}_a^{(n)} = (F_n, G_n)$, as it should be. Note that PT kernels in Eq. (301) are no longer a separable function of wave vectors and time.

From the recursion relations given by Eq. (301), it is possible to find the recursion relations for the smoothed vertices v_n and μ_n as functions of scale factor a and smoothing scale R , and therefore infer the values of the cumulants as functions of the γ_p 's [Eq. (280)] as in Section 5.4, but with additional dependence with the scale factor. For the skewness parameters, one finds in the Einstein–de Sitter case

$$S_3(a) = \frac{[4 + \gamma_1]}{a} + \left\{ \frac{34}{7} + \gamma_1 \right\} - \frac{\gamma_1 + \frac{26}{5}}{a} + \frac{12}{35a^{7/2}} \quad (302)$$

$$= \frac{34}{7} + \gamma_1 - \frac{6}{5a} + \frac{12}{35a^{7/2}}, \quad (303)$$

$$T_3(a) = -\frac{[2 + \gamma_1]}{a} - \left\{ \frac{26}{7} + \gamma_1 \right\} + \frac{\gamma_1 + \frac{16}{5}}{a} + \frac{18}{35a^{7/2}} \quad (304)$$

$$= -\frac{26}{7} - \gamma_1 + \frac{6}{5a} + \frac{18}{35a^{7/2}}, \quad (305)$$

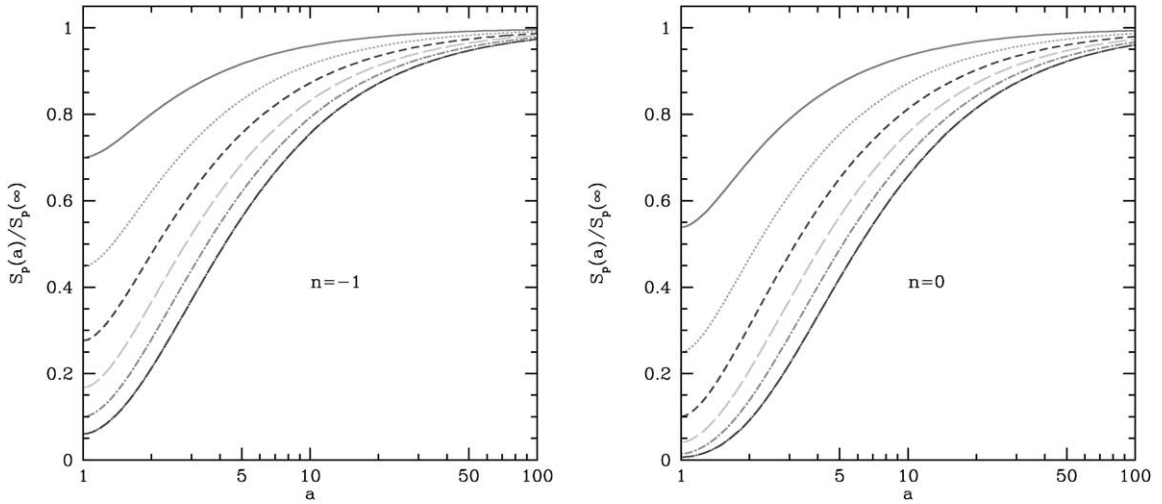


Fig. 30. The ratio of the tree-level S_p parameters at scale factor a to their asymptotic exact dynamics value for scale-free initial spectra with spectral indices $n = -1, 0$. From top to bottom $p = 3, \dots, 8$. The values at $a = 1$ represent those set by the ZA initial conditions.

where we have assumed ZA initial velocities. On the other hand, for initial velocities set from perturbed particle positions, we have

$$S_3(a) = \frac{[4 + \gamma_1]}{a} + \left\{ \frac{34}{7} + \gamma_1 \right\} - \frac{\gamma_1 + \frac{22}{5}}{a} - \frac{16}{35a^{7/2}} \quad (306)$$

$$= \frac{34}{7} + \gamma_1 - \frac{2}{5a} - \frac{16}{35a^{7/2}}, \quad (307)$$

$$T_3(a) = -\frac{[4 + \gamma_1]}{a} - \left\{ \frac{26}{7} + \gamma_1 \right\} + \frac{\gamma_1 + \frac{22}{5}}{a} - \frac{24}{35a^{7/2}} \quad (308)$$

$$= -\frac{26}{7} - \gamma_1 + \frac{2}{5a} - \frac{24}{35a^{7/2}}. \quad (309)$$

For $\Omega_m \neq 1$, these expressions are approximately valid upon replacing the scale factor a by the linear growth factor $D_1(\tau)$. The first term in square brackets in Eqs. (302) and (304) represents the initial skewness given by the ZA (e.g. [46]), which decays with the expansion as a^{-1} , as expected from the discussion on non-Gaussian initial conditions in the previous section. The second and remaining terms in Eqs. (302) and (304) represent the asymptotic exact values (in between braces) and the transient induced by the exact dynamics respectively; their sum vanishes at $a = 1$ where the only correlations are those imposed by the initial conditions. Similar results to these are obtained for higher-order moments, we refer the reader to [561] for explicit expressions. Note that for scale-free initial conditions, the transient contributions to S_p and T_p break self-similarity. Transients turn out to be somewhat less important for velocities set from perturbed particle positions, than in the ZA prescription, as in this case higher-order correlations are closer to those in the exact dynamics.

Fig. 30 illustrates these results for the skewness and higher-order S_p parameters as functions of scale factor a for different spectral indices, assuming that velocities are set as in the ZA. The plots show the ratio of $S_p(a)$ to its “true” asymptotic value predicted by PT, $S_p(\infty)$, for $3 \leq p \leq 8$. The

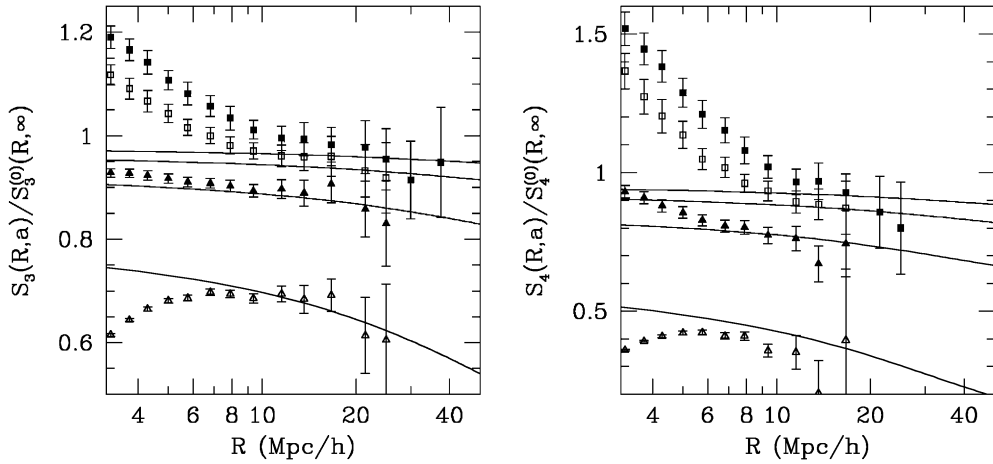


Fig. 31. Symbols show the ratio of the S_p parameters for different scale factor a (simulation began at $a=1$) measured in SCDM numerical simulations [28] to their asymptotic tree-level exact dynamics value as a function of smoothing scale R . Symbols represent $a=1$ (\triangle), $a=1.66$ (\blacktriangle), $a=2.75$ (\square) and $a=4.2$ (\blacksquare). Error bars denote the variance of measurements in 10 realizations. Solid lines correspond to the predictions of transients in tree-level PT, expected to be valid at large scales.

values at $a=1$ correspond to the ratio of ZA to exact dynamics S_p 's, which becomes smaller as either p or n increases. For the skewness, it takes as much as $a=6$ for $n=0$ to achieve 10% of the asymptotic exact PT value, whereas spectra with more large-scale power, where the ZA works better, require less expansion factors to yield the same accuracy. As p increases, however, the transients become worse and at $p=8$ an expansion by a factor $a=40$ is required for $n=0$ to achieve 10% accuracy in S_8 . This suggests that the tails of the PDF could be quite affected by transients from initial conditions.

Fig. 31 presents a comparison of the perturbative predictions for transients in S_p parameters with the standard CDM numerical simulations measurements of [28]. In this case, initial velocities are set as in [199] rather than using the ZA. The error bars in the measurements correspond to the variance over 10 realizations. If there were no transients and no other sources of systematic uncertainties, all the curves would approach unity at large scales, where tree-level PT applies. Unfortunately, there are other sources of systematic uncertainties which prevents a clean test of the transients predictions from PT, as we now briefly discuss, but more details will be given in Section 6.12.

The different symbols correspond to different outputs of the simulation: open triangles denote initial conditions ($a=1$, $\sigma_8=0.24$), solid triangles ($a=1.66$, $\sigma_8=0.40$), open squares ($a=2.75$, $\sigma_8=0.66$) and solid squares ($a=4.2$, $\sigma_8=1.0$). For the initial conditions measurements (open triangles) there is some disagreement with the ZA predictions, especially at small scales, due to discreteness effects, which have not been corrected for. The initial particle arrangement is a grid; therefore, the Poisson model commonly used to correct for discreteness is not necessarily a good approximation (see [28] for further discussion of this point and Section 6.12.2). The second output time (solid triangles) is perhaps the best for testing the predictions of transients: discreteness corrections become much smaller due to evolution away from the initial conditions, and the system has not yet evolved long enough so that finite volume corrections are important (see also Section 6.12.1). For S_3 we see

excellent agreement with the predictions of Eq. (307), with a small excess at small scales due to non-linear evolution away from the tree-level prediction. For $p > 3$ the numerical results show a similar behavior with increased deviation at small scales due to non-linear evolution, as expected. For the last two outputs we see a further increase of non-linear effects at small scales, then a reasonable agreement with the transients predictions, and lastly a decrease of the numerical results compared to the PT predictions at large scales due to finite volume effects, which increase with σ_8 , R and p [147,28,150,472].

5.8. The density PDF

Up to now, we have given exhaustive results on the local density moments. In the following we show how these results can be used to reconstruct the one-point density PDFs [44].

5.8.1. Reconstruction of the PDF from the generating function

We use here the relation between the probability distribution function and the generating function $\varphi(y)$, Eq. (142). To be able to use such a relation one needs a supplementary non-trivial hypothesis. Indeed $\varphi(y)$ is a priori σ dependent through every S_p parameter. We assume here that we have

$$\varphi(y, \sigma) \rightarrow \varphi(y) \quad \text{when } \sigma \rightarrow 0, \quad (310)$$

in a *uniform* way as suggested by numerical simulation results on S_p . No proof has, however, been given of such a property. It has even been challenged recently by calculations presented in [661,663], which suggest that $\varphi(y, \sigma)$ is not analytic at $y \rightarrow 0^-$ for finite values of σ . That would affect results presented below (in particular, the shape of the large density tails). In the following, we will ignore these subtleties and assume that, when the variance is small enough, it is legitimate to compute the density PDF from

$$p(\delta) d\delta = \int_{-i\infty}^{+i\infty} \frac{dy}{2\pi i \sigma^2} \exp \left[-\frac{\varphi(y)}{\sigma^2} + \frac{y\delta}{\sigma^2} \right] d\delta, \quad (311)$$

where $\varphi(y)$ is given by system (262), (263) by analytic continuation from the point $\varphi(0) = 0$.

From this equation numerous results can be obtained. The different forms of $p(\delta)$ have been described in detail in [16,17]. Taking advantage of the approximation $\sigma \ll 1$ one can apply the saddle-point approximation to get

$$p(\delta) d\delta = \frac{d\delta}{-\mathcal{G}'_\delta(\tau)} \left[\frac{1 - \tau \mathcal{G}''_\delta(\tau) / \mathcal{G}'_\delta(\tau)}{2\pi\sigma^2} \right]^{1/2} \exp \left(-\frac{\tau^2}{2\sigma^2} \right), \quad \mathcal{G}_\delta(\tau) = \delta. \quad (312)$$

This solution is valid when $\delta \leq \delta_c$ where δ_c is the value of the density contrast for which $1 = \tau \mathcal{G}''_\delta(\tau) / \mathcal{G}'_\delta(\tau)$. Here function $\mathcal{G}_\delta(\tau)$ is equal to $\mathcal{G}_\delta^L(\tau)$ or $\mathcal{G}_\delta^E(\tau)$ whether one works in Lagrangian space or Eulerian space while taking smoothing into account (Section 5.4.3).

When δ is larger than δ_c , the saddle-point approximation is no longer valid. The shape of $p(\delta)$ is then determined by the behavior of $\varphi(y)$ near its singularity on the real axis,

$$\varphi(y) \simeq \varphi_s + r_s(y - y_s) - a_s(y - y_s)^{3/2}, \quad (313)$$

Table 9

Parameters of the singularity (313) for different values of the spectral index n (there is no singularity for $n \geq 0$)

n	δ_c	y_s	r_s	a_s	φ_s
-3	0.656	-0.184	1.66	1.84	-0.030
-2.5	0.804	-0.213	1.80	2.21	-0.041
-2	1.034	-0.253	2.03	2.81	-0.058
-1.5	1.44	-0.310	2.44	3.93	-0.093
-1	2.344	-0.401	3.34	6.68	-0.172
-0.5	5.632	-0.574	6.63	18.94	-0.434

and we have

$$p(\delta) d\delta = \frac{3 a_s \sigma}{4 \sqrt{\pi}} (1 + \delta - r_s)^{-5/2} \exp[-|y_s| \delta / \sigma^2 + |\varphi_s| / \sigma^2] d\delta . \tag{314}$$

Table 9 gives the parameters describing the singularity corresponding to different values of the spectral index, for the PDF of the smoothed density field in Eulerian space.⁴⁴ One sees that the shape of the cut-off is very different from that of a Gaussian distribution. This shape is due to the analytic properties of the generating function $\varphi(y)$ on the real axis. We explicitly assume here that Eq. (310) is valid, in particular that the position of the first singularity is at finite distance from the origin when σ is finite. It has been pointed out in [663] that Eq. (263) admits a second branch for $y_s < y < 0$ which cannot be ignored in the computation of the density PDF for finite values of σ . In practice its effect is modest. It, however, affects the analytical properties of $\varphi(y)$ and therefore the shape of the large density tail, Eq. (314).

Numerically, it is always possible to integrate Eq. (311) without using the saddle-point approximation. It is then useful to take advantage of the weak Ω_m and Ω_Λ dependence of the vertex generating function. In particular, one can use

$$\mathcal{G}_\delta^L(\tau) = \left(1 + \frac{2\tau}{3}\right)^{-3/2} - 1 , \tag{315}$$

which is the exact result for the spherical collapse dynamics when $\Omega_m \rightarrow 0$, $\Omega_\Lambda = 0$. This leads to slight over-estimation of the low-order vertex [in this case $S_3 = 5 - (n + 3)$, for instance] but the power-law behavior at large τ is correctly reproduced. For this \mathcal{G}_δ^L and for a power-law spectrum, τ can be explicitly written in terms of \mathcal{G}_δ^E . It is interesting to note that for $n=0$ there is no singularity, the saddle-point approximation reduces to Eq. (312) and the Eulerian PDF of the smoothed density field reads

$$p_{n=0}(\delta) d\delta = \sqrt{(1 + \delta)^{-5/3} + (1 + \delta)^{-7/3}} \times \exp \left[-\frac{9((1 + \delta)^{2/3} - 1)^2}{8(1 + \delta)^{1/3} \sigma^2} \right] \frac{d\delta}{\sqrt{\pi} \sigma} . \tag{316}$$

One can also obtain the PDF from the SC model using the local Lagrangian mapping [256,554]. The PDFs that are obtained are in good agreement with the results of numerical simulations. In Fig. 32,

⁴⁴ The case $n = -3$ corresponds as well to the PDF in Lagrangian space or to the unsmoothed case.

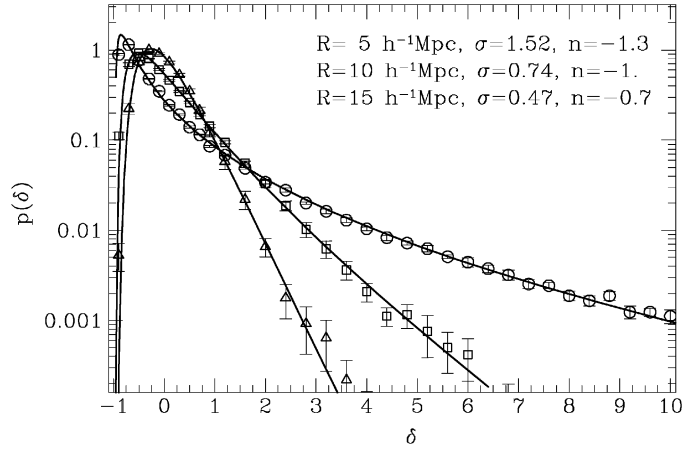


Fig. 32. Comparison between predictions of tree-level PT with results of N -body simulations in the standard CDM model [predictions were calculated assuming Eq. (315)]. From [44].

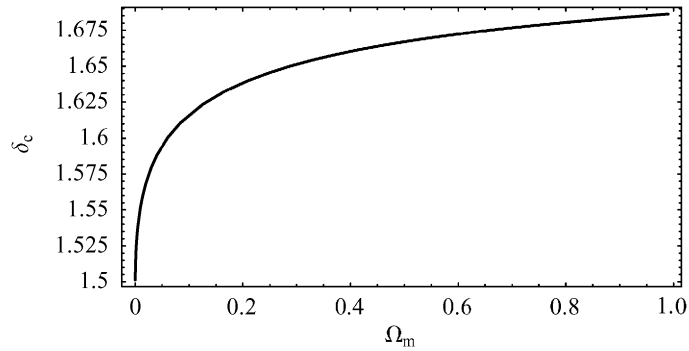


Fig. 33. Variation of the position of the critical (linear) value for the density contrast as a function of Ω_m for open cosmologies.

PT predictions for different smoothing scales are compared to measurements in a P^3M simulation for the standard CDM model. The predicted shape for the PDF (computed from the measured variance and known linear spectral index) is in remarkable agreement with the N -body results.

5.8.2. Dependence on cosmological parameters

The dependence of the shape of the PDF on cosmological parameters is entirely contained in the spherical collapse dynamics when the density field is expressed in terms of the linear density contrast. It can be examined, for instance, in terms of the position of the critical density contrast, δ_c . The variation of δ_c with cosmology is rather modest as shown in Fig. 33 for $\Omega_\Lambda = 0$. This result applies also to the overall shape of \mathcal{G}_δ (see [44,45]), for which the dependence on cosmological parameters remains extremely weak, at percent level. This extends what has been found explicitly for the S_3 and S_4 parameters.

5.8.3. The PDF in the Zel'dovich approximation

For approximate dynamics such as ZA, the previous construction can also be done. It follows exactly the same scheme and the tree-order cumulant generating function can be obtained through the ZA spherical collapse dynamics [471,49].⁴⁵ It is given by

$$\mathcal{G}_\delta^{ZA} = \left(1 - \frac{\tau}{3}\right)^{-3}. \quad (317)$$

One could then compute the Laplace inverse transform of the cumulant generating function to get the one-point density PDF. As in the previous case, this result is not exact in the sense that it is based on the leading order result for the cumulants.

In the case of the ZA, it is actually possible to do an a priori much more accurate calculation with a direct approach. Indeed, the local density contrast neglecting filtering effects is given by the inverse Jacobian of the deformation tensor, Eq. (93), and the joint PDF of the eigenvalues can then be explicitly calculated [190]:

$$p(\lambda_1, \lambda_2, \lambda_3) = \frac{5^{5/2} \times 27}{8\pi\sigma^6} (\lambda_3 - \lambda_1)(\lambda_3 - \lambda_2)(\lambda_2 - \lambda_1) \\ \times \exp \left\{ \left[-3(\lambda_1 + \lambda_2 + \lambda_3)^2 - \frac{15}{2}(\lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3) \right] / \sigma^2 \right\}, \quad (318)$$

where we have assumed that $\lambda_1 < \lambda_2 < \lambda_3$. From this, it is possible to compute the shape of the one-point density PDF [382,49],

$$p(\delta) = \frac{9 \cdot 5^{3/2}}{4\pi N_s (1 + \delta)^3 \sigma^4} \int_{3/(1+\delta)^3}^{\infty} ds e^{-(s-3)^2/2\sigma^2} \\ \times (1 + e^{-6s/\sigma^2})(e^{-\beta_1^2/2\sigma^2} + e^{\beta_2^2/2\sigma^2} - e^{\beta_3^2/2\sigma^2}), \quad (319)$$

$$\beta_n(s) \equiv s^{5/2} \left(\frac{1}{2} + \cos \left[\frac{2}{3}(n-1)\pi + \frac{1}{3} \arccos \left(\frac{54}{s^3(1+\delta)^3} - 1 \right) \right] \right), \quad (320)$$

where N_s is the mean number of streams; $N_s = 1$ in the single-stream regime.

The above prediction for the PDF is however of limited value because, in the absence of smoothing, there is an accumulation of density values at infinity. This is due to the fact that there is always a finite probability of forming caustics (where the Jacobian vanishes). An unfortunate consequence of this is that the moments of this distribution are always infinite! This does not, however, contradict the results given in Section 5.5 as shown in [49]: when a cut-off is applied to the large-density tail, the moments remain finite, and behave as expected from the PT calculations. This has been explicitly verified up to one-loop order [557].

⁴⁵ Extension to other non-linear approximations discussed in Section 2.8 is considered as well in [471]. In addition, recent works have focussed on the PDF generated by second-order PT [644,682]; however, these neglect the effects of smoothing.

5.9. Two-dimensional dynamics

The case of gravitational instability in two spatial dimensions (2D) might be viewed as quite academic. It is, however, worth investigating for different reasons: (i) it is a good illustration of the general method; (ii) numerical simulations in 2D dynamics can be done with a much larger dynamical range than in 3D; and, perhaps most importantly, (iii) the 2D results turn out to be of direct use to study statistical properties of the projected density (Section 7.2), relevant for observations of angular clustering and weak gravitational lensing.

The dynamics we are interested in corresponds actually to density fluctuations embedded in a 3D space but which are uniform along one direction. The general equations of motion are left unchanged; here, we consider again only the Einstein–de Sitter case.

Let us review the different stages of the calculation [48]. For the naked vertices, without smoothing effects, the only change introduced is due to the $\cos^2(\mathbf{k}_1, \mathbf{k}_2)$ factor that in 2D averages to $\frac{1}{2}$ instead of $\frac{1}{3}$. The resulting recursion relations between the vertices v_n and μ_n then read

$$v_n = \sum_{m=1}^{n-1} \binom{n}{m} \frac{\mu_m}{(2n+3)(n-1)} [(2n+1)v_{n-m} + \mu_{n-m}], \quad (321)$$

$$\mu_n = \sum_{m=1}^{n-1} \binom{n}{m} \frac{\mu_m}{(2n+3)(n-1)} [3v_{n-m} + n\mu_{n-m}], \quad (322)$$

instead of Eqs. (50) and (51). No simple solution for the generating function of v_n , $\mathcal{G}_\delta^{2D}(\tau)$, is known although it again corresponds to the equation describing the “spherical” collapse in 2D.⁴⁶ It can, however, be shown that $\mathcal{G}_\delta^{2D}(\tau) - 1 \sim \tau^{-(\sqrt{13}-1)/2}$ when $\tau \rightarrow \infty$, and the expression

$$\mathcal{G}_\delta^{2D}(\tau) = \left(1 + \frac{\tau}{v}\right)^{-v} - 1 \quad \text{with } v = \frac{\sqrt{13}-1}{2} \quad (323)$$

provides a good fit. More precisely one can rigorously calculate the expansion of $\mathcal{G}_\delta(\tau)$ near $\tau = 0$ and it reads

$$\mathcal{G}_\delta^{2D}(\tau) = -\tau + \frac{12}{14} \tau^2 - \frac{29}{42} \tau^3 + \frac{79}{147} \tau^4 - \frac{2085}{5096} \tau^5 + \dots \quad (324)$$

The resulting values for the S_p^{2D} parameters when smoothing is neglected are $S_3^{2D} = \frac{36}{7}$, $S_4^{2D} = \frac{2540}{49}$, $S_5^{2D} = 793$, $S_6^{2D} = 13\,370$. When filtering is taken into account the vertex generating function becomes⁴⁷

$$\mathcal{G}_\delta^E(\tau) = \mathcal{G}_\delta^{2D}(\tau [1 + \mathcal{G}_\delta^E(\tau)]^{-2-n}) \quad (325)$$

for a power-law spectrum of index n . This leads to [48]

$$S_3^{2D} = \frac{36}{7} - \frac{3(n+2)}{2}, \quad (326)$$

$$S_4^{2D} = \frac{2540}{49} - 33(n+2) + \frac{21(n+2)^2}{4}, \quad (327)$$

⁴⁶ To our knowledge there is no closed analytical solution for the 2D spherical collapse.

⁴⁷ In 2D dynamics if $P(k) \sim k^n$, then $\sigma(R) \propto R^{-(n+2)}$.

Table 10

Parameters of the singularity, Eq. (313), for the 2D case. There is no singularity for $n \geq 0$

n	y_s	φ_s	r_s	a_s
-2	-0.172	-0.197	1.60	-1.72
-1.5	-0.212	-0.252	1.81	-2.25
-1	-0.277	-0.350	2.23	-3.41
-0.5	-0.403	-0.581	3.55	-7.73

$$S_5^{2D} = \frac{271\,960}{343} - \frac{38\,900(n+2)}{49} + \frac{3705(n+2)^2}{14} - \frac{235(n+2)^3}{8}, \quad (328)$$

$$S_6^{2D} = \frac{510\,882\,660}{31\,213} - \frac{7\,721\,415(n+2)}{343} + \frac{2\,272\,395(n+2)^2}{196} - \frac{74\,205(n+2)^3}{28} + \frac{1815(n+2)^4}{8}. \quad (329)$$

Obviously, these results can also be obtained from a direct perturbative calculation using the geometrical properties of the 2D top-hat window function given in Appendix C. The position and shape of the singularity are also changed in 2D dynamics. In Table 10 we give the parameters of the singularity in $\varphi(y)$.

5.10. The velocity divergence PDF

So far our description has been focussed on the density field. The structure of the equations for the velocity divergence is the same as for the local density. We briefly account here for the results that have been obtained at tree level for the velocity divergence [44]. Loop corrections with exact PT are discussed in e.g. [557]. Note that the SC model approximation described in Section 5.5.2 does not do as well as for the density contrast, due to tidal contributions,⁴⁸ but can provide again approximate loop corrections for the cumulants while still giving exact tree-level results [223].

5.10.1. The velocity divergence cumulants hierarchy

In what follows, we assume that the velocity divergence is expressed in units of the conformal expansion rate, $\mathcal{H} = aH$. For convenience, we define the vertex generating function for the velocity divergence as

$$\mathcal{G}_v(\tau) \equiv -f(\Omega_m, \Omega_\Lambda) \sum_{p \geq 1} \mu_p \frac{(-\tau)^p}{p!} \equiv \sum_{p \geq 1} \check{\mu}_p \frac{(-\tau)^p}{p!}. \quad (330)$$

This definition corresponds to slightly different vertices from those given by Eq. (49):

$$\check{\mu}_p \equiv \langle \theta^{(n)} [\delta^{(1)}]^p \rangle_c / \langle [\delta^{(1)}]^2 \rangle^p. \quad (331)$$

⁴⁸ Velocities are more affected by previrialization effects, as shown in Fig. 12.

When the filtering effect is not taken into account, the vertex generating function can be obtained from that of the density field. From the continuity equation we have [43]

$$\mathcal{G}_\theta(a, \tau) = - \left[a \frac{d}{da} \mathcal{G}_\delta(a, \tau) + f(\Omega_m, \Omega_\Lambda) \tau \frac{d}{d\tau} \mathcal{G}_\delta(a, \tau) \right] [1 + \mathcal{G}_\delta(a, \tau)]^{-1}. \quad (332)$$

One can use the fact that function $\mathcal{G}_\delta(a, \tau)$ is nearly insensitive to the values of Ω_m and Ω_Λ to obtain a simplified form for the function $\mathcal{G}_\theta(a, \tau)$,

$$\mathcal{G}_\theta(\tau) \approx -f(\Omega_m, \Omega_\Lambda) \tau \frac{d}{d\tau} \mathcal{G}_\delta^L(\tau) / [1 + \mathcal{G}_\delta^L(\tau)], \quad (333)$$

so that $\mathcal{G}_\theta(\tau) \approx f(\Omega_m, \Omega_\Lambda) \tau (1 + 2\tau/3)^{-1}$ if approximation in Eq. (315) is used. This, in fact, fully justifies the definition of the vertices μ_p which are seen to be almost independent of the cosmological parameters, as already discussed in Section 2.4.3.

From now on, we use again for clarity the Lagrangian and Eulerian superscripts, in particular $\mathcal{G}_\theta^L \equiv \mathcal{G}_\theta$, $\mathcal{G}_\delta^L \equiv \mathcal{G}_\delta$. Including filtering effects requires taking into account the mapping from Lagrangian to Eulerian space, as explained in Section 5.4.3. As a consequence of this we have

$$\mathcal{G}_\theta^E(\tau) = \mathcal{G}_\theta^L \left[\tau \frac{\sigma([1 + \mathcal{G}_\delta^E(\tau)]^{1/3} R)}{\sigma(R)} \right], \quad (334)$$

which amounts to say that the velocity divergence should be calculated at the correct mass scale. This remapping does not further complicate the dependence on cosmological parameters: $\mathcal{G}_\theta^E(\tau)/f(\Omega_m, \Omega_\Lambda)$ remains independent of $(\Omega_m, \Omega_\Lambda)$ to a very good accuracy.

It is possible to derive the cumulants T_p from the implicit equation (334), relying on the usual relations given in Section 5.4.1 between the cumulants and what would be the genuine intrinsic velocity divergence vertices, $\mu_p^{\text{intr}} \equiv \langle \theta^{(n)} [\theta^{(1)}]^p \rangle_{c,E} / \langle [\theta^{(1)}]^2 \rangle_E^p$ that are straightforwardly related to $\tilde{\mu}_p^E$ through $\mu_p^{\text{intr}} = \tilde{\mu}_p^E [-f(\Omega_m, \Omega_\Lambda)]^{-p}$. The corresponding vertex generating function, $\mathcal{G}_\theta^{\text{intr}}(\tau)$, is given by $\mathcal{G}_\theta^{\text{intr}}(\tau) = \mathcal{G}_\theta^E[-f(\Omega_m, \Omega_\Lambda)\tau]$ together with Eqs. (260), (226) and (263), and replacing S_p with T_p and \mathcal{G}_δ with $\mathcal{G}_\theta^{\text{intr}}$ can be used to compute the velocity divergence cumulant parameters. For an Einstein–de Sitter universe, the first two read

$$T_3(\Omega_m = 1, \Omega_\Lambda = 0) = - \left(\frac{26}{7} + \gamma_1 \right), \quad (335)$$

$$T_4(\Omega_m = 1, \Omega_\Lambda = 0) = \frac{12\,088}{441} + \frac{338\gamma_1}{21} + \frac{7\gamma_1^2}{3} + \frac{2\gamma_2}{3}, \quad (336)$$

⋮

where the parameters γ_p are given by Eq. (280). Furthermore, the dependence on cosmological parameters is straightforwardly given by⁴⁹

$$T_p(\Omega_m, \Omega_\Lambda) \approx \frac{1}{f(\Omega_m, \Omega_\Lambda)^{(p-2)}} T_p(\Omega_m = 1, \Omega_\Lambda = 0), \quad (337)$$

which implies a relatively strong Ω_m dependence for the shape of $p(\theta)$ as we now discuss.

⁴⁹ To be compared, for example, to the more accurate result given for T_3 in Eq. (245).

5.10.2. The shape of the PDF

The above line of arguments provides a general rule for the dependence of the PDF on cosmological parameters:

$$p[f(\Omega_m, \Omega_\Lambda), \theta, \sigma_\theta] d\theta \approx p \left[1, \frac{\theta}{f(\Omega_m, \Omega_\Lambda)}, \frac{\sigma_\theta}{f(\Omega_m, \Omega_\Lambda)} \right] \frac{d\theta}{f(\Omega_m, \Omega_\Lambda)}. \quad (338)$$

Otherwise, the PDF can be calculated exactly the same way as for the density contrast.

The case $n = -1$ is worth further investigations since it is then possible to derive a closed form that fits extremely well the exact numerical integration, as for the PDF of δ for $n = 0$. This approximation is based on the approximate form in Eq. (315) for the function \mathcal{G}_δ^1 . With $n = -1$ it leads to

$$\mathcal{G}_\delta^{E, n=-1}(\tau) = \left[-\frac{\tau}{3} + \left(1 + \frac{\tau^2}{9} \right) \right]^3 - 1. \quad (339)$$

One can then show that

$$\mathcal{G}_\theta^{E, n=-1}(\tau) = f(\Omega_m, \Omega_\Lambda) \left[\tau \left(1 + \frac{\tau^2}{9} \right)^{1/2} - \frac{\tau^2}{3} \right]. \quad (340)$$

The calculation of the PDF of the velocity divergence from the saddle-point approximation [e.g. Eq. (312)] then leads to the expression

$$p(\theta) d\theta = \frac{([2\kappa - 1]/\kappa^{1/2} + [\lambda - 1]/\lambda^{1/2})^{-3/2} \exp \left[-\frac{\theta^2}{2\lambda\sigma_\theta^2} \right] d\theta}{\kappa^{3/4}(2\pi)^{1/2}\sigma_\theta} \quad (341)$$

with

$$\kappa = 1 + \frac{\theta^2}{9\lambda f(\Omega_m, \Omega_\Lambda)^2}, \quad \lambda = 1 - \frac{2\theta}{3f(\Omega_m, \Omega_\Lambda)}, \quad (342)$$

where θ is expressed in units of the conformal expansion rate, \mathcal{H} .

5.10.3. Comparison with N -body simulations

Measurements in numerical simulations turn out to be much more non-trivial for the velocity field than for the density field. The reason is that in N -body simulations, the density field is traced by a Poisson realization. Although it suffices to count points, in grid cells for instance, to get the filtered density,⁵⁰ the velocity field is only known in a non-uniform way where particles happen to be. Therefore, simple averages of velocities do not lead to good estimations of the statistical properties one is interested in, especially when the number density of particles is small.

For this purpose, specific methods have been developed to deal with velocity field statistics [52]. The idea is to use tessellations to obtain a continuous description of the velocity field; two alternative prescriptions have been proposed. One makes use of the Voronoi tessellation; in this case the velocity is assumed to be uniform within each Voronoi cell, in other words, the local velocity at any space point is the one of the closest particle. The second method makes use of the Delaunay tessellation. In this case the local velocity is assumed to vary linearly within each Delaunay tetrahedron (such ensemble of tetrahedra forms a unique partition of space); the local velocity is then defined by a linear combination of its closest neighbors, see [52,54] for details.

⁵⁰ Corrected for discreteness effects using factorial moments as discussed in Section 6.7.

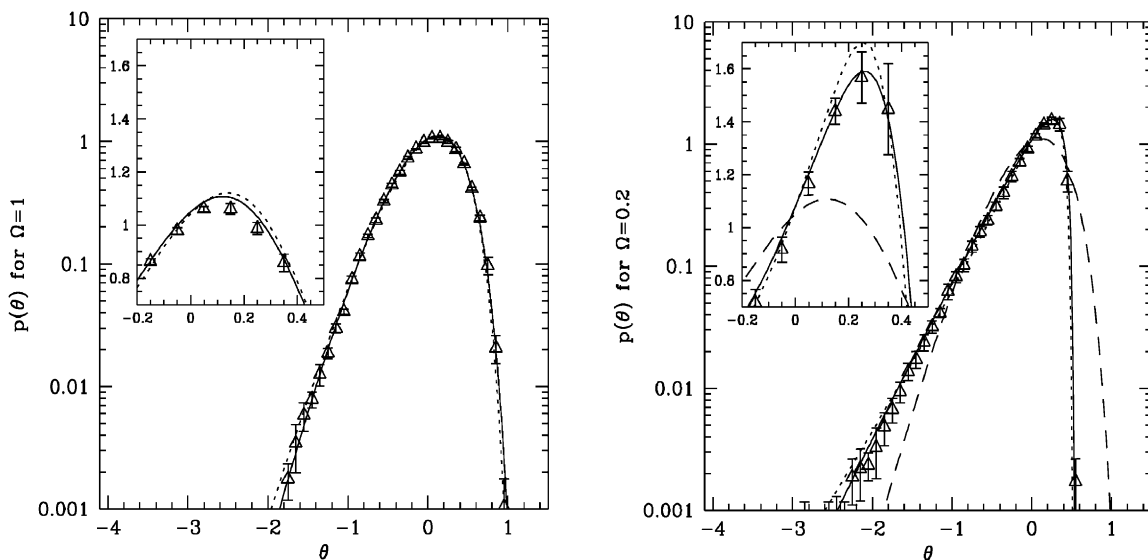


Fig. 34. The PDF of the velocity divergence for two different values of Ω_m ($\Omega_m = 1$, left panel and $\Omega_m = 0.2$, right panel). The dotted lines correspond to the approximate analytic fit [Eq. (341)] and the solid lines to the theoretical predictions obtained from a direct numerical integration of the inverse Laplace transform with $n = -0.7$. In the right panel the dashed line is the prediction for $\Omega_m = 1$ and the same $\sigma_\theta \approx 0.4$. From [54].

These methods have been applied to results of numerical simulations [54,387]. Comparisons between theoretical predictions, in particular form (341), and the measurements are shown in Fig. 34. The simulation used here is a PM simulation with a scale-free spectrum with $n = -1$. The prediction, Eq. (341), gives a good account of the shape of the divergence PDF, especially in the tails. The detailed behavior of the PDF near its maximum requires a more exact computation. We obtained it here by an exact inverse Laplace computation using Eq. (315) for the density vertex generating function [and Eq. (333)] to get the velocity vertices. Because this expression does not accurately predict the low-order cumulants⁵¹ the integration has been made with $n = -0.7$, instead of $n = -1$, to compensate for this problem. The agreement with simulations is quite remarkable.

5.11. The velocity–density relation

PT also allows one to consider multivariate PDFs such as the joint distribution of the local density contrast and the local divergence θ . An example of such PDF is shown in Fig. 35. It illustrates, in particular, the fact that the local density and local divergence do not follow in general a one-to-one correspondence, as it would be the case in linear perturbation theory. Deviations from this regime induce not only a nonlinear relation between δ and θ , i.e. a bending in the δ – θ relation, but also a significant scatter.

In general, the statistical properties of these two fields can be studied through their joint cumulants, $\langle \delta^p \theta^q \rangle_c$. Similar to cases involving only one variable, it is possible to compute such quantities at

⁵¹ For example, $T_3 = 4 - (n + 3)$ instead of $T_3 = \frac{26}{7} - (n + 3)$.

leading order, or at next to leading order (involving loop corrections) in PT. One can define the parameters U_{pq} as

$$\langle \delta^p \theta^q \rangle_c = U_{pq} \langle \delta^2 \rangle^{p+q-1}, \quad (343)$$

where θ is expressed in units of the conformal expansion rate, \mathcal{H} . The U_{pq} 's are finite (and non-zero) at large scales for Gaussian initial conditions and can be easily computed at tree order. Their calculation follows a tree construction from the vertices v_p and μ_q . For instance, one obtains

$$U_{11} = v_1 \check{\mu}_1 = \check{\mu}_1 = -f(\Omega_m, \Omega_\Lambda),$$

$$U_{21} = 2v_2 \check{\mu}_1 + \check{\mu}_2,$$

$$U_{31} = 3v_3 \check{\mu}_1 + \check{\mu}_3 + 6v_2^2 \check{\mu}_1 + 6v_2 \check{\mu}_2,$$

$$U_{22} = 2v_3 \check{\mu}_1^2 + 2\check{\mu}_3 \check{\mu}_1 + 8v_2 \check{\mu}_2 \check{\mu}_1 + 2v_2^2 \check{\mu}_1^2 + 2\check{\mu}_2^2,$$

with $\check{\mu}_p \equiv -f(\Omega_m, \Omega_\Lambda) \mu_p$.

These expressions are straightforward when the smoothing effects are not taken into account. They are still true otherwise, but they rely on the fact that the same mapping applies to the density and the velocity divergence. More generally, it is possible to derive explicitly the generating function of the joint cumulants. The demonstration is presented in Appendix B.2.

An interesting application of these results is the computation of the joint density–velocity PDF. Assuming that the leading order contributions to cumulants provide a reliable description, we have

$$\begin{aligned} p(\delta, \theta) d\delta d\theta &= \int_{-i\infty}^{+i\infty} \frac{dy_1}{2\pi i} \int_{-i\infty}^{+i\infty} \frac{dy_2}{2\pi i} \exp \left[\frac{\delta y_1}{\sigma^2} + \frac{\theta y_2}{\sigma^2} - \frac{\varphi(y_1, y_2)}{\sigma^2} \right], \\ \varphi(y_1, y_2) &= y_1 \mathcal{G}_\delta(\tau) + y_2 \mathcal{G}_\theta(\tau) - \frac{1}{2} y_1 \tau \frac{d}{d\tau} \mathcal{G}_\delta(\tau) - \frac{1}{2} y_2 \tau \frac{d}{d\tau} \mathcal{G}_\theta(\tau), \\ \tau &= -y_1 \frac{d}{d\tau} \mathcal{G}_\delta(\tau) - y_2 \frac{d}{d\tau} \mathcal{G}_\theta(\tau), \end{aligned} \quad (344)$$

where σ^2 is the variance of the *density* field.

As a consequence of this relation one can compute *constrained* averages such as the expectation value of θ under the constraint that the local density is known, $\langle \theta \rangle_\delta$. For a vanishing variance (that is, at tree level) the result turns out to be extremely simple and reads [42]

$$\langle \theta \rangle_\delta = \mathcal{G}_\theta(\tau) \quad \text{with} \quad \mathcal{G}_\delta(\tau) = \delta. \quad (345)$$

This relation can obviously be inverted to get $\langle \delta \rangle_\theta$. It is interesting to note that this result is not quantitatively changed when top-hat smoothing effects are taken into account (nor it depends on the shape of the power spectrum), which is not true anymore with Gaussian smoothing [125].

Table 11

The coefficients a_1, \dots, a_3 and r_1, \dots, r_3 as functions of the spectral index n for scale-free power spectra and Gaussian smoothing. Results are given at leading order, except for a_1 and r_1 for which one-loop corrections are included when available (correction is infinite for $n \geq -1$)

Index n	a_1	a_2	a_3	r_1	r_2	r_3
-3.0	—	0.190	-0.0101	$1 + 0.3\sigma^2$	-0.190	0.0826
-2.5	—	0.192	-0.00935	$1 + 0.202\sigma^2$	-0.192	0.0822
-2.0	$1 - 0.172\sigma_\theta^2$	0.196	-0.00548	$1 + 0.077\sigma^2$	-0.196	0.0821
-1.5	$1 + 0.187\sigma_\theta^2$	0.203	-0.000127	$1 - 0.296\sigma^2$	-0.203	0.0822
-1.0	$1 + [\infty]$	0.213	0.00713	$1 + [\infty]$	-0.213	0.0835
-0.5	$1 + [\infty]$	0.227	0.0165	$1 + [\infty]$	-0.227	0.0865
0	$1 + [\infty]$	0.246	0.0279	$1 + [\infty]$	-0.246	0.0928
0.5	$1 + [\infty]$	0.270	0.0408	$1 + [\infty]$	-0.270	0.1051
1.0	$1 + [\infty]$	0.301	0.0532	$1 + [\infty]$	-0.301	0.1283

A more pedestrian approach should be used when the variance is not negligible:

$$\langle \delta \rangle_\theta = a_0 + a_1 \theta + a_2 \theta^2 + a_3 \theta^3 + \dots, \quad (346)$$

$$\langle \theta \rangle_\delta = r_0 + r_1 \delta + r_2 \delta^2 + r_3 \delta^3 + \dots. \quad (347)$$

Computations should be made order by order and it becomes inevitable to introduce next-to-leading order corrections, i.e. loop corrections.

The coefficients a_0, \dots, a_3 and r_0, \dots, r_3 have been computed explicitly up to third order in PT [125,127,56]. It is to be noted that at leading order one has $a_0 = -a_2 \sigma_\theta^2$ and $r_0 = -r_2 \sigma^2$ to ensure that the global ensemble averages of θ and δ vanish. Note also that the third order PT results for a_1 and r_1 involve a loop correction that diverges for $n \geq -1$. The known results are given in Table 11 for the Einstein–de Sitter case and Gaussian smoothing.

The Ω_m dependence of these coefficients can be explicitly derived. For instance, the coefficient r_2 can be expressed in terms of the skewness of the two fields (at leading order only), which leads to $r_2 = f(\Omega_m, \Omega_A)(S_3 + f(\Omega_m, \Omega_A)T_3)/6$. For a top-hat filter, r_2 is always given by $f(\Omega_m, \Omega_A)4/21$ and, for a Gaussian window it varies slightly with the power-spectrum index but shows a similarly strong $f(\Omega_m, \Omega_A)$ (and therefore Ω_m) dependence. Comparisons with numerical simulations have demonstrated the accuracy and robustness of these predictions (except for the loop terms) [56,387].

Such results are of obvious observational interest since one can, in principle, measure the value of Ω_m from velocity–density comparisons, see [179]. In particular, a detailed analysis of the curvature in the δ – θ relation (through a_2 or r_2) would provide a way to break the degeneracy between biasing parameters (Section 7.1) and Ω_m [128,56].⁵² Moreover, these results can be extended to take into account redshift distortion effects (Section 7.4) as described in [129]. The main practical issue is that current velocity surveys are not sufficiently large to provide accurate density–velocity comparisons going beyond linear PT.

⁵² The scatter in this relation seen in Fig. 35 can be reduced by including also off-diagonal components of the velocity deformation tensor [273,429,126].

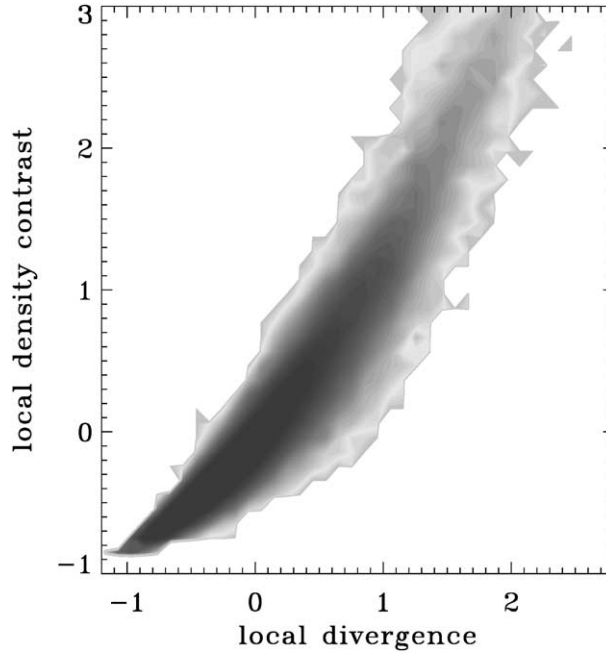


Fig. 35. Example of a joint PDF of the density and the velocity divergence. The color is in logarithmic scale, the smoothing scale is 15 Mpc/h, the spectrum is scale free with $n = -1.5$ and $\sigma_8 \equiv 1$, see [56] for details.

It is finally worth noting that these investigations are also useful for detailed analysis of the Lyman- α forest [483].

5.12. The two-point density PDF

Perturbation theory can obviously be applied to any combination of the density taken at different locations. In particular, for sound cosmic error computations (see Section 6), the bivariate density distribution is an important quantity that has been investigated in some detail.

The object of this section is to present the exact results that have been obtained at tree level for the two-point density cumulants [51]. We consider the joint densities at positions \mathbf{x}_1 and \mathbf{x}_2 and we are interested in computing the cumulants $\langle \delta^p(\mathbf{x}_1) \delta^q(\mathbf{x}_2) \rangle_c$ where the field is supposed to be filtered at a given scale R . In general, such cumulants are expected to have quite complicated expressions, depending on both the smoothing length R and the distance $|\mathbf{x}_1 - \mathbf{x}_2|$. We make here the approximation that the distance between the two points is large compared to the smoothing scale. In other words, we neglect short-distance effects.

Let us define the parameters C_{pq} :

$$C_{pq} = \frac{\langle \delta^p(\mathbf{x}_1) \delta^q(\mathbf{x}_2) \rangle_c}{\langle \delta(\mathbf{x}_1) \delta(\mathbf{x}_2) \rangle \langle \delta^2 \rangle^{p+q-2}}. \quad (348)$$

Because of the tree structure of the correlation hierarchy, we expect the coefficients C_{pq} to be finite in both the large-distance limit and at leading order in the variance. This expresses the fact that

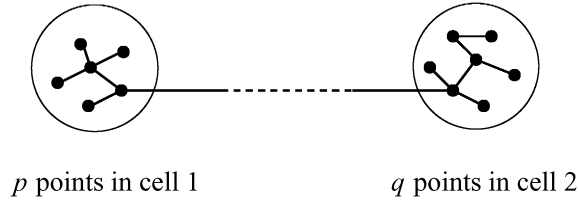


Fig. 36. Structure of the coefficient C_{pq} in large separation limit: C_{pq} is given by the sum of all possible trees joining p points in first cell to q points in the second with only one crossing line. The sums can be done separately on each side leading to $C_{pq} = C_{p1} C_{q1}$.

among all the diagrams that connect the two cells, the ones that involve only one line between the cells are expected to be dominant in cases when $\langle \delta(\mathbf{x}_1)\delta(\mathbf{x}_2) \rangle \ll \langle \delta^2 \rangle$.

The next remarkable property is directly due to the tree structure of the high-order correlation functions. The coefficients C_{pq} are dimensionless quantities that correspond to some geometrical averages of trees. It is quite easy to realize (see Fig. 36) that such averages can be factorized into two parts, corresponding to the end points of the line joining the two cells. In other words, one should have

$$C_{pq} = C_{p1} C_{q1} . \quad (349)$$

This factorization property is specific to tree structures. It was encountered originally in previous work in the fully non-linear regime [40]. It has specific consequences on the behavior of the two-point density PDF, namely we expect that

$$p[\rho(\mathbf{x}_1), \rho(\mathbf{x}_2)] = p[\rho(\mathbf{x}_1)]p[\rho(\mathbf{x}_2)](1 + b[\rho(\mathbf{x}_1)]\langle \delta(\mathbf{x}_1)\delta(\mathbf{x}_2) \rangle b[\rho(\mathbf{x}_2)]) . \quad (350)$$

The joint density PDF is thus entirely determined by the shape of the “bias” function, $b(\rho)$.⁵³

The general computation of the C_{p1} series is not straightforward, although the tree structure of the cumulants is indicative of a solution. Indeed the generating function $\psi(y)$ of C_{p1} ,

$$\psi(y) = \sum_{p=1}^{\infty} C_{p1} \frac{y^p}{p!} , \quad (351)$$

corresponds to the generating function of the diagrams with one external line. For exact trees this would be $\tau(y)$. However, the Lagrangian to Eulerian mapping affects the relation between $\varphi(y)$ and $\tau(y)$ and this should be taken into account. We give here the final expression of $\psi(y)$, derived in detail in [51]:

$$\psi(y) = \tau(y) \frac{\sigma(R)}{\sigma(R[1 + \mathcal{G}_\delta^E]^{1/3})} , \quad (352)$$

where $\tau(y)$ is the solution of implicit equation (263). A formal expansion of $\psi(y)$ with respect to y gives the explicit form of the first few coefficients C_{p1} . They can be expressed in terms of the successive logarithmic derivatives of the variance, γ_i [Eq. (280)]:

$$C_{21} = \frac{68}{21} + \frac{\gamma_1}{3} , \quad (353)$$

⁵³ The interpretation of this function as a bias function is discussed in Section 7.1.2.

$$C_{31} = \frac{11\,710}{441} + \frac{61}{7} \gamma_1 + \frac{2}{3} \gamma_1^2 + \frac{\gamma_2}{3}, \quad (354)$$

$$C_{41} = \frac{107\,906\,224}{305\,613} + \frac{90\,452\gamma_1}{441} + \frac{116\gamma_1^2}{3} + \frac{7\gamma_1^3}{3} + \frac{758\gamma_2}{63} \\ + \frac{20\gamma_1\gamma_2}{9} + \frac{2\gamma_3}{9}. \quad (355)$$

These numbers provide a set of correlators that describe the joint density distribution in the weakly non-linear regime. They generalize the result found initially in [231] for C_{21} . Numerical investigations (e.g. [51]) have shown that the large separation approximation is very accurate even when the cells are quite close to each other.

For a comparison of the above results with N -body simulations and the spherical collapse model see [263].

5.13. Extended perturbation theories

The range of validity of perturbation theory results suggests that they provide, on a sole phenomenological basis, a robust model for describing the correlation hierarchy in all regimes. In the extended perturbation theory (EPT) ansatz, the S_p 's are assumed to be given by Eqs. (281)–(284) with $\gamma_1 \equiv -(n+3)$ and $\gamma_i = 0$, $i \geq 2$, where $n = n_p(\sigma)$ is an *adjustable* parameter inferred from the *measured* value of S_p as a function of the measured variance σ^2 :

$$S_p[n = n_p(\sigma)] \equiv S_p^{\text{measured}}(\sigma). \quad (356)$$

As observed in [151], for scale-free initial conditions, the function $n_p(\sigma)$ does not depend on cumulant order p to a very good approximation:

$$n_p(\sigma) \simeq n_{\text{eff}}(\sigma) \quad (357)$$

in any regime, from very small⁵⁴ value of σ to a very large value of σ . A simple form has been proposed to account for these results [151]:

$$n_{\text{eff}} = n + (n_{\text{nonlinear}} - n) \frac{x^\tau}{x^\tau + x^{-\tau}}, \\ x = \exp[\log_{10}(\sigma^2/\sigma_0^2)], \quad (358)$$

where n_{eff} is varying from the value of the initial power spectrum index, n , to a value corresponding to the stable clustering regime, $n_{\text{nonlinear}}$. The location and the width of the transition between these two regimes depend on the initial power spectrum index and are described respectively by σ_0 and τ . Values of the parameters involved in Eq. (358) are listed in Table 12 for n ranging from -2 to 1 . These values can be approximately obtained by the following fitting formulae valid for $n \lesssim -1$:

$$n_{\text{nonlinear}}(n) \simeq 3 \frac{(n-1)}{(3+n)}, \quad (359)$$

⁵⁴ Of course, in this regime $n_{\text{eff}} = n$, where n is the linear spectral index.

Table 12
Parameters used in fit (358)

n	$n_{\text{nonlinear}}$	$n_{\text{nonlinear}}^-$	$n_{\text{nonlinear}}^+$	σ_0	τ
-2	-9.5	-12.4	-7.22	1.6	1.4
-1	-3	-3.8	-2.24	1.4	1.2
0	-1.2	-1.6	-0.86	1.25	0.6
+1	-0.85	-1.17	-0.57	0.7	0.3

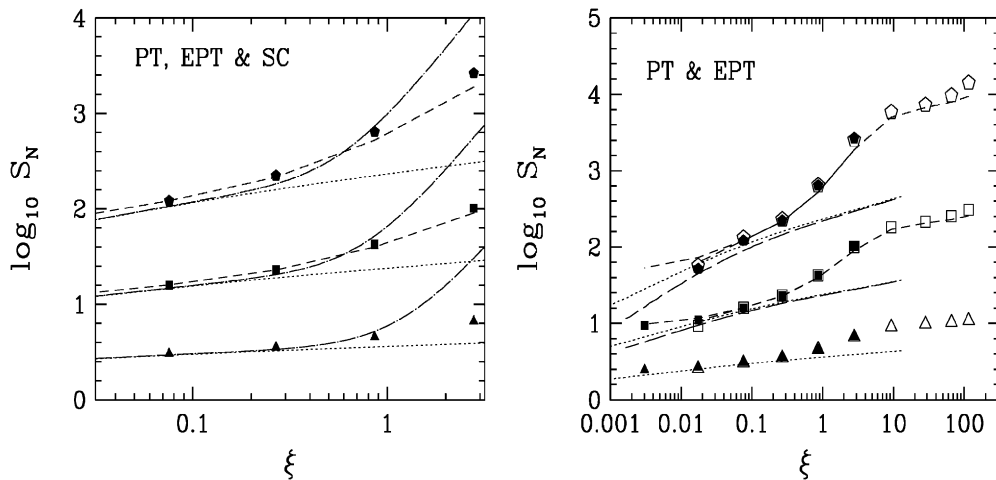


Fig. 37. The cumulants S_p in the τ CDM model as functions of $\xi \equiv \sigma^2$, for $p=3,4$ and 5 (with respectively triangles, squares and pentagons) compared to tree order PT predictions assuming a local power spectrum (dots), taking into account spectral index variation, i.e. corrections γ_p , $p > 2$ in Eqs. (281)–(284) (long dashes on right panel), EPT where n_{eff} is inferred from the measured S_3 (short dashes) and one-loop perturbation theory predictions based on the spherical model (dots–long dashes on left panel). From [153].

$$\tau(n) \simeq 0.8 - 0.3n, \quad (360)$$

$$\log_{10} \sigma_0^2(n) \simeq 0.2 - 0.1n. \quad (361)$$

Eq. (359) is in good agreement with measurements of the bispectrum [234] in N -body simulations as well as predictions from HEPT (Section 4.5.6). For a realistic, scale-dependent spectral index (such as CDM models), the situation becomes slightly more complicated since Eq. (357) is in principle not valid anymore, at least in the weakly non-linear regime, due to the γ_p corrections in Eqs. (281)–(284), which should be taken into account. However, these corrections are in practice quite small [44,28,153] and can be neglected in a first approximation as illustrated in the right panel of Fig. 37. Then, Eq. (357) extends as well to non-scale-free spectra such as CDM models [151,153,629] (see Fig. 37).

It is even possible to use scale-free power spectra results, Eq. (358), with appropriate choice of n in Eqs. (359)–(361), $n = -\gamma_1(R) - 3$ obtained from the linear variance computed at smoothing scale R , to obtain an approximate fit of function $n_{\text{eff}}(\sigma)$ [151]. It is worth noting as well that EPT is a good

approximation for the S_p 's measured in 2D galaxy catalogs, with n_{eff} varying from approximately -2 to -5 depending on the angular scale considered [622].

This description can be extended to the joint moments [623], giving the so-called E²PT framework [630,153]. This provides a reasonable description of the joint cumulants in the non-linear regime, but not as accurate as EPT for one-point cumulants [153]. However, a first application suggests that this is in disagreement with observations [623].

Both EPT and E²PT provide useful ways of describing higher-order statistics as functions of a single parameter n_{eff} and can be used for estimating cosmic errors on statistics measured in galaxy catalogs as discussed in the next section. However, except in the weakly non-linear regime, these prescriptions lack any rigorous theoretical background, although some elements toward their justification can be found in HEPT (see Section 4.5.6).

6. From theory to observations: estimators and errors

6.1. Introduction

This section focuses on issues regarding accurate estimation of clustering statistics in large-scale galaxy surveys and their uncertainties, in order to properly constraint theories against observations. We also consider applications to measurements in N -body simulations, as briefly described in Section 6.12.

In many respects, the theory of estimators of large-scale structure statistics was triggered in the 1970s and the early 1980s by Peebles and his collaborators. In a series of seminal works, starting with a fundamental paper [500], these authors developed the statistical theory of the two-point correlation function in real and Fourier space, in two- and three-dimensional catalogs, including estimates of the cosmic errors and the cosmic bias (formulated as an integral constraint problem), followed soon by investigations on higher-order statistics. They used several estimators, including count-in-cell statistics. These results are summarized in [508].

Since then, and particularly in the 1990s, a number of techniques were put forward to allow a more precise testing of cosmological theories against observations. These include:

- Detailed studies of two-point and higher-order correlation function estimators.
- Accurate estimation of errors going beyond the simple (and often severe underestimate) Poisson error bars, to include finite-volume effects, survey geometry and non-Gaussian contributions due to non-linear evolution.
- The treatment of covariance between measurements at different scales. In order to properly test theoretical predictions, this is equally important to an accurate treatment of errors, which are just the diagonal elements of the covariance matrix. Neglecting off-diagonal elements can lead to a substantial overestimate of the constraining power of observations (see e.g. Section 8).
- Implementation of techniques for data compression, error decorrelation and likelihood analysis for cosmological parameters estimation.

It is clear that the upcoming large-scale galaxy surveys such as 2dFGRS and SDSS will certainly have to rely heavily on these new developments to extract all the information encoded by galaxy

clustering to constrain cosmological parameters, primordial non-Gaussianity and galaxy formation models. In addition to standard second-order statistics such as the power spectrum or the two-point correlation function, our review focuses on higher-order statistics for several reasons:

- As detailed in previous sections, non-linear evolution leads to deviations from Gaussianity, so two-point statistics are not enough to characterize large-scale structure. They do not contain all the information available to constrain cosmological theories.⁵⁵
- The additional information encoded by higher-order statistics can be used, for example, to constrain galaxy biasing (Section 7.1), primordial non-Gaussianity (Sections 4.4 and 5.6) and break degeneracies present in measurements of two-point statistics, e.g. those obtained from measurements of the redshift-space power spectrum (Section 7.4). PT provides a framework for accomplishing this.⁵⁶
- The significant improvement in accuracy for higher-order statistics measurements expected in upcoming large-scale surveys, see e.g. Fig. 40.

Needless is to say that measurements in galaxy catalogs are subject to a number of statistical and systematic uncertainties, that must be properly addressed before comparing to theoretical predictions, succinctly:

- (i) *Instrumental biases and obscuration*: There are technical limitations due to the telescopes and the instruments attached to it. For example, in spectroscopic surveys using multifiber devices such as the SDSS, close pairs of galaxies are not perfectly sampled unless several passes of the same part of the sky are done (e.g. see [74]). This can affect the measurement of clustering statistics, in particular higher-order correlations. Also, the sky is contaminated by sources (such as stars), dust extinction from our galaxy, etc.
- (ii) *Dynamical biases and segregation*: Unfortunately, it is not always possible to measure directly quantities of dynamical interest: in three-dimensional catalogs, the estimated object positions are contaminated by peculiar velocities of galaxies. In 2D catalogs, the effects of projection of the galaxy distribution along the line of sight must be taken into account. Furthermore, galaxy catalogs sample the visible matter, whose distribution is in principle different from that of the matter. The resulting *galaxy bias* might depend on environment, galaxy type and brightness. Objects selected at different distances from the observer do not necessarily have the same properties: e.g. in magnitude-limited catalogs, the deeper objects are intrinsically brighter. One consequence in that case is that the number density of galaxies decreases with distance and thus corrections for this are required unless using volume-limited catalogs.
- (iii) *Statistical biases and errors*: The finite nature of the sample induces uncertainties and systematic effects on the measurements, denoted below as *cosmic bias* and *cosmic error*. These cannot be avoided (although it is possible to estimate corrections in some cases), only reduced by increasing the size of the catalog and optimizing its geometry.

⁵⁵ For example, although one could construct a matter linear power spectrum that evolves non-linearly into the observed *galaxy* power spectrum (see Fig. 51), it is not possible to match at the same time the higher-order correlations at small scales (see Fig. 54). This implies non-trivial galaxy biasing in the non-linear regime, as we discuss in detail in Sections 8.2.4 and 8.2.5.

⁵⁶ A quantitative estimate of how much information is added by considering higher-order statistics is presented in [645].

In this section, we concentrate mainly on point (iii). Dynamical biases mentioned in point (ii) will be addressed in the next section. These effects can also be taken into account in the formalism, by simply replacing the values of the statistics intervening in the equations giving cosmic errors and cross-correlations with the “distorted” ones, as we shall implicitly assume in the rest of this section.⁵⁷ Segregation effects and incompleteness due to instrument biases, obscuration or to selection in magnitude will be partly discussed here through weighted estimators, and in Section 8 when relevant.

This section is organized as follows. In Section 6.2, we discuss the basic concepts of *cosmic bias*, *cosmic error* and the *covariance matrix*. Before entering into technical details, it is important to discuss the fundamental assumptions implicit in any measurement in a galaxy catalog, namely the *fair sample hypothesis* [500] and the *local Poisson approximation*. This is done in Section 6.3, where basic concepts on count-in-cell statistics and discreteness effect corrections are introduced to illustrate the ideas. In Section 6.4, we study the most widely used statistic, the two-point correlation function, with particular attention to the *Landy and Szalay* estimator [393] introduced in Section 6.4.1. The corresponding cosmic errors and biases are given and discussed in several regimes. Section 6.5 is similar to Section 6.4, but treats the Fourier counterpart of ξ , the power spectrum. Generalization to higher-order statistics is discussed in Section 6.6.

Section 6.7 focuses on the count-in-cell distribution function, which probes the density field smoothed with a top-hat window. In that case, a full analytic theory for estimators and corresponding cosmic errors and biases is available. Section 6.8 discusses multivariate counts-in-cells statistics. In Section 6.9 we introduce the notion of optimal weighting: each galaxy or fraction of space can be given a specific statistical weight chosen to minimize the cosmic error. Section 6.10 deals with cross-correlations and the shape of the cosmic distribution function and discusses the validity of the Gaussian approximation, useful for maximum likelihood analysis. Section 6.11 reinvestigates the search for optimal estimators in a general framework in order to give account of recent developments. In particular, error decorrelation and the discrete Karhunen–Loève transforms are discussed. Finally, Section 6.12 discusses the particular case of measurements in N -body simulations.

In what follows, we assume we have a \mathcal{D} -dimensional galaxy catalog D of volume V and containing N_g objects, with $N_g \gg 1$, corresponding to an average number density $\bar{n}_g = N_g/V$. Similarly, we define a pure random catalog R of same geometry and same number of objects.⁵⁸ Despite the fact that we use three-dimensional notations ($\mathcal{D} = 3$) most of the results below are valid as well for angular surveys except when specified otherwise. Simply, $\xi(r)$ has to be replaced with $w(\theta)$, Q_N with q_N , etc.

⁵⁷ Of course, this step can be non-trivial. Measurements in galaxy catalogs (Section 8) and in N -body simulations suggest that in the non-linear regime the hierarchical model is generally a good approximation (e.g. [87,234,147,150,472]), but it can fail to describe fine statistical properties (e.g. for the power spectrum covariance matrix [564,296]). In the weakly non-linear regime, PT results including redshift distortions (Section 7.4), projection along the line of sight (Section 7.2) and biasing (Section 7.1) can help to compute the quantities determining cosmic errors, biases and cross-correlations. In addition to the hierarchical model, extensions of PT to the non-linear regime, such as EPT, E²PT (Section 5.13) and HEPT (Section 4.5.6), coupled with a realistic description of galaxy biasing can be used to estimate the errors.

⁵⁸ Note that R stands as well for a smoothing scale, but the meaning of R will be easily determined by the context.

6.2. Basic concepts

6.2.1. Cosmic bias and cosmic error

In order to proceed we need to introduce some new notation. If A is a statistic, its estimator will be designated by \hat{A} . The probability $\Upsilon(\hat{A})$ of measuring the value \hat{A} in a galaxy catalog (given a theory) will be called the *cosmic distribution function*. The ensemble average of \hat{A} (the average over a large number of virtual realizations of the galaxy catalog) is

$$\langle \hat{A} \rangle = \int d\hat{A} \hat{A} \Upsilon(\hat{A}) . \quad (362)$$

Due to their non-linear nature many estimators (such as ratios) are biased, i.e. their ensemble average is not equal to the real value A : the *cosmic bias* (to distinguish it from the bias between the galaxy distribution and the matter distribution)

$$b_A = \frac{\langle \hat{A} \rangle - A}{A} \quad (363)$$

does not vanish, except when the size of the catalog becomes infinite (if the estimator is properly normalized).

A good estimator should have minimum cosmic bias. It should as well minimize the *cosmic error*, which is usually obtained by calculating the variance of the function Υ :

$$(\Delta A)^2 = \langle (\delta \hat{A})^2 \rangle = \int (\delta \hat{A})^2 \Upsilon(\hat{A}) d\hat{A} \quad (364)$$

with

$$\delta \hat{A} \equiv \hat{A} - \langle \hat{A} \rangle . \quad (365)$$

The cosmic error is most useful when the function $\Upsilon(\hat{A})$ is Gaussian. If this is not the case, full knowledge of the shape of the cosmic distribution function, including its skewness, is necessary to interpret correctly the measurements.⁵⁹

6.2.2. The covariance matrix

As for correlation functions, a simple generalization of the concept of variance is that of covariance between two different quantities. This can be, for example, between two estimators \hat{A} and \hat{B} ,

$$\text{Cov}(\hat{A}, \hat{B}) = \langle \delta \hat{A} \delta \hat{B} \rangle = \int \delta \hat{A} \delta \hat{B} \Upsilon(\hat{A}, \hat{B}) d\hat{A} d\hat{B} , \quad (366)$$

or simply between estimates of the same quantity at different scales; say, for the power spectrum, the covariance matrix between estimates of the power at k_i and k_j reads

$$C_{ij}^P \equiv \langle \hat{P}(k_i) \hat{P}(k_j) \rangle - \langle \hat{P}(k_i) \rangle \langle \hat{P}(k_j) \rangle , \quad (367)$$

where $\hat{P}(k_i)$ is the estimator of the power spectrum at a band power centered about k_i .

In general, testing theoretical predictions against observations requires knowledge of the joint covariance matrix for all the estimators (e.g. power spectrum, bispectrum) at all scales considered. We will consider some examples below in Sections 6.4.4, 6.5.4 and 6.10.2.

⁵⁹ For example, it could be very desirable to impose in this case that a good estimator should have minimum skewness [610].

The cosmic error and the cosmic bias can be roughly separated into three contributions [621] if the scale R (or separation) considered is small enough compared to the typical survey size L , or equivalently, if the volume $v \equiv v_R \equiv (4/3)\pi R^3$ is small compared to the survey volume V :

- (i) *Finite volume effects*: They are due to the fact that we can have access to only a finite number of structures of a given size in surveys (whether they are 2D or 3D surveys), in particular the mean density itself is not always well determined. These effects are roughly proportional to the average of the two-point correlation function over the survey, $\bar{\xi}(L)$. They are usually designated by “cosmic variance”.
- (ii) *Edge effects*: They are related to the geometry of the catalog. In general, estimators give less weight to galaxies near the edge than those far away from the boundaries. As we shall see later, edge effects can be partly corrected for, at least for N -point correlation functions. At leading order in v/V , they are proportional to roughly $\xi v/V$. Note that even 2D surveys cannot avoid edge effects because of the need to mask out portions of the sky due to galaxy obscuration, bright stars, etc. Edge effects vanish only for N -body simulations with periodic boundary conditions.
- (iii) *Discreteness effects*: One usually assumes that the observed galaxy distribution is a discrete, local Poisson representation of an underlying smooth field whose statistical properties one wants to extract. This discrete nature has to be taken into account with appropriate corrections, not only to the mean of a given statistic but also to the error. Discreteness errors, which are proportional to $1/N_g$ at some power where N_g is the number of objects in the catalog, become negligible for large enough N_g .

The above separation into three contributions is convenient but somewhat artificial since all the effects are correlated with each other. For example, there are edge-discreteness effects and edge-finite-volume effects [624]. At next to leading order in R/L , there is a supplementary edge-effect contribution proportional to the perimeter of the survey, which is most important when the geometry of the survey is complex, and dominant when $R/L \approx 1$ [537,154].

6.3. Fair sample hypothesis and local Poisson approximation

6.3.1. The fair sample hypothesis

A stochastic field is called *ergodic* if all information about its multi-point probability distributions (or its moments) can be obtained from a single realization of the field. For example, Gaussian fields with continuous power spectrum are ergodic [3].

The *fair sample hypothesis* [500] states that the finite part of the universe accessible to observations is a fair sample of the whole, which is represented by a statistically homogeneous and isotropic (as defined in Section 3.2.1) ergodic field. Together with the ergodic assumption, the fair sample hypothesis states that well-separated parts of the (observable) universe are independent realizations of the same physical process and that there are enough of such independent samples to obtain all the information about its probability distributions (e.g. [508,61]). Under the fair sample hypothesis, ensemble averages can be replaced with spatial averages. In the simplest inflationary models leading to Gaussian primordial fluctuations, the fair sample hypothesis holds, but special cases can be encountered in models of universe with non-trivial global topological properties (see e.g. [389]) where apparently well-separated parts of the universe may be identical.

6.3.2. Poisson realization of a continuous field

In general, statistical properties of the density field are measured in a discrete set of points, composed e.g. of galaxies or N -body particles. It is natural to assume that such point distributions result from a Poisson realization of an underlying continuous field. This means that the probability of finding N points in a volume v at location \mathbf{r} is given by $P_N^{\text{Poisson}}[\bar{n}_g v(1 + \delta(\mathbf{r}))]$, where $P_N^{\text{Poisson}}(\bar{N})$ is the probability of finding N objects in a Poisson process with expectation number $\bar{N} = \bar{n}_g v$,

$$P_N^{\text{Poisson}}(\bar{N}) \equiv \frac{\bar{N}^N}{N!} e^{-\bar{N}}, \quad (368)$$

$\delta(\mathbf{r})$ is the overall density contrast within the volume and \bar{n}_g is the average number density of the random process. It implies that the *count probability distribution function*, hereafter CPDF, defined as the probability P_N of finding N galaxies in a cell of size R and volume v thrown at random in the catalog can be expressed through the convolution,

$$P_N = \int_{-1}^{+\infty} d\delta p(\delta) P_N^{\text{Poisson}}[\bar{N}(1 + \delta)], \quad (369)$$

where the average number of objects per cells, \bar{N} , reads

$$\bar{N} = \sum_N N P_N. \quad (370)$$

In the continuous limit, $\bar{N} \rightarrow \infty$, the CPDF of course tends to the PDF of the underlying density field

$$P_N \rightarrow \frac{P[\bar{N}(1 + \delta)]}{\bar{N}}. \quad (371)$$

It is worth at this point to mention the void probability function, P_0 , which can be defined in discrete samples only. From Eqs. (369) and (368), it reads

$$P_0 = \int_{-1}^{+\infty} d\delta p(\delta) \exp[-\bar{N}(1 + \delta)], \quad (372)$$

which can be expressed in terms of the cumulant generating function [687,16,619] (see Section 3.3),

$$P_0 = \exp[-\bar{N} + \mathcal{C}(-\bar{N})] = \exp\left[\sum_{n=1}^{\infty} \frac{(-\bar{N})^n}{n!} \langle \delta^n \rangle_c\right]. \quad (373)$$

This property was used in practice to obtain directly the cumulant generating function from the void probability function (e.g. [445,205,92]), relying on the local Poisson approximation.

Obviously, the validity of the local Poisson approximation is questionable. A simple argument against it is that galaxies have an extended size which defines zones of mutual exclusion and suggests that at very small scales, galaxies do not follow a local Poisson process because they must be anti-correlated. One way to bypass this problem is of course to choose the elementary volume such that it has a sufficiently large size, say $\ell \gtrsim$ a few tens of kpc. One might still argue that short-range physical processes depending on environment might influence small-scale statistics in such a way that it might be impossible to find a reasonably small-scale ℓ for which the Poisson process is valid. Also, the galaxy distribution might keep memory of initial fluctuations of the density field, even at small, non-linear scales, particularly in underdense regions which do not experience

shell-crossing and violent relaxation. If, for example, these initial conditions were locally fractal up to some very small scale, obviously the local Poisson approximation would break down. Note, on the other hand, that sparse sampling strategies [361] which were used to build a number of galaxy catalogs, make the samples “closer” to Poisson.

It is generally assumed that the observed galaxy distribution follows the local Poisson approximation. To our knowledge there exists no direct rigorous check of the validity of this statement, but it is supported indirectly, for example by the fact that the measured count probability distribution function (CPDF, see Section 6.7) in galaxy catalogs compares well with models relying on the local Poisson approximation (see, e.g. [92]).

In N -body simulations, the local Poisson assumption is in general very good.⁶⁰ However, this depends on the statistic considered and there are some requirements on the degree of evolution of the system into the non-linear regime, as discussed in Section 6.12.2.

Under the assumption of local Poisson approximation, it is possible to derive the correlation functions of the discrete realization in terms of the underlying continuous one. In particular, from Eq. (369) the moment generating function of the discrete realization, $\mathcal{M}_{\text{disc}}$, is related to that of the continuous field, \mathcal{M} (Section 3.3.3), by $\mathcal{M}_{\text{disc}}(t) = \mathcal{M}(t)[\exp(t) - 1]$. This leads to the standard expressions for moments and spectra of discrete realizations in terms of continuous ones, e.g. see [396,508,233,619,247,434]. Here we give the first few low-order moments:

$$\langle \delta_n^2 \rangle = \frac{1}{\bar{N}} + \bar{\xi}_2, \quad (374)$$

$$\langle \delta_n^3 \rangle = \frac{1}{\bar{N}^2} + 3 \frac{\bar{\xi}_2}{\bar{N}} + \bar{\xi}_3, \quad (375)$$

where $\delta_n \equiv (N - \bar{N})/\bar{N}$ denotes the discrete number density contrast. In Section 6.7, which discusses in more detail count-in-cells statistics, we shall see that there exists an elegant way of correcting for discreteness effects using factorial moments.

Similarly, for the power spectrum and bispectrum,

$$\langle \delta_n(\mathbf{k}_1) \delta_n(\mathbf{k}_2) \rangle = \left[\frac{1}{N_g} + P(k_1) \right] \delta_D(\mathbf{k}_{12}), \quad (376)$$

$$\langle \delta_n(\mathbf{k}_1) \delta_n(\mathbf{k}_2) \delta_n(\mathbf{k}_3) \rangle = \left[\frac{1}{N_g^2} + \frac{1}{N_g} (P_1 + P_2 + P_3) + B_{123} \right] \delta_D(\mathbf{k}_{123}), \quad (377)$$

where $P_i \equiv P(k_i)$, $B_{123} \equiv B(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)$, $\mathbf{k}_{12} = \mathbf{k}_1 + \mathbf{k}_2$ and $\mathbf{k}_{123} = \mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3$.

6.4. The two-point correlation function

In this section, we present the traditional estimators of the two-point correlation function based on pairs counting.⁶¹ We assume that the catalog under consideration is statistically homogeneous.

⁶⁰ Except when dealing with the clustering of dark matter halos; in this case exclusion effects can lead to sub-Poisson sampling, see e.g. [599].

⁶¹ For a review on existing estimators, see, e.g. [372,525].

Optimal weighting and correction for selection effects will be treated in Section 6.9. More elaborate estimates taking into account cross-correlations between bins will be discussed in Section 6.10.

6.4.1. Estimators

In practice, due to the discrete nature of the studied sample, the function ξ [Eq. (115)] is not measured at separation exactly equal to r but rather one must choose a bin, e.g. $[r, r + \Delta r[$. More generally, the quantity measured is

$$\frac{1}{G_p^\infty V_\infty^2} \int_{V_\infty} d^{\mathcal{D}} \mathbf{r}_1 d^{\mathcal{D}} \mathbf{r}_2 \Theta(\mathbf{r}_1, \mathbf{r}_2) \xi(r_{12}), \quad (378)$$

where the function $\Theta(\mathbf{r}_1, \mathbf{r}_2)$ is symmetric in its arguments (e.g. [624]). In what follows, we assume that the function Θ is invariant under translations and rotations, $\Theta(\mathbf{r}_1, \mathbf{r}_2) = \Theta(r)$, $r = r_{12} = |\mathbf{r}_1 - \mathbf{r}_2|$, is unity on a domain of values of r , for example in the interval $[r, r + \Delta r[$ and vanishes otherwise. The values where Θ is non-zero define a “bin” which we call Θ as well. We assume that $\xi(r)$ is sufficiently smooth and that the bin and the normalization, G_p^∞ , are such that Eq. (378) would reduce with a good accuracy to $\xi(r)$ in a survey of very large volume V_∞ .

Practical calculation of the two-point correlation function relies on the fact that it can be defined in terms of the excess probability over random δP of finding two galaxies separated by a distance (or an angle) r [as discussed already in Section 3, Eq. (127)]

$$\delta P = \bar{n}_g^2 [1 + \xi(r)] \delta V_1 \delta V_2, \quad (379)$$

where δV_1 and δV_2 are volume (surface) elements and \bar{n}_g is the average number density of objects.

Let DD be the number of pairs of galaxies in the galaxy catalog belonging to the bin Θ and RR defined likewise but in a random (Poisson distributed) catalog with same geometry and same number of objects, $N_r = N_g$. They read

$$DD = \int_{\mathbf{r}_1 \neq \mathbf{r}_2} d^{\mathcal{D}} \mathbf{r}_1 d^{\mathcal{D}} \mathbf{r}_2 \Theta(\mathbf{r}_1, \mathbf{r}_2) n_g(\mathbf{r}_1) n_g(\mathbf{r}_2), \quad (380)$$

$$RR = \int_{\mathbf{r}_1 \neq \mathbf{r}_2} d^{\mathcal{D}} \mathbf{r}_1 d^{\mathcal{D}} \mathbf{r}_2 \Theta(\mathbf{r}_1, \mathbf{r}_2) n_r(\mathbf{r}_1) n_r(\mathbf{r}_2), \quad (381)$$

where n_g and n_r are local number density fields respectively in the galaxy catalog and the random catalog:

$$n_g = \sum_{j=1}^{N_g} \delta_D(\mathbf{x} - \mathbf{x}_j), \quad (382)$$

where \mathbf{x}_j are the galaxy positions and likewise for n_r . It is easy to derive from Eq. (379) a simple estimator commonly used in the literature [503]:

$$\hat{\xi}(r) = \frac{DD}{RR} - 1. \quad (383)$$

Various alternatives have been proposed to improve the estimator given by Eq. (383), in particular to reduce the cosmic bias induced by edge effects at large separations. Detailed studies [373] suggest

that the best of them is the Landy and Szalay (LS) estimator [393]⁶²

$$\hat{\xi}(r) = \frac{DD - 2DR + RR}{RR}, \quad (384)$$

where DR is the number of pairs selected as previously but the first object belongs to the galaxy sample and the second one to the random sample

$$DR = \int_{\mathbf{r}_1 \neq \mathbf{r}_2} d^{\mathcal{D}}\mathbf{r}_1 d^{\mathcal{D}}\mathbf{r}_2 \Theta(\mathbf{r}_1, \mathbf{r}_2) n_g(\mathbf{r}_1) n_r(\mathbf{r}_2). \quad (385)$$

The LS estimator, which formally can be written $(D_1 - R_1)(D_2 - R_2)/R_1 R_2$ corresponds to the “intuitive” procedure of first calculating overdensities and then expectation values; this has the obvious generalization to higher-order correlation functions [624], see Section 6.6 for more details.

Note that the calculations of DR and RR can be arbitrarily improved by arbitrarily increasing N_r and applying the appropriate corrections to DR and RR , i.e. multiplying DR and RR by the ratio N_g/N_r and $N_g(N_g - 1)/[N_r(N_r - 1)]$ respectively, to preserve normalization. Actually, DR and RR can be computed numerically as integrals with a different method than generating a random catalog, the latter being equivalent to Monte-Carlo simulation. It amounts to replace DR and RR by DF and FF with

$$DF = \bar{n}_g \int_{\mathbf{r}_1 \neq \mathbf{r}_2} d^{\mathcal{D}}\mathbf{r}_1 d^{\mathcal{D}}\mathbf{r}_2 \Theta(\mathbf{r}_1, \mathbf{r}_2) n_g(\mathbf{r}_1), \quad (386)$$

$$FF = \bar{n}_g^2 \int_{\mathbf{r}_1 \neq \mathbf{r}_2} d^{\mathcal{D}}\mathbf{r}_1 d^{\mathcal{D}}\mathbf{r}_2 \Theta(\mathbf{r}_1, \mathbf{r}_2). \quad (387)$$

In that case, the actual measurements are performed on pixelized data.

The LS estimator is theoretically optimal with respect to both cosmic bias and cosmic error at least in the weak correlation limit [393]; numerical studies [373] show moreover that for practical purposes it is better than any other known estimators based on pair counting, among which one can mention $(DD - DR)/RR$ [311], the popular $DD/DR - 1$ [172,68] and $DDRR/(DR)^2 - 1$ [291] which are actually almost as good as LS [373]. In Section 6.8 we shall mention other ways of measuring $\xi(r)$ and higher-order correlation functions, based on multiple counts-in-cells.

Finally, it is worth mentioning a few efficient methods used to measure $\xi(r)$, which apply to any of the estimators discussed in this section. The brute force approach is indeed rather slow since it scales typically as $\mathcal{O}(N_g^2)$. To improve the speed of the calculation, one often interpolates the sample onto a grid and creates a linked list where each object points to a neighbor belonging to the same grid site. For separations smaller than the grid step, λ , this method scales roughly as $\mathcal{O}(N_g N_{\text{cell}})$, where N_{cell} is the typical number of objects per grid cell. This approach is however limited by the step of the grid: measuring the correlation function at scales large compared to λ is rather inefficient and can become prohibitive. Increasing λ makes N_{cell} larger and for too large λ , the method is slow again.

Another scheme relies on a double walk in a quad-tree or a oct-tree according to the dimension of the survey (a hierarchical decomposition of space in cubes/squares and subcubes/subsquares

⁶² See however [525] for a more reserved point of view.

[461]). This approach is potentially powerful, since it scales as $\mathcal{O}(N_g^{3/2})$ according to its authors [461]. It is also possible to rely on FFTs or fast harmonic transforms at large scales [636], but it requires appropriate treatment of the Fourier coefficients to make sure that the quantity finally measured corresponds to the estimator of interest, e.g. the LS estimator (see [636] for a practical implementation in harmonic space).

6.4.2. Cosmic bias and integral constraint of the LS estimator

The full calculation of the cosmic bias and the cosmic error of the LS estimator was done by Landy and Szalay [393] in the weak correlation limit and by Bernstein [59] for the general case but neglecting edge effects, $r \ll L$, where L is the smallest size of the survey.⁶³ At leading order in r/L and assuming that the density variance at the scale of the survey is small, the cosmic bias reads

$$b_\xi \simeq \left(3 - \frac{1}{\xi}\right) \bar{\xi}(L) - 2 \frac{\check{\xi}_3}{\xi} - \frac{1}{2N_g^2}, \quad r/L, |\bar{\xi}(L)|, \left| \frac{\bar{\xi}(L)}{\xi} \right| \ll 1, \quad (388)$$

where

$$\bar{\xi}(L) = \frac{1}{V^2} \int d^{\mathcal{D}} \mathbf{r}_1 d^{\mathcal{D}} \mathbf{r}_2 \xi(r) \quad (389)$$

is the average of the correlation function over the survey volume (or area). The quantity $\check{\xi}_3$ is defined as

$$\check{\xi}_3 = \frac{1}{G_p V^3} \int d^{\mathcal{D}} \mathbf{r}_1 d^{\mathcal{D}} \mathbf{r}_2 d^{\mathcal{D}} \mathbf{r}_3 \Theta(r_{12}) \xi_3(r_1, r_2, r_3), \quad (390)$$

where G_p is the form factor defined in [393] as

$$G_p = \frac{1}{V^2} \int d^{\mathcal{D}} \mathbf{r}_1 d^{\mathcal{D}} \mathbf{r}_2 \Theta(r_{12}), \quad (391)$$

i.e. the probability of finding a pair included in the survey in bin Θ . When r/L is small enough it is simply given by $G_p \simeq 4\pi r^2 \Delta r / V$ (for a bin $\Theta = [r, r + \Delta r]$). Assuming the hierarchical model, Eq. (214), we get $\check{\xi}_3 \simeq 2Q_3 \xi \bar{\xi}(L)$ and the cosmic bias simplifies to

$$b_\xi \simeq \left(3 - 4Q_3 - \frac{1}{\xi}\right) \bar{\xi}(L) - \frac{1}{2N_g^2}, \quad r/L, |\bar{\xi}(L)|, \left| \frac{\bar{\xi}(L)}{\xi} \right| \ll 1. \quad (392)$$

In the weak correlation limit, it simply reduces to [393]

$$b_\xi \simeq \frac{-\bar{\xi}(L)}{\xi}, \quad |\xi|, |\bar{\xi}(L)| \ll 1. \quad (393)$$

The LS estimator, although designed to minimize both the cosmic error and the cosmic bias and thus quite insensitive to edge effects and discreteness effects, is still affected by finite-volume effects, proportional to $\bar{\xi}(L)$ (indeed the latter cannot be reduced without prior assumptions about clustering at scales larger than those probed by the survey, as discussed below). The corresponding cosmic

⁶³ It is however important to notice a subtle difference between the two approaches: Landy and Szalay use conditional averages with fixed number of galaxies in the catalog N_g , while N_g is kept random in Bernstein's approach. This difference is analyzed in Section 6.10.

bias is negative, of small amplitude in the highly non-linear regime, but becomes significant when the separation r becomes comparable to the survey size. In this regime, where $\xi(r)$ is expected to be much smaller than unity, Eq. (393) is generally valid: the correct value of ξ is obtained by adding an unknown constant to the measured value. This corresponds to the so-called *integral constraint* problem [502,508]. Physically, it arises in a finite survey because one is estimating the mean density and fluctuations about it from the same sample, and thus the fluctuation must vanish at the survey scale. In other words, one cannot estimate correlations at the survey scale since there is only one sample available of that size.

This bias cannot be a priori corrected for unless a priori assumptions are made on the shape of the two-point correlation function at scales larger than those probed by the survey. One can, for instance, decide to model the two-point correlation as a power law and do a joint determination of all parameters [502]. We will come back to this problem when discussing the case of the power spectrum, where other corrections have been suggested, see Section 6.5.2.

6.4.3. Cosmic error of the LS estimator

The general computation of the cosmic error for such estimator is quite involved and has been derived in the literature in various cases. For instance, the covariance of $DD - 2DF + FF$ between two bins Θ_a and Θ_b reads [500,291,634]

$$\begin{aligned} \text{Cov}(DD - 2DF + FF) &= \bar{n}_g^4 \int d^{\mathcal{D}}\mathbf{r}_1 d^{\mathcal{D}}\mathbf{r}_2 d^{\mathcal{D}}\mathbf{r}_3 d^{\mathcal{D}}\mathbf{r}_4 \Theta_a(\mathbf{r}_1, \mathbf{r}_2) \Theta_b(\mathbf{r}_3, \mathbf{r}_4) \\ &\quad \times [\xi_4(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4) + \xi(\mathbf{r}_1, \mathbf{r}_3)\xi(\mathbf{r}_2, \mathbf{r}_4) + \xi(\mathbf{r}_1, \mathbf{r}_4)\xi(\mathbf{r}_2, \mathbf{r}_3)] \\ &\quad + 4\bar{n}_g^3 \int d^{\mathcal{D}}\mathbf{r}_1 d^{\mathcal{D}}\mathbf{r}_2 d^{\mathcal{D}}\mathbf{r}_3 \Theta_a(\mathbf{r}_1, \mathbf{r}_2) \Theta_b(\mathbf{r}_1, \mathbf{r}_3) [\xi(\mathbf{r}_2, \mathbf{r}_3) + \xi_3(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)] \\ &\quad + 2\bar{n}_g^2 \int d^{\mathcal{D}}\mathbf{r}_1 d^{\mathcal{D}}\mathbf{r}_2 \Theta_a(\mathbf{r}_1, \mathbf{r}_2) \Theta_b(\mathbf{r}_1, \mathbf{r}_2) [1 + \xi(\mathbf{r}_1, \mathbf{r}_2)]. \end{aligned} \quad (394)$$

This is a general expression, i.e. it applies to the two-point correlation function as well as the power spectrum, or any pairwise statistics of the density field, depending on the choice of the binning function Θ . It does not, however, take into account the possible cosmic fluctuations of the denominator in the LS estimator. This latter effect is more cumbersome to compute because one has to deal with moments of the inverse density. This is possible if one assumes that fluctuations are small. This leads to the cosmic covariance derived in [59] for the LS estimator. We give here a simplified expression of the diagonal term, the cosmic error:

$$\begin{aligned} \left(\frac{\Delta \xi}{\xi} \right)^2 &\simeq 2 \frac{\bar{\xi}^2}{\xi^2} + 4(1 - 2Q_3 + Q_4)\bar{\xi}(L) + \frac{4}{N_g} \left[\frac{\xi_{\text{ring}}(1 + 2Q_3\xi)}{\xi^2} + Q_3 - 1 \right] \\ &\quad + \frac{2}{N_g^2} \left[\left(\frac{1}{G_p} - 1 \right) \frac{1 + \xi}{\xi^2} - \frac{1}{\xi} - 1 \right], \quad r/L, |\bar{\xi}(L)|, |\bar{\xi}(L)/\xi| \ll 1, \end{aligned} \quad (395)$$

where $\bar{\xi}^2$ is the average of the square of the two-point correlation function over the survey volume,

$$\bar{\xi}^2 = \frac{1}{G_p^2 V^4} \int d^{\mathcal{D}}\mathbf{r}_1 \cdots d^{\mathcal{D}}\mathbf{r}_4 \Theta(r_{12}) \Theta(r_{34}) \xi_2(r_{13}) \xi_2(r_{24}), \quad (396)$$

and ξ_{ring} is the average of the two-point correlation function for pairs inside the shell of radius r and thickness Δr ,

$$\xi_{\text{ring}} = \frac{1}{G_t V^3} \int d^{\mathcal{D}} \mathbf{r}_1 d^{\mathcal{D}} \mathbf{r}_2 d^{\mathcal{D}} \mathbf{r}_3 \Theta(r_{12}) \Theta(r_{13}) \xi(r_{23}). \quad (397)$$

We have introduced the new geometrical factor G_t given by [393]

$$G_t = \frac{1}{V^3} \int d^{\mathcal{D}} \mathbf{r}_1 d^{\mathcal{D}} \mathbf{r}_2 d^{\mathcal{D}} \mathbf{r}_3 \Theta(r_{12}) \Theta(r_{13}), \quad (398)$$

i.e. G_t is the probability, given one point, of finding two others in bin Θ , for example the interval $[r, r + \Delta r]$. As pointed out in [59], $\xi_{\text{ring}} \gtrsim \xi$, but

$$\xi_{\text{ring}} \simeq \xi \quad (399)$$

is a good approximation. In Eq. (395), a degenerate hierarchical model (Section 4.5.5) has been assumed to simplify the results. A more general expression can be found in [59] (see also [291,634]).

The finite volume errors are given by a term in ξ^2 and one proportional to $\bar{\xi}(L)$. It is interesting to compare which two contributions. For a power-law spectrum of index n , $\bar{\xi}^2/\xi^2$ scales like $(r/L)^{\mathcal{D}}$ whereas $\bar{\xi}(L)$ scales like $(r_0/L)^{-(\mathcal{D}+n)}$ if r_0 is the correlation length ($\xi(r_0) \equiv 1$). Therefore, in the quasi-linear regime for which $r \gg r_0$ and for surveys with a large number of objects, the first term is likely to dominate (this is the case typically for wide angular surveys), whereas for surveys which probe deeply into the non-linear regime, the other terms are more likely to dominate.

The discreteness error is given by the term in $1/N_g$ which vanishes for a randomized purely Poisson catalog. The intrinsic Poisson error is encoded in the term in $(1/N_g)^2$. This estimate of the cosmic error neglects however edge effects that become significant at scales comparable to the size of the survey. In this latter regime, correlations are expected to be weak, and from [393] one finds that the cosmic error is dominated by edge-discreteness effects [624]:

$$\left(\frac{\Delta \xi}{\xi} \right)^2 \simeq \frac{2}{N_g^2 \xi^2} \left[\frac{1}{G_p} - 2 \frac{G_t}{G_p^2} + 1 \right], \quad |\xi|, |\bar{\xi}(L)| \ll 1. \quad (400)$$

One can note that when r/L is small enough, the term in square brackets is roughly equal to $1/G_p$ [as in Eq. (395)], that is the fraction of pairs available in the survey. This is obviously the dominant contribution of the error when the bin size Δr is very small. This pure Poisson contribution can generally be computed exactly given the geometry of the survey.

Expressions (395) and (400) can be used to estimate the full cosmic error. This method however requires prior assumptions about the hierarchical model parameters Q_3 and Q_4 and for the integral of the two-point correlation function over the survey volume, $\bar{\xi}(L)$. For this reason, the Gaussian limit is often used to compute errors (that is the contribution of $\bar{\xi}^2$, e.g. [410]), but this might be a bad approximation when $\xi \gtrsim 1$ as we discussed above.⁶⁴

6.4.4. The covariance matrix

As discussed above, Eq. (394) gives the cosmic covariance matrix of the two-point correlation function assuming that \bar{n}_g is perfectly determined, while the calculation of Bernstein [59], for which

⁶⁴ Fig. 38, extracted from [564], illustrates that for the power spectrum.

we gave a simplified expression of the diagonal terms, takes into account possible fluctuations in \bar{n}_g . We refer the reader to [59] for the full expression of C_ξ which is rather cumbersome.

Interestingly, the pure Poisson contribution vanishes for non-overlapping bins in Eq. (394). A simplified formula can be obtained in the Gaussian limit where non-Gaussian and discreteness contributions can be neglected,

$$\begin{aligned} C_\xi(r_a, r_b) &= \langle \hat{\xi}_2(r_a) \hat{\xi}_2(r_b) \rangle - \langle \hat{\xi}_2(r_a) \rangle \langle \hat{\xi}_2(r_b) \rangle \\ &= \frac{2}{G_p(r_a) G_p(r_b) V^4} \int d^{\mathcal{D}} \mathbf{r}_1 \cdots d^{\mathcal{D}} \mathbf{r}_4 \Theta_a(r_{12}) \Theta_b(r_{34}) \xi_2(r_{13}) \xi_2(r_{24}), \end{aligned} \quad (401)$$

in particular, $C_\xi(r, r) = \overline{\xi^2}$ [Eq. (396)]. This expression can be conveniently expressed in terms of the power spectrum. It reads, for $\mathcal{D} = 3$,

$$C_\xi(r_a, r_b) = \frac{(2\pi)^5}{V} \int k^2 dk [P(k)]^2 J_{1/2}(kr_a) J_{1/2}(kr_b), \quad (402)$$

where $J_{1/2}$ is a Bessel function. A similar expression has been derived for 2D fields [204],

$$\begin{aligned} C_w(\theta_a, \theta_b) &= \langle \hat{w}_2(\theta_a) \hat{w}_2(\theta_b) \rangle - \langle \hat{w}_2(\theta_a) \rangle \langle \hat{w}_2(\theta_b) \rangle \\ &= \frac{2(2\pi)^3}{A_\Omega} \int_0^\infty k dk [P(k)]^2 J_0(k\theta_a) J_0(k\theta_b), \end{aligned} \quad (403)$$

where A_Ω is the area of the survey, $w_2(\theta)$ represents the angular two-point function and \hat{w}_2 its estimator.

Note that as the volume/area of the survey increases, the diagonal terms in Eq. (401) do not, in general, become dominant compared to the off-diagonal ones. This is because correlation function measurements are statistically correlated, even in the Gaussian limit, unlike binned power spectrum measurements, e.g. see Section 6.5.4.

6.4.5. Recipes for error calculations

The issue of cosmic error computation is recurrent in cosmological surveys and the previous computations clearly show that this is a complex issue. Various recipes have been proposed in the literature. A particularly popular one is the bootstrap method [24]. We stress that bootstrap resampling is not suited for correlation function measurements. Indeed, as shown explicitly in [597], such method does not lead, in general, to a reliable estimate of the cosmic error [525,373].

Another popular and elementary way of estimating the errors consists in dividing the catalog into a number of smaller subsamples of same volume and compute the dispersion in the measurements corresponding to each subsample (e.g. [249]). This method is not free of bias and generally overestimates the errors, since the obtained dispersion is an estimator of the cosmic error on the subsamples and not the parent catalog. Recent studies on error estimation [572,704] also suggest that the Jackknife method, which is a variant of the subsample method where the i th sample is obtained

by *removing* the i th subsample, gives a very good estimate of the cosmic error on the two-point correlation function. Unlike the subsample method, it does not lead to overestimation of the cosmic error at large scales.⁶⁵

Of course, methods such as Jackknife and subsamples cannot lead to an accurate estimation of finite-volume errors at the scale of the survey since only one realization of such a volume is available to the observer. This can only be achieved through a detailed computation of the cosmic errors [Eqs. (395) and (400)] with prior assumptions about the behavior of statistics involved at scales comparable to the survey size, or else numerically by constructing multiple realizations of the survey, e.g. mock catalogs relying on N -body simulations or simplified versions thereof (e.g. [571]). On the other hand, methods that use the actual data are very useful to assess systematic errors, by comparing to other external estimates such as those just mentioned.

6.5. The power spectrum

The power spectrum $P(k)$ is simply the Fourier transform of the two-point correlation function (see Section 3.2.2), and therefore it is formally subject to the same effects. In fact, a common theoretical framework can be set up for $\xi(r)$ and $P(k)$ in order to find the best estimators (e.g. [293,294,624]). In practice, however, power spectrum measurements have been undertaken mostly in linear or weakly non-linear scales which are subject to edge effects, difficult to correct for. In this section, we introduce simple (unweighted) estimators and discuss the biases and cosmic error introduced by the finiteness of the survey. The techniques developed to measure $P(k)$ are numerous and sometimes very elaborate (a good review can be found in [648]), but most of them rely on the assumption that the underlying statistics is Gaussian. In this section we prefer to keep the statistical framework general and thus restrict ourselves to traditional estimators. More sophisticated methods, using spatial weighting and cross-correlations between bins, will be discussed in Sections 6.9 and 6.11.

6.5.1. Simple estimators

For convenience, in finite surveys the adopted normalization convention for the Fourier transforms and the power spectra is often different. This is the reason why in this subsection, we also adopt following convention:

$$\tilde{A}(\mathbf{k}) = \frac{1}{V} \int_V d^{\mathcal{D}} \mathbf{x} e^{-i\mathbf{k}\cdot\mathbf{x}} A(\mathbf{x}), \quad (404)$$

where $\tilde{A}(\mathbf{k})$ are the Fourier modes of $A(\mathbf{x})$ and V is the survey volume (and to recover the convention used in Eq. (36), one can simply use the formal correspondence $V \leftrightarrow (2\pi)^{\mathcal{D}}$). The power spectrum is defined as the Fourier transform of the two-point correlation function. It differs thus by a $V/(2\pi)^{\mathcal{D}}$ normalization factor compared to the adopted normalization in the other sections. The higher-order spectra are defined similarly from the higher-order correlation functions in such a way that the functional relation between spectra is preserved [e.g. the coefficients \tilde{Q} in Eq. (154) are left unchanged].

⁶⁵ An alternative to these methods has been suggested by Hamilton [291], in which many realizations from a given sample are generated by effectively varying the pair-weighting function.

As shown in previous sections, estimating the correlation function consists in counting pairs in bins, both in the galaxy catalog and in random realizations with the same survey geometry. This procedure can be generalized to the measurement of the power spectrum (e.g. [212]) for which the binning function Θ defined in Section 6.4.2 is now different. For one single mode the straightforward choice would be (e.g. [624]) $\Theta(\mathbf{r}_1, \mathbf{r}_2) = (e^{i\mathbf{k}\cdot(\mathbf{r}_1-\mathbf{r}_2)} + e^{i\mathbf{k}\cdot(\mathbf{r}_2-\mathbf{r}_1)})/2$. Actual estimation of the power is made over a k bin defined for instance so that the magnitude of wave vectors belong to a given interval $[k, k + \Delta k[$. It means that the function Θ to use actually reads

$$\Theta(\mathbf{r}_1, \mathbf{r}_2) = \langle e^{i\mathbf{k}\cdot(\mathbf{r}_1-\mathbf{r}_2)} \rangle_{\Theta} \equiv \frac{1}{V_k} \int_{|\mathbf{k}'| \in [k, k+\Delta k[} d^{\mathcal{D}} \mathbf{k}' e^{i\mathbf{k}'\cdot(\mathbf{r}_1-\mathbf{r}_2)}, \quad (405)$$

where V_k is the volume of the bin in k space. Note that for a rectangular-shaped survey with periodic boundaries modes are discrete and the number of modes in V_k is

$$N_k = \frac{V V_k}{(2\pi)^{\mathcal{D}}}. \quad (406)$$

In the following we assume that V_k is large enough to encompass a sufficient number of modes to make any measurement possible. With this expression of Θ the quantities DD , DR , RR , DF and FF defined in (380)–(387) where Θ is replaced by Eq. (405) can be used to estimate the power spectrum [624].

Traditionally, the estimate of the power spectrum is done in the following way: the density contrast is Fourier transformed directly (e.g. [500,492,679,215,489]):

$$\hat{\delta}_{\mathbf{k}} = \frac{1}{V} \int \left[\frac{n_{\mathbf{g}}(\mathbf{r})}{\bar{n}_{\mathbf{g}}} - 1 \right] e^{i\mathbf{k}\cdot\mathbf{x}} d^{\mathcal{D}} \mathbf{x} = \frac{1}{N_{\mathbf{g}}} \sum_{j=1}^{N_{\mathbf{g}}} e^{i\mathbf{k}\cdot\mathbf{x}_j} - W_{\mathbf{k}}, \quad (407)$$

where $W_{\mathbf{k}}$ is the Fourier transform of the window function of the survey,

$$W_{\mathbf{k}} = \frac{1}{V} \int e^{i\mathbf{k}\cdot\mathbf{x}} d^{\mathcal{D}} \mathbf{x}. \quad (408)$$

The power spectrum estimator is then given by

$$\hat{P}(k) = \langle |\hat{\delta}_{\mathbf{k}}|^2 \rangle_{\Theta} - \frac{1}{N_{\mathbf{g}}}, \quad (409)$$

where $\langle \dots \rangle_{\Theta}$ stands for summation in the k bin [e.g. Eq. (405)], which can also be written as

$$\hat{P}(k) = \frac{1}{N_{\mathbf{g}}^2} (DD - 2DF + FF). \quad (410)$$

Note that the correction for shot noise contribution is automatically taken into account by the exclusion $\mathbf{r}_1 \neq \mathbf{r}_2$ in the integral DD . One can see that this is analogous to LS estimator (384) in Fourier space [624].

6.5.2. Cosmic bias and integral constraint

As for the two-point correlation function, it is possible to show that the estimator in Eq. (410) is biased [500,492], at least due to finite volume effects. Again this is generally described as the integral constraint problem.

The expressions for the cosmic bias can be directly inferred from Eqs. (388) and (393). More specifically, at large, weakly non-linear scales, where the Gaussian limit is a good approximation, the cosmic bias reads [492]

$$b_{\hat{P}(k)} \simeq -P_*(0) \frac{\langle |W_{\mathbf{k}}|^2 \rangle_{\Theta}}{\langle P_*(\mathbf{k}) \rangle_{\Theta}}. \quad (411)$$

The quantity P_* is the true power spectrum convolved with the Fourier transform of the window function of the survey:

$$P_*(\mathbf{k}) = P(\mathbf{k}) * |W_{\mathbf{k}}|^2. \quad (412)$$

Note that $P_*(0)$ is nothing but $\bar{\xi}(L)$ [Eq. (389)].

At smaller scales, in the regime $k \gg 1/L$, the cosmic bias reads

$$b_{\hat{P}(k)} \simeq P_*(0) \left[3 - \frac{\langle |W_{\mathbf{k}}|^2 \rangle_{\Theta}}{\langle P(\mathbf{k}) \rangle_{\Theta}} \right] - \frac{2\langle B_*(\mathbf{k}, -\mathbf{k}, 0) \rangle_{\Theta}}{\langle P(\mathbf{k}) \rangle_{\Theta}}, \quad (413)$$

where B_* is the bispectrum (convolved with the Fourier transform of the survey window).

In general, the cosmic bias is approximated by the white noise value in the Gaussian limit [489]

$$b_{\hat{P}(k)} \simeq -\langle |W_{\mathbf{k}}|^2 \rangle_{\Theta} = -FF/N_g^2, \quad (414)$$

and the corresponding correction is applied to estimator (410).

An interesting approach to correct for the cosmic bias takes advantage of the Gaussian limit expression, Eq. (411). Since the bias is proportional to the Fourier transform of the window of the survey, construction of a tailored window such that $W_{\mathbf{k}} = 0$ for each mode \mathbf{k} of interest makes Eq. (411) vanish [215,648]. However, one must keep in mind that this procedure is approximate; even in the Gaussian limit there are higher-order corrections to the result in Eq. (411) which are not proportional to $W_{\mathbf{k}}$.⁶⁶

6.5.3. The cosmic error

The calculation of the cosmic error on the power spectrum is formally equivalent to that of the two-point correlation function. However, existing results assume that the average number density of galaxies in the universe is an external parameter, i.e. the ensemble average $\langle [\delta\hat{P}(k)]^2 \rangle$ is calculated with N_g fixed in Eq. (410).

In the limit when $k \gg 1/L$, where L is the smallest size of the survey, for the power spectrum Eq. (394) reads

$$\begin{aligned} \left[\frac{\Delta\hat{P}(k)}{P(k)} \right]^2 &\simeq \frac{2}{N_k} + \frac{\bar{T}(k, k)}{[P(k)]^2} + \frac{4}{N_g} \left[\frac{1}{N_k P(k)} + \frac{\bar{B}(k, k)}{[P(k)]^2} \right] \\ &\quad + \frac{2}{N_g^2} \left[\frac{1}{N_k [P(k)]^2} + \frac{\bar{P}(k, k)}{[P(k)]^2} \right] \end{aligned} \quad (415)$$

⁶⁶ The cosmic bias in this expression comes in fact from the uncertainty in the mean density \bar{n}_g from the numerator in $\delta = (n_g - \bar{n}_g)/\bar{n}_g$; uncertainties from the denominator lead to additional contributions, see e.g. [328].

with

$$\begin{aligned} \bar{T}(k_i, k_j) &\equiv \langle T(\mathbf{k}_1, -\mathbf{k}_1, \mathbf{k}_2, -\mathbf{k}_2) \rangle_{\theta_{k_i}, \theta_{k_j}} \\ &\equiv \int_{|\mathbf{k}_1| \in [k_i, k_i + \Delta k_i]} \frac{d^{\mathcal{D}} \mathbf{k}_1}{V_{k_i}} \int_{|\mathbf{k}_2| \in [k_j, k_j + \Delta k_j]} \frac{d^{\mathcal{D}} \mathbf{k}_2}{V_{k_j}} T(\mathbf{k}_1, -\mathbf{k}_1, \mathbf{k}_2, -\mathbf{k}_2), \end{aligned} \quad (416)$$

$$\bar{B}(k_i, k_j) \equiv \langle B(\mathbf{k}_1, \mathbf{k}_2, -\mathbf{k}_1 - \mathbf{k}_2) \rangle_{\theta_{k_i}, \theta_{k_j}}, \quad (417)$$

$$\bar{P}(k_i, k_j) \equiv \frac{1}{2} \langle P(\mathbf{k}_1 + \mathbf{k}_2) + P(\mathbf{k}_1 - \mathbf{k}_2) \rangle_{\theta_{k_i}, \theta_{k_j}}. \quad (418)$$

This result assumes that the true power spectrum is sufficiently smooth and the bin in k space thin enough that $\langle P(k) \rangle_{\theta_k} \simeq P(k)$, $\langle P(k)^2 \rangle_{\theta_k} \simeq [P(k)]^2$. The continuous limit $N_g \rightarrow \infty$ of Eq. (415) was computed in [564], and the Gaussian limit, $B = T = 0$ in [212].

From the calculations of [564], one gets

$$\bar{T}(k, k) \simeq \frac{232}{441} [P(k)]^3 \quad (419)$$

in the regime where PT applies, and

$$\bar{T}(k, k) \simeq (8Q_{4,a} + 4Q_{4,b}) [P(k)]^3, \quad (420)$$

if the hierarchical model applies (Section 4.5.5) [564,296]. Similar calculations can be done to evaluate $\bar{B}(k, k)$ and $\bar{P}(k, k)$.

One must emphasize [452,564] again the fact that the Gaussian limit, traditionally used to compute errors and optimal weighting (see Section 6.9), is invalid when $k \gtrsim k_{\text{nl}}$, where k_{nl} is the transition scale to the non-linear regime defined from the power spectrum, $4\pi k_{\text{nl}}^3 P(k_{\text{nl}}) \equiv 1$. This is clearly illustrated by top panel of Fig. 38. It compares the measured cosmic error obtained from the dispersion over 20 PM simulations of Λ CDM with the Gaussian limit [564]. This shows that the Gaussian limit underestimates the cosmic error, increasingly with k/k_{nl} . Note, however, that the correction brought by Eq. (419) is rather small. As a result, the regime where the Gaussian limit is a reasonable approximation for estimating the cosmic error extends up to values of k/k_{nl} of order of a few. This is unfortunately not true for the full cosmic covariance matrix $C_{ij}^P \equiv \text{Cov}(P_{k_i}, P_{k_j})$, which deviates from the Gaussian predictions (vanishing non-diagonal terms) as soon as $k \simeq k_{\text{nl}}$ [452,564], as we now discuss.

6.5.4. The covariance matrix

The covariance of the power spectrum, Eq. (367), can be easily written beyond the Gaussian approximation neglecting shot noise and the window of the survey [452,564],⁶⁷

$$C_{ij}^P = \frac{2P^2(k_i)}{N_{k_i}} \delta_{ij} + \bar{T}(k_i, k_j), \quad (421)$$

where δ_{ij} is a Kronecker delta and \bar{T} is the bin-averaged trispectrum (416).

⁶⁷ See e.g. [293] for expressions including shot noise.

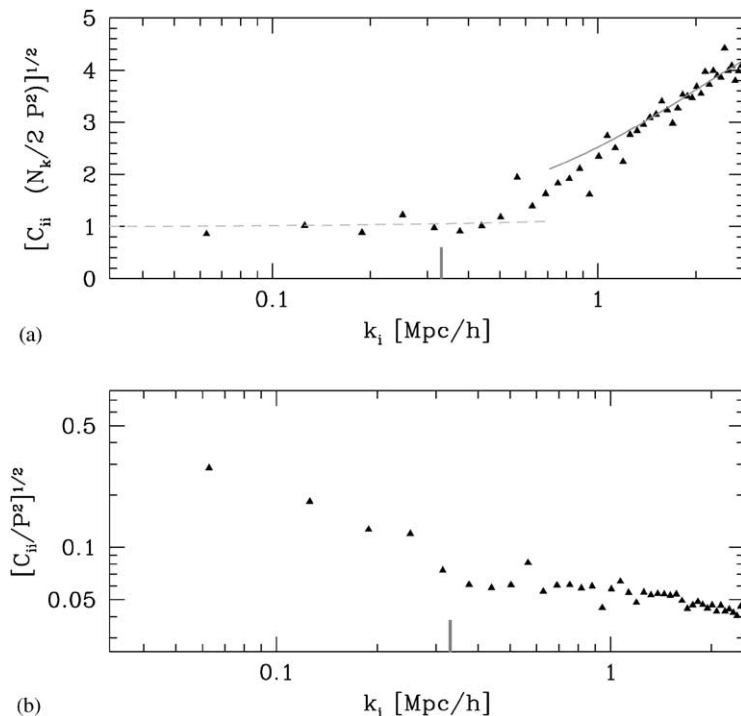


Fig. 38. The top panel shows the measured cosmic error on the power spectrum normalized by the Gaussian variance, obtained from the dispersion over 20 PM simulations of SCDM. The dashed line shows the predictions of PT, and the solid line the hierarchical scaling. The bottom panel shows the fractional error in the band-power estimates. This fractional error scales with the size of the survey or simulation box, the results in the figure correspond to a volume $V_0 = (100 h^{-1} \text{ Mpc})^3$. Results for other volumes can be obtained by scaling by $(V_0/V)^{1/2}$. The vertical line on the x-axis indicates the non-linear scale. The width of shells in k space is $\Delta k = 2\pi/100 h/\text{Mpc}$.

The first term in Eq. (421) is the Gaussian contribution. In the Gaussian limit, each Fourier mode is an independent Gaussian random variable. The power estimates of different bands are therefore uncorrelated, and the covariance is simply given by $2/N_{k_i}$ where $N_{k_i}/2$ is the number of independent Gaussian variables. The second term in Eq. (421) arises because of non-Gaussianity, which generally introduces correlations between different Fourier modes, and hence it is not diagonal in general.

Both terms in the covariance matrix in Eq. (421) are inversely proportional to V for a fixed bin size (recall that with the adopted convention $P(k)$ scales like $1/V$ and T like $1/V^3$). But while the Gaussian contribution decreases when N_k increases, the non-Gaussian term remains constant. Therefore, when the covariance matrix is dominated by the non-Gaussian contribution, the only way to reduce the variance of the power spectrum is to increase the volume of the survey instead of averaging over more Fourier modes.

The importance of the non-Gaussian contribution to the cross-correlation between band powers was studied with numerical simulations in [452,564], in particular Meiksin and White [452] show in detail that the correlations induced by non-linearities are not negligible even at scales $k \lesssim k_{\text{nl}}$, in agreement with PT predictions [564]. In the non-linear regime, as expected, the cross-correlations are very strong; indeed, the cross-correlation coefficient $r_{ij} \equiv C_{ij}/\sqrt{C_{ii}C_{jj}}$ is very close to unity.

Predictions for r_{ij} from the hierarchical ansatz using HEPT amplitudes (see Section 4.5.6) are in reasonable agreement with simulations [564], although at large separations ($k_i \gg k_j$) there are significant deviations [564,296].

An efficient (although approximate) numerical approach to computing the covariance matrix of the power spectrum is presented in [571], using a combination of 2LPT at large scales, and knowledge about dark matter halos at small scales (see e.g. Sections 7.1.3 and 7.1.4), which also allows to take into account the effects of redshift distortions and galaxy biasing.

6.6. Generalization to higher-order correlation functions

Higher-order statistics such as correlation functions in real and Fourier space were not studied in as much detail as the power spectrum and the two-point correlation function. In particular, there is no accurate analytic estimate of the cosmic bias and error on such statistics,⁶⁸ although a general formalism (relying on a statistical framework set up by Ripley [537]) which we summarize below was recently developed by Szapudi and collaborators [624,633,634].

The LS estimator presented in Section 6.4.1 for the two-point correlation function, $\langle \delta_1 \delta_2 \rangle$, can be formally written as $(D_1 - R_1)(D_2 - R_2)/R_1 R_2$. As suggested in [624], a simple generalization for a statistic of order N , for example the unconnected N -point correlation function, $f_N \equiv \langle \delta_1 \cdots \delta_N \rangle$, is simply $(D_1 - R_1)(D_2 - R_2) \cdots (D_N - R_N)/R_1 \cdots R_N$. More exactly, [Szapudi and Szalay [624] and Szapudi [634]] define symbolically an estimator $D^p R^q$ with $p + q = N$ for a function Θ symmetric in its arguments,

$$D^p R^q = \sum \Theta(\mathbf{x}_1, \dots, \mathbf{x}_p, \mathbf{y}_1, \dots, \mathbf{y}_q) \quad (422)$$

where $\mathbf{x}_i \neq \mathbf{x}_j \in D$ and $\mathbf{y}_i \neq \mathbf{y}_j \in R$ are objects positions in the galaxy catalog and the random catalog, respectively. The generalized LS estimator reads

$$\hat{f}_N = \frac{1}{S} \sum_i \binom{N}{i} (-1)^{N-i} \left(\frac{D}{\bar{n}_g}\right)^i \left(\frac{R}{\bar{n}_r}\right)^{N-i}, \quad (423)$$

where the normalization number S is given by

$$S \equiv \int \Theta(\mathbf{x}_1, \dots, \mathbf{x}_N) d^{\mathcal{D}} \mathbf{x}_1 \cdots d^{\mathcal{D}} \mathbf{x}_N. \quad (424)$$

If \bar{n}_g is determined with arbitrary accuracy estimator (423) is unbiased, optimally edge corrected in the weak-correlation limit [624]. For practical measurements, however, \bar{n}_g is determined from the catalog itself, and the integral constraint problem arises again, as described in Section 6.4.3.

The cosmic covariance of \hat{f}_N assuming that n_g is perfectly determined was given in [634],

$$\begin{aligned} \text{Cov}(f_{N_1}, f_{N_2}) &\equiv \langle \hat{f}_{N_1,a} \hat{f}_{N_2,b} \rangle - \langle \hat{f}_{N_1,a} \rangle \langle \hat{f}_{N_2,b} \rangle \\ &= \frac{1}{S^2} \sum_{i,j} \binom{N_1}{i} \binom{N_2}{j} (-1)^{i+j} [E(i,j,N_1,N_2) \\ &\quad - \mathcal{P}_0\{f_i(1, \dots, i) f_j(N_1 + 1, \dots, N_1 + j)\}] \end{aligned} \quad (425)$$

⁶⁸ See however the attempt in [458] about estimating the error on ξ_3 in various approximations.

with

$$\begin{aligned}
 E(p_1, p_2, N_1, N_2) &\equiv \left\langle \left(\frac{D}{\bar{n}_g} \right)^{p_1} \left(\frac{R}{\bar{n}_r} \right)^{N_1 - p_1} \left(\frac{D}{\bar{n}_g} \right)^{p_2} \left(\frac{R}{\bar{n}_r} \right)^{N_2 - p_2} \right\rangle \\
 &= \sum_i \binom{p_1}{i} \binom{p_2}{i} i! \bar{n}_g^{-i} \mathcal{S}_i \{ f_{N_1 + p_1 + p_2 - i} \}, \tag{426}
 \end{aligned}$$

where the operator \mathcal{S}_i is defined by

$$\begin{aligned}
 \mathcal{S}_k \{ g \} &\equiv \int d^{\mathcal{D}} \mathbf{x}_1 \cdots d^{\mathcal{D}} \mathbf{x}_{N_1 + N_2 - k} \Theta_a(1, \dots, N_1) \Theta_b(1, \dots, k, N_1 + 1, \dots, N_1 + N_2 - k) \\
 &g(1, \dots, p_1, N_1 + 1, \dots, N_1 + p_2 - k), \tag{427}
 \end{aligned}$$

and the convention that $\binom{k}{l}$ is non-zero only for $k \geq 0$, $l \geq 0$ and $k \geq l$. In these equations we have used the short-hand notations, $1 = \mathbf{x}_1, \dots, i = \mathbf{x}_i$, etc., and g should be viewed as $f_i(1, \dots, i) f_j(N_1 + 1, \dots, N_1 + j)$ in Eq. (427) to compute the \mathcal{S}_0 term in Eq. (425).

Eq. (425) assumes that the random catalog contains a very large number of objects, $\bar{n}_r \rightarrow \infty$, i.e. does not take into account errors brought by the finiteness of N_r (see [634] for more details). Using a computer algebra package, one can derive from this formalism Eq. (394). Similar but cumbersome expression for the three-point correlation function can be found in [634].

Note, as suggested in [624], that this formalism can be applied to Fourier space, i.e. to the power spectrum (see [636] for a practical implementation of estimator \hat{f}_2 in harmonic space) and to the bispectrum. It can also be theoretically applied to one-point distribution functions, such as count-in-cells, studied below, but it was not done so far. Therefore, we shall instead present results relying on a more traditional approach in the next section.

Note that for the bispectrum, some work has been done in computing its covariance matrix and cosmic bias in particular cases. In [434], the bispectrum covariance matrix is estimated including shot-noise terms and beyond the Gaussian approximation⁶⁹ by using second-order Eulerian PT.⁷⁰ A numerical calculation of the bispectrum covariance matrix and the cosmic bias expected for IRAS surveys is presented in [566] using 2LPT.⁷¹

6.7. One-point distributions: counts-in-cells

6.7.1. Definitions

The count probability distribution function (CPDF) was introduced in Section 6.3.2. Here we give more definitions on count-in-cells statistics, such as factorial moments and their relation to cumulants and the CPDF in terms of generating functions. Some additional information can be found as well in Appendix E.

⁶⁹ Estimation of the cosmic error in the Gaussian approximation is given in [234,560].

⁷⁰ This is, however, only approximate since a consistent calculation of the connected six-point function requires up to fifth-order Eulerian PT, a quite complicated calculation.

⁷¹ This is also not a consistent calculation of non-Gaussian terms in the covariance matrix; however, 2LPT does include significant contributions to any order in Eulerian PT, and comparison for one-point moments suggests 2LPT is a very good approximation [561].

Following the presentation in Section 6.3.2, we discuss in more detail here an elegant way of correcting for discreteness effects, which makes use of the *factorial moments*. These are defined as follows:

$$F_k \equiv \langle (N)_k \rangle = \langle N(N-1)\cdots(N-k+1) \rangle = \sum_N (N)_k P_N . \quad (428)$$

Note thus that $\bar{N} = F_1$. We have

$$F_k = \bar{N}^k \langle (1 + \delta)^k \rangle , \quad (429)$$

so F_k/\bar{N}^k estimates directly the moment of order k of the underlying (smoothed) density field.

The generating function of the counts

$$\mathcal{P}(t) \equiv \sum_N t^N P_N \quad (430)$$

is related to the moment generating function through

$$\mathcal{M}(\bar{N}t) = \mathcal{P}(t+1) . \quad (431)$$

Factorial moments thus verify

$$F_k = \left(\frac{\partial}{\partial t} \right)^k \mathcal{P}(t+1) \Big|_{t=0} . \quad (432)$$

It is easy to find, using Eq. (141), the following useful recursion [619] relating factorial moments to quantities of physical interest, S_p ,

$$S_p = \frac{\bar{\xi}_2 F_p}{N_c^p} - \frac{1}{p} \sum_{q=1}^{p-1} \binom{p}{q} \frac{(p-q) S_{p-q} F_q}{N_c^q} , \quad (433)$$

where N_c is the typical number of object in a cell in overdense regions, $N_c \equiv \bar{N} \bar{\xi}_2$.

6.7.2. Estimators

In practice, the measurement of the CPDF and its factorial moments is very simple. It consists of throwing C cells at random in the catalog and computing

$$\hat{P}_N^C = \frac{1}{C} \sum_{i=1}^C \delta_{N_i, N} , \quad (434)$$

where $\delta_{N,M}$ is the Kronecker delta function and N_i denotes the number of objects in cell “ i ”. Similarly, the estimator for the factorial moment of order k is

$$\hat{F}_k^C = \frac{1}{C} \sum_{i=1}^C (N_i)_k \quad (435)$$

or can be derived directly from \hat{P}_N^C using Eq. (428). Estimators (434) and (435) are unbiased. However, if one uses relation (433) to compute cumulants from factorial moments, i.e.

$$\hat{\xi} = \frac{\hat{F}_2}{\hat{F}_1^2} - 1 , \quad (436)$$

$$\hat{S}_3 = \frac{\hat{F}_1(\hat{F}_3 - 3\hat{F}_1\hat{F}_2 + 2\hat{F}_1^3)}{(\hat{F}_2 - \hat{F}_1^2)^2}, \quad (437)$$

$$\hat{S}_4 = \frac{\hat{F}_1^2(\hat{F}_4 - 4\hat{F}_3\hat{F}_1 - 3\hat{F}_2^2 + 12\hat{F}_2\hat{F}_1^2 - 6\hat{F}_1^4)}{(\hat{F}_2 - \hat{F}_1^2)^3}, \quad (438)$$

the corresponding estimators are biased because non-linear combinations of estimators are generally biased (e.g. [328,630]).

To reduce the bias and the errors on direct measurements of cumulants from Eqs. (436)–(438) it is possible to use some prior information, for example by assuming that the PDF of the underlying density field is given by the Edgeworth expansion, Eq. (144), convolved with a Poisson distribution to take into account discreteness, Eq. (369). This procedure was actually applied to the IRAS 1.2Jy galaxy catalog [377]. The advantage of such a method is that it can be less sensitive to finite-volume effects by using the shape of the PDF near its peak (since finite-volume effects mainly affect the tails). One disadvantage is that the validity of the Edgeworth expansion is quite restricted, even in the weakly non-linear regime (see, e.g. [356]). In particular, the PDF is not positive definite. Convolution with the Poisson distribution to account for discreteness alleviates this problem for the sparse IRAS surveys [377]; however, for applications to the next generation of galaxy surveys this will likely not be the case. Another difficulty of this approach is that error estimation is not straightforward. On the other hand, the idea of using prior information on the shape of the PDF to estimate moments is certainly worth pursuing with a more detailed modeling of the density PDF.

6.7.3. Error propagation: cosmic bias vs. cosmic error

We now review the theory of error propagation in a general setting for functions of correlated random variables, following the treatment in [630].⁷² This theory was actually behind the calculation of the errors on the two-point correlation function in Section 6.4. Since the calculations are necessarily technical, we only present computations of the cosmic bias and error on non-linear estimators such as those given by Eqs. (436)–(438).

Let us suppose that we measure a quantity $f(\hat{\mathbf{x}})$, where $\hat{\mathbf{x}}$ is a vector of unbiased estimators, such as the factorial moments, and that the measurement of $\hat{\mathbf{x}}$ is sufficiently close to the ensemble average $\langle \hat{\mathbf{x}} \rangle = \mathbf{x}$. Then f can be expanded around the mean value

$$f(\hat{\mathbf{x}}) = f(\mathbf{x}) + \sum_k \frac{\partial f}{\partial x_k} \delta \hat{x}_k + \frac{1}{2} \sum_{k,l} \frac{\partial^2 f}{\partial x_k \partial x_l} \delta \hat{x}_k \delta \hat{x}_l + \mathcal{O}(\delta x^3), \quad (439)$$

where x_k is the k th component of $\hat{\mathbf{x}}$ and

$$\delta \hat{x}_k = \hat{x}_k - x_k. \quad (440)$$

After ensemble average of Eq. (439) one obtains

$$\langle f \rangle = f(\mathbf{x}) + \frac{1}{2} \sum_{k,l} \frac{\partial^2 f}{\partial x_k \partial x_l} \langle \delta \hat{x}_k \delta \hat{x}_l \rangle + \mathcal{O}(\delta x^3). \quad (441)$$

⁷² For a different approach, based on an expansion in terms of the variance at the scale of the survey see [328].

To second order the cosmic bias [Eq. (363)] thus reads

$$b_f \simeq \frac{1}{2f(\mathbf{x})} \frac{\partial^2 f}{\partial x_k \partial x_l} \langle \delta \hat{x}_k \delta \hat{x}_l \rangle. \quad (442)$$

Similarly, the covariance between two functions f and g is

$$\text{Cov}(f, g) = \langle \delta \hat{f} \delta \hat{g} \rangle = \sum_{k,l} \frac{\partial f}{\partial x_k} \frac{\partial g}{\partial x_l} \langle \delta \hat{x}_k \delta \hat{x}_l \rangle + \mathcal{O}(\delta x^3). \quad (443)$$

In particular, the relative cosmic error is given by

$$\sigma_f \equiv \frac{\Delta f}{\langle f \rangle} = \sqrt{\text{Cov}(f, f) / \langle f \rangle}. \quad (444)$$

It is important to notice the following point, from Eqs. (442) and (443):

$$b_f \sim \mathcal{O}(\sigma_f^2). \quad (445)$$

The range of applicability of this perturbative theory of error propagation is $\langle \delta \hat{x}_k \delta \hat{x}_l \rangle / x_k x_l \ll 1$: errors and cross-correlations of the vector $\hat{\mathbf{x}}$ must be weak. In this regime *the cosmic bias is always smaller than the relative cosmic error*, except for accidental cancellations in Eq. (442) (in that case, the next order would be needed in the expansion). When the cosmic bias becomes large the expansion in Eq. (439) breaks down; in this case, numerical simulations show that the cosmic bias can be larger than the relative cosmic error [328].

6.7.4. Cosmic error and cross-correlations of factorial moments

According to the above formalism, the knowledge of errors and cross-correlations on a complete set of unbiased estimators, such as the factorial moments, F_k , $k = 1, \dots, \infty$, or count-in-cells themselves, P_N , allows the calculation of the cosmic error (or cross-correlations) on any counts-in-cells statistics. The general theoretical framework for computing the cosmic error on factorial moments can be found in [621,630].⁷³ Here we review the main results.

First, it is important to notice that there is a source of error due to the finiteness of the number of cells C used in Eqs. (434) and (435). This source of error, which is estimated in [621], can be rendered arbitrarily small by taking very large number of sampling cells, C , or by using an algorithm equivalent to infinite sampling, $C \rightarrow \infty$ as proposed in [625]. Measurements are often done using $C \simeq V/v$, i.e. the number of cells necessary to cover the sample, which *is not* a good idea. Indeed, such small number of sampling cells *does not, in general, extract all the statistically significant information* from the catalog, except in some particular regimes in the Poisson limit. The best way to measure count-in-cells statistics is thus to do *as massive oversampling as possible*⁷⁴ and estimate the cosmic error independently, as explained below. Similarly, when measuring the two-point correlation function using a Poisson sample R to estimate RR and DR , in order to avoid adding noise to the measurements, the random catalog R should have as many objects as possible. Having that in mind, we shall assume from now that C is very large.

⁷³ See the earlier work in [149] for detailed calculations of the void probability function cosmic error.

⁷⁴ This is because missing clusters cores, which occupy a very small fraction of the volume, leads to underestimation of higher-order moments.

The error-generating function is defined as follows

$$\mathcal{E}(x, y) = \sum_{N, M} [\langle \hat{P}_N \hat{P}_M \rangle - \langle \hat{P}_N \rangle \langle \hat{P}_M \rangle], \quad (446)$$

where the ensemble average $\langle \dots \rangle$ denotes the average over a large number of realizations of the catalog with same geometry and same underlying statistics. Then, the cosmic covariance on factorial moments and count-in-cells reads

$$\Delta_{k, l} \equiv \text{Cov}(F_k, F_l) = \left(\frac{\partial}{\partial x} \right)^k \left(\frac{\partial}{\partial y} \right)^l \mathcal{E}(x+1, y+1) \Big|_{x=y=0}, \quad (447)$$

$$\text{Cov}(P_N, P_M) = \left(\frac{\partial}{\partial x} \right)^N \left(\frac{\partial}{\partial y} \right)^M \mathcal{E}(x, y) \Big|_{x=y=0}. \quad (448)$$

The error generating function can be written in terms of bivariate distributions

$$\mathcal{E}(x, y) = \frac{1}{\hat{V}^2} \int_{\hat{V}} d^{\mathcal{D}} \mathbf{r}_1 d^{\mathcal{D}} \mathbf{r}_2 [\mathcal{P}(x, y) - \mathcal{P}(x)\mathcal{P}(y)]. \quad (449)$$

In this equation, \hat{V} is the volume covered by cells included in the catalog and $\mathcal{P}(x, y)$ is the generating function of bicounts $P_{N, M}$ for cells separated by a distance $|\mathbf{r}_1 - \mathbf{r}_2|$ (see also Section 6.8):

$$\mathcal{P}(x, y) \equiv \sum_{N, M} x^N y^M P_{N, M}. \quad (450)$$

The calculation of the function $\mathcal{E}(x, y)$, detailed in Appendix F, is simplified by separating the integral in Eq. (449) into two components, $E_{\text{overlap}}(x, y)$ and $E_{\text{disjoint}}(x, y)$, according to whether cells overlap or not.

At leading order in v/V , $\Delta_{k, l}$ has three contributions

$$\Delta_{k, l} = \Delta_{k, l}^{\text{F}} + \Delta_{k, l}^{\text{E}} + \Delta_{k, l}^{\text{D}}, \quad (451)$$

where $\Delta_{k, l}^{\text{F}}$, $\Delta_{k, l}^{\text{E}}$ and $\Delta_{k, l}^{\text{D}}$ are the finite volume, edge and discreteness effect contributions, respectively. From [621,630], the first few terms in the three-dimensional case are listed in Appendix F.

The finite-volume error comes from the disjoint cells contribution in the error generating function. The corresponding relative error, or cross-correlation, $\Delta_{k, l}^{\text{F}}/(F_k F_l)$ does not depend on the number of objects in the catalog, and is proportional to the integral of the two-point correlation function over the survey volume:

$$\bar{\xi}(\hat{L}) \equiv \frac{1}{\hat{V}} \int_{r_{12} \geq 2R} d^{\mathcal{D}} \mathbf{r}_1 d^{\mathcal{D}} \mathbf{r}_2 \xi(r_{12}). \quad (452)$$

The edge effect term, $\Delta_{k, l}^{\text{E}}/(F_k F_l)$, is the contribution remaining from overlapping cells in the continuous limit, $\bar{N} \rightarrow \infty$. It does not depend on the number of objects in the catalog and is proportional to $\bar{\xi}v/V$. A pure Poisson sample does not have edge effect error at leading order in v/V , in agreement with intuition. The discreteness effect error, $\Delta_{k, l}^{\text{D}}/(F_k F_l)$, is the contribution from overlapping cells which depends on \bar{N} and thus disappears in the continuous limit. As discussed in the introduction of this section, the separation between these three contributions is useful but somewhat arbitrary.

For example, Eq. (452) actually contains some edge effects through the constrain $r_{12} \geq 2R$, as shown in Appendix F.

Furthermore, if next to leading order contributions in v/V are considered, the corresponding correction is proportional to the contour of the survey, ∂V [537,154]. Each contribution, $\Delta_{k,l}^X/(F_k F_l)$, $X = F, E$ or D contains a term proportional to ∂V . This correction is an edge correction, leading to terms such as edge-finite-volume and edge-discreteness contributions in our nomenclature.

It is important to emphasize that the expressions given in Appendix F *are of direct practical use*⁷⁵ for estimating errors on factorial moments or on cumulants (Section 6.7.5) using the theory of propagation of errors explained above (e.g. [319,632,635] for applications to actual measurements in real galaxy catalogs). Similar to Eq. (395), a careful examination of these expressions shows that prior knowledge of the shape of the two-point correlation function ξ [namely, $\bar{\xi}$ and $\bar{\xi}(\hat{L})$] and higher-order statistics, S_p and C_{pq} up to some value of p and q is necessary to compute $\Delta_{k,l}$. To estimate cumulants $\bar{\xi}$ and S_p , one can simply use the values directly measured in the catalog or other existing estimates (e.g. [249,622]), as well as existing fitting formulae for $\bar{\xi}$ ([289,493,335,494], see Section 4.5.4) and PT, EPT ([151], see Section 5.13) or HEPT ([563], see Section 4.5.6) for S_p . To compute $\bar{\xi}(\hat{L})$, it is necessary to make assumptions about the cosmological model. The cumulant correlators C_{pq} can be estimated directly from the catalog or from various models which further simplify the calculations (e.g. [41,619,630]). These models can be particular cases of the hierarchical model, Eq. (214), or can rely on PT results (Section 5.12) or extensions such as E²PT (Section 5.13).

Among the models tested, the best known so far is E²PT as illustrated in Fig. 39. In this figure, taken from [153], the cosmic error on factorial moments is measured from the dispersion over 4096 subsamples of size $L = 125h^{-1}$ Mpc, extracted from a τ CDM simulation of size $2000h^{-1}$ Mpc involving 1000^3 particles [206]. The accuracy of theoretical predictions is quite good, especially at large, weakly non-linear scales. At small scales, all the models tend to overestimate the magnitude of the errors, including E²PT, but the disagreement between theory and measurements is at most a factor two approximately. This discrepancy suggests that details of the dynamics still need to be understood in order to describe appropriately multivariate distribution functions in the highly non-linear regime.

6.7.5. Cosmic error and cosmic bias of cumulants

Using the results in Sections 6.7.3 and 6.7.4 it is possible to compute the cosmic bias and the cosmic error on estimators (436)–(438) (see also [328]). It would be too cumbersome to put all the results here, but getting analytic expressions similar to what was obtained for $\Delta_{k,l}$ is very easy with standard mathematical packages. For example, simple algebraic calculations give for the cosmic bias

$$b_{\bar{\xi}} = \frac{F_2}{\bar{\xi}\bar{N}^2} \left(\frac{3\Delta_{11}}{\bar{N}^2} - \frac{2\Delta_{12}}{\bar{N}F_2} \right), \quad (453)$$

$$b_{S_3} = b_{\bar{\xi}_3} - 3b_{\bar{\xi}} - \frac{2\Delta_{23}}{F_2F_3} + \frac{3\Delta_{22}}{F_2^2}, \quad (454)$$

⁷⁵ They have been implemented in the publically available FORCE package [630].

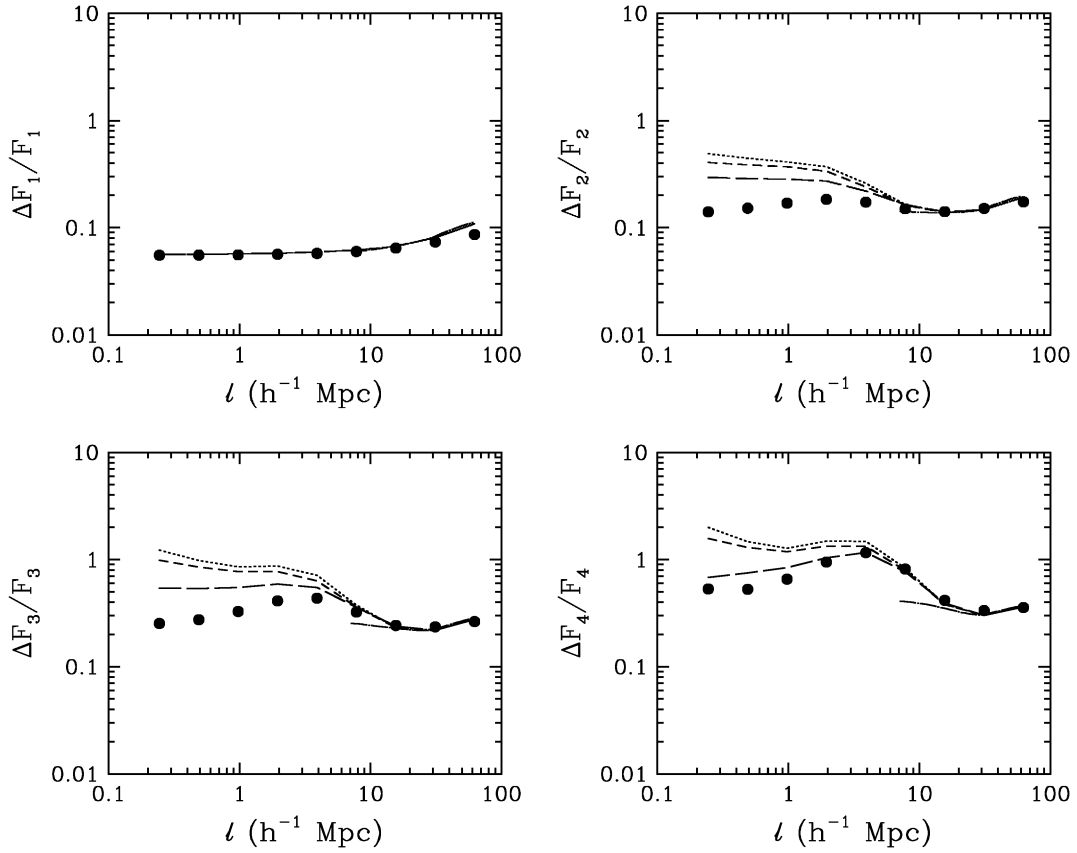


Fig. 39. The relative cosmic error on factorial moments measured as a function of scale [153], obtained from the dispersion over a large ensemble of subsamples extracted from one of the Hubble volume simulations [206], as explained in the text. The dotted, dashed, long dashed, dot–long dashed curves correspond respectively to theoretical predictions based on two particular cases of the hierarchical model, namely SS and BeS, E²PT and PT. The SS model [619] assumes $Q_{NM} = Q_{N+M}$ with the definition in Eq. (F.24). The BeS model [41] is more complicated, but obeys $Q_{NM} = Q_{N1}Q_{M1}$, as in the E²PT framework, described in Section 5.13. The PT results are shown only in the weakly non-linear regime, $\bar{\xi} \lesssim 1$.

with

$$b_{\bar{\xi}_3} = \frac{F_3}{\bar{\xi}_3} \bar{N}^3 \left(\frac{6\Delta_{11}}{\bar{N}^2} - \frac{3\Delta_{13}}{\bar{N}F_3} \right) - 3 \frac{F_2}{\bar{\xi}_3 F_1^2} \left(\frac{3\Delta_{11}}{\bar{N}^2} - \frac{2\Delta_{12}}{\bar{N}F_2} \right). \quad (455)$$

Similarly, the cosmic errors read

$$\begin{aligned} \sigma_{\bar{\xi}}^2 &\simeq \frac{1}{\bar{N}^6} \left(4F_2^2 \Delta_{11} - 4\bar{N}F_2 \Delta_{21} + \bar{N}^2 \Delta_{22} \right), \quad (456) \\ \sigma_{\bar{\xi}_3}^2 &\simeq \frac{1}{\bar{N}^{12} \bar{\xi}^6 S_3^2} \left[(2\bar{N}^3 F_2 - 6\bar{N}F_2^2 + 3\bar{N}^2 F_3 + F_2 F_3)^2 \Delta_{11} \right. \\ &\quad \left. + 2\bar{N}(-2\bar{N}^6 F_2 + 12\bar{N}^4 F_2^2 - 18\bar{N}^2 F_2^3 - 3\bar{N}^5 F_3) \right] \end{aligned}$$

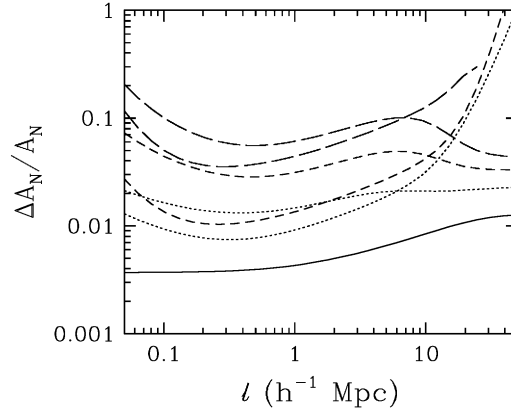


Fig. 40. Comparison of the cosmic errors for the factorial and connected moments expected in the SDSS [630]. Standard CDM is assumed for the two-point correlation function and E²PT with $n_{\text{eff}} = -2.5$ for higher-order statistics. Solid, dotted, dash, and long dash lines correspond to orders 1,2,3, and 4, respectively. Of each pair of curves with the same line-types the one turning up on large scales relates to the cumulant. Note that the perturbative approach used to compute the cosmic error on the cumulants fails at large scales, explaining the right stopping point of the long dash curve for S_4 .

$$\begin{aligned}
 &+ 4\bar{N}^3 F_2 F_3 + 15\bar{N} F_2^2 F_3 - 6\bar{N}^2 F_3^2 - 2F_2 F_3^2) A_{12} \\
 &+ 2\bar{N}^3 \bar{\xi} (2\bar{N}^3 F_2 - 6\bar{N} F_2^2 + 3\bar{N}^2 F_3 + F_2 F_3) A_{13} \\
 &+ \bar{N}^2 (\bar{N}^3 - 3\bar{N} F_2 + 2F_3)^2 A_{22} \\
 &+ 2\bar{N}^4 \bar{\xi} (\bar{N}^3 - 3\bar{N} F_2 + 2F_3) A_{23} + \bar{N}^6 \bar{\xi}^2 A_{33}]. \quad (457)
 \end{aligned}$$

It is interesting to compare the results obtained for $\bar{\xi}$ to what was derived for function $\xi(r)$. For example, replacing Δ_{kl} and F_k with their value as functions of \bar{N} and cumulants leads to the following result for the cosmic bias in the 3D case [630]:

$$b_{\bar{\xi}} \simeq \left(0.04 - \frac{1}{\bar{\xi}}\right) \frac{v}{\bar{N}V} + \left(16.5 - 7.6S_3 - \frac{0.53}{\bar{\xi}}\right) \frac{\bar{\xi}v}{V} + \left(3 - 2C_{12} - \frac{1}{\bar{\xi}}\right) \bar{\xi}(\hat{L}). \quad (458)$$

In this equation, valid in the perturbative regime ($|b_{\bar{\xi}}| \ll \sigma_{\bar{\xi}} \ll 1$) and at leading order in v/V one can recognize in the first, second and third terms the discreteness, edge and finite-volume effect contributions, respectively. As expected, the last line is very similar to Eq. (392). Note that the discreteness effect term is rather small and can be neglected in most realistic situations, in agreement with Eq. (392). An alternative calculation of $b_{\bar{\xi}}$ can be found in [328] with similar conclusions.

Fig. 40 displays the cosmic error as a function of scale for factorial moments and cumulants expected in the SDSS. It illustrates how these different estimators perform and shows that the relative error on the cumulants $\bar{\xi}$, S_3 and S_4 is expected to be smaller by 3%, 5% and 15%, respectively, in the scale range $1-10h^{-1}$ Mpc [630].

6.8. Multivariate count-in-cells

The generalization of count-in-cells to the multivariate case is quite straightforward. Here we focus on bivariate statistics, which were used to compute the cosmic error on count-in-cells estimators in Section 6.7.4.

For a pair of cells at position \mathbf{r}_1 and \mathbf{r}_2 separated by distance $r = |\mathbf{r}_1 - \mathbf{r}_2|$, factorial moment correlators [620] are defined as

$$W_{kl}(r_{12}) \equiv \frac{F_{kl} - F_k F_l}{\bar{N}^{k+l}}, \quad (459)$$

$$W_{k0} \equiv \frac{F_{k0}}{\bar{N}^k} \equiv \frac{F_k}{\bar{N}^k}, \quad (460)$$

where the joint factorial moment is given by

$$F_{kl}(r_{12}) \equiv \langle (N)_k (N)_l \rangle. \quad (461)$$

As with factorial moments, F_{kl} estimates joint moments of the smoothed density field:

$$F_{kl}(r_{12}) = \bar{N}^{k+l} \langle [1 + \delta(\mathbf{r}_1)]^k [1 + \delta(\mathbf{r}_2)]^l \rangle. \quad (462)$$

The joint factorial moments and thus the factorial moment correlators can be easily related to the quantities of physical interest, namely the two-point density normalized cumulants—also designed by cumulant correlators [623], C_{pq} [Eq. (348)]. Indeed, as for the monivariate case, one can write

$$F_{kl} = \left(\frac{\partial}{\partial x} \right)^k \left(\frac{\partial}{\partial y} \right)^l \mathcal{P}(x+1, y+1) \Big|_{x=y=0}, \quad (463)$$

$$\mathcal{M}(\bar{N}x, \bar{N}y) = \exp[\mathcal{C}(x, y)] = \mathcal{P}(x+1, y+1), \quad (464)$$

where $\mathcal{P}(x, y)$ is the generating function for bicounts defined previously in Eq. (450), $\mathcal{M}(x, y) = \langle \exp[x\delta(\mathbf{r}_1) + y\delta(\mathbf{r}_2)] \rangle$ is the moment generating function (Section 3.3.3) and $\mathcal{C}(x, y)$ is the two-point density cumulant generating function [Eq. (138)]. For example, the first few cumulant correlators are [623]

$$C_{12} \bar{\xi} \xi = W_{12} - 2\xi, \quad (465)$$

$$C_{13} \bar{\xi}^2 \xi = W_{13} - 3W_{12} - 3W_{20} + 6\xi, \quad (466)$$

$$C_{22} \bar{\xi}^2 \xi = W_{22} - 4W_{12} + 4\xi - 2\xi^2, \quad (467)$$

with $\xi \equiv \xi(r_{12})$. We have used the approximation $W_{11} \simeq \xi$, valid when $r_{12} \gg R$.

An unbiased estimator for the joint factorial moment F_{kl} analogous to Eq. (435) is simply, for a set of P pairs of cells in the catalog separated by distance r and thrown at random (with random direction),

$$\hat{F}_{kl}^P(r) = \frac{1}{2P} \sum_{\text{pairs } (i,j)} [(N_i)_k (N_j)_l + (N_i)_l (N_j)_k]. \quad (468)$$

A possible (biased) estimator for the factorial moments correlators is then, for the same set of cells,

$$W_{kl}^P = \frac{\hat{F}_{kl}^P - \hat{F}_{k0}^P \hat{F}_{l0}^P}{[\hat{F}_{10}^P]^{k+l}} \quad (469)$$

with the definition

$$\hat{F}_{k0}^P \equiv \frac{1}{2^P} \sum_{\text{pairs } (i,j)} [(N_i)_k + (N_i)_l]. \quad (470)$$

At this point, it is interesting to notice again that W_{11} can be used directly as an estimator of the two-point correlation function, if the cell size R is small compared to the separation r (e.g. [503,275]). In that case, the averages are done on sets of pairs of cells in a bin Θ as defined in Section 6.4.1.

Further generalization to higher-order multivariate statistics is trivial. For example, W_{111} can be used to estimate the three-point correlation function (e.g. [275]), W_{1111} to estimate the four-point correlation function (e.g. [226]) and so on.

6.9. Optimal weighting

To optimize the measurements of N -point statistics, the data can be given a varying spatial weight $\omega(\mathbf{r}_1, \dots, \mathbf{r}_N)$ symmetric in its arguments and properly normalized. Furthermore, in realistic redshift surveys, the average number density of galaxies changes with distance r from the observer:

$$\bar{n}_g(r) = \bar{n}_g \phi(r), \quad (471)$$

where $\phi(r) \leq 1$ is the selection function. Now, the estimators defined so far are valid only for *statistically homogeneous* catalogs, i.e. with constant $\bar{n}_g(r)$. One way to avoid this problem is to use volume-limited catalogs. This method consists in extracting from the parent catalog, subsamples of depth R_i such that the apparent magnitude of objects in these catalogs at distance $r = R_i$ from the observer would be larger than the magnitude limit. Such a selection criterion renders the number density of galaxies in the subsamples independent of distance at the price of a significant information loss.⁷⁶ In order to be able to extract all the information from the catalog, it is however possible to correct the estimators for the spatial variation of $\bar{n}_g(r)$. Moreover, the signal to noise can be further improved by appropriate choice of the weight function ω .

The generalization of Eq. (422) reads

$$D^p R^q = \sum \frac{\omega(\mathbf{x}_1, \dots, \mathbf{x}_p, \mathbf{y}_1, \dots, \mathbf{y}_q)}{\phi(\mathbf{x}_1) \cdots \phi(\mathbf{x}_p) \phi(\mathbf{y}_1) \cdots \phi(\mathbf{y}_q)} \Theta(\mathbf{x}_1, \dots, \mathbf{x}_p, \mathbf{y}_1, \dots, \mathbf{y}_q). \quad (472)$$

(We assume that same selection effects are applied to the random catalog R .) Note that the weight could be included in the bin function Θ , but we prefer to separate the idea of spatial weighting from the idea of binning. In principle, the binning can change slightly the nature of the measured statistic A in $A_\Theta \neq A$. Of course, up to now we have assumed that the binned quantity is always close to the quantity of interest, $A_\Theta \simeq A$, but this condition is not absolutely necessary: the binning function Θ can be chosen arbitrarily and determined a priori. Then the statistic of interest becomes A_Θ instead

⁷⁶ However, a number of volume-limited samples can be constructed from the parent catalog to compensate for this.

of the original A . For example, count-in-cells represent a particular choice of the binning function. On the other hand, the spatial weight should not bring any change, i.e. the weighted quantity, should be, after ensemble average, equal to the real value (or at least, very close to it): $\langle \hat{A}_{\theta, \omega} \rangle = A_{\theta}$.

The optimal weight by definition minimizes the cosmic error. In what follows, we assume that \bar{n}_g and $\phi(r)$ are externally determined with very good accuracy. As a result the cosmic error for N -point statistics is given by Eq. (425), with the obvious correction to the functional \mathcal{S}_i with the weights and selection function. The optimal weight can then be found by solving an integral equation for the function ω [291,293,152]. There are several methods to solve this equation, for example by pixelizing the data, thus transforming the integral into a sum. In this way, solving the integral equation corresponds to inverting a matrix. We shall come back to that in the end of this section and in Section 6.11.2.

Otherwise, it has been shown that within the following approximations,

- (1) the considered N -uplets occupy a region \mathcal{R} small enough compared to the size of the catalog that variations of function ϕ in the vicinity of a N -uplet are negligible, $\phi(\mathbf{r}_1) \simeq \dots \simeq \phi(\mathbf{r}_N)$;
- (2) edge effects are insignificant;
- (3) the function ω depends only on position \mathbf{r} of the region \mathcal{R} , i.e. the variations of ω within \mathcal{R} are negligible;

the function $\omega(\mathbf{r})$ that gives the optimal weight for the two-point function (but it is likely to be the case for the higher-order functions) appears to be a functional of the selection function only [291].

Within this simplifying framework,⁷⁷ the solution for the optimal weight is very simple [291,152]

$$\omega(r) \propto 1/\sigma^2(r), \quad (473)$$

where $\sigma(r)$ is the relative cosmic error on the considered statistics in a statistically homogeneous catalog with same geometry and same underlying statistics as the studied one, but with a number of objects such that its number density is $\bar{n}_g \phi(r)$. This result actually applies as well to Fourier space (at least for the power spectrum [212]) and to counts-in-cells statistics [152].

To find the optimal weight, one has to make assumptions about the higher-order statistics in order to compute the cosmic error since the latter depends on up to the $2k$ th order for estimators of k th-order statistics. To simplify the calculation of $\sigma(r)$, the Gaussian limit is often assumed. This is valid only in the weakly non-linear regime and leads to the following weight for the two-point correlation function, commonly used in the literature [410,291,462,196,293]:

$$\omega(r) \propto \frac{1}{[1/\bar{n}_g(r) + J(r)]^2}, \quad (474)$$

where

$$J(r) = \int_{r' \leq r} d^{\mathcal{D}} \mathbf{r}' \xi(r'). \quad (475)$$

⁷⁷ Hamilton [293,294] developed a general formalism for optimizing the measurement of the two-point correlation function in real and Fourier space, relying on the covariance matrix of the statistic $\langle \delta(\mathbf{r}_i) \delta(\mathbf{r}_j) \rangle$, which would correspond to the binning function $\Theta(\mathbf{r}_1, \mathbf{r}_2) = \delta_D(\mathbf{r}_1) \delta_D(\mathbf{r}_2)$. He proposed a way of computing the optimal sampling weight without requiring these simplifying assumptions.

In Fourier space the result is [212]

$$\omega(r) \propto \frac{1}{[1/V\bar{n}_g(r) + P(k)]^2}, \quad (476)$$

a result that can be easily guessed from Eq. (415). This equation is valid for $\{k, \Delta k\} \gg 1/L$, where L is the size of the catalog in the smallest direction and Δk is the width of the considered bin.

Note that the function $\omega(r)$ is of pairwise nature. It corresponds to weighting the data with

$$n_g(\mathbf{r}) \rightarrow n_g(\mathbf{r})\sqrt{\omega(r)}. \quad (477)$$

Now, we turn to a more detailed discussion of optimal weighting in count-in-cell statistics. The problem of finding the optimal sampling weight was studied in [152]. Similar to Eq. (472), the weighted factorial moment estimator reads

$$\hat{F}_k^C = \frac{1}{C} \sum_{i=1}^C \frac{(N_i)_k \omega(\mathbf{r}_i)}{[\phi_R(r_i)]^k}, \quad (478)$$

where $\phi_R(r)$ is the average of the selection function over a cell.

To simplify the writing of the cosmic error as a function of the sampling weight, the variations of the function ω and of the selection function are assumed to be negligible within the cells, which is equivalent to points (1) and (3) above. Then the relative cosmic error $\sigma_{F_k}[\omega, \phi] = (\Delta \hat{F}_k / F_k)^2$ is

$$\sigma_{F_k}^2[\omega, \phi] = \sigma_F^2[\omega] + \sigma_E^2[\omega] + \sigma_D^2[\omega, \phi], \quad (479)$$

where the finite volume, edge effect and discreteness contributions read, respectively,

$$\sigma_F^2[\omega] = \frac{\sigma_F^2}{\bar{\xi}(\hat{L})\hat{V}} \int_{\hat{V}} d^3\mathbf{r}_1 d^3\mathbf{r}_2 \omega(\mathbf{r}_1)\omega(\mathbf{r}_2)\xi(r_{12}), \quad (480)$$

$$\sigma_E^2[\omega] = \frac{\sigma_E^2}{\hat{V}} \int_{\hat{V}} d^3\mathbf{r} \omega(\mathbf{r}), \quad (481)$$

$$\sigma_D^2[\omega, \phi] = \frac{1}{\hat{V}} \int_{\hat{V}} d^3\mathbf{r} \omega^2(\mathbf{r}) \sigma_D^2(r). \quad (482)$$

In these equations, there are terms such as $\sigma_F^2 = \sigma_F^2[1]$ or $\sigma_E^2 = \sigma_E^2[1]$. They correspond to the finite volume and edge effect errors in the case of homogeneous sampling weight. They do not depend on the number density and are given by analytical expressions in Appendix F. The term $\sigma_D^2(r)$ is similar, but there is a supplementary r dependence because the average count \bar{N} is proportional to the selection function ϕ .

Using Lagrange multipliers, it is easy to write the following integral equation which determines the optimal weight [152]:

$$\frac{\sigma_F^2}{\bar{\xi}(\hat{L})\hat{V}} \int_{\hat{V}} d^3\mathbf{u} \omega(\mathbf{u})\xi(|\mathbf{r} - \mathbf{u}|) + [\sigma_E^2 + \sigma_D^2(r)]\omega(\mathbf{r}) + \lambda = 0. \quad (483)$$

The constant λ is determined by appropriate normalization of the weight function

$$\frac{1}{\hat{V}} \int_{\hat{V}} d^3\mathbf{r} \omega(\mathbf{r}) = 1. \quad (484)$$

The solution of this integral equation can be found numerically. However, approximation (473) was found to be excellent, i.e. almost perfectly minimizes the cosmic error [152].

Using the leading order theory of propagation of errors in Section 6.7.3, it is easy to see that these calculations apply as well to the variance and the cumulants, provided that errors are small enough: in Eqs. (436)–(438), \hat{F}_k would be computed with Eq. (478), using the sampling weight minimizing the cosmic error of the cumulant of interest.

This result shows as well that for a statistically homogeneous catalog, a weight unity $\omega=1$ is very close to optimal in most practical cases for count-in-cell statistics. This statement, of course, is not necessarily true for N -point correlation functions, particularly if the catalog presents a complicated geometry. In that case, the use of a weight might help to correct for edge effects at large scales, although the LS estimator and its generalization to higher order perform already well in this respect with a uniform weight. For traditional counts-in-cells estimators, the finite extension of the cells prevents from correcting for edge effects. This is actually the main weakness of these statistics compared to the N -point correlation functions, and often the latter are preferred to the former, particularly when the geometry of the catalog is complicated by the presence of numerous masks which reduce considerably the range of scales available to counts-in-cells.

Finally, it is worth noting the following point: the optimal weight is actually difficult to compute because it requires knowledge of statistics of order $l \leq 2k$ for an estimator of order k . Therefore, the Gaussian limit, given by Eqs. (474) and (476) for functions $\xi(r)$ and $P(k)$, respectively, was widely used in the literature. However, this is rigorously valid only in the weakly non-linear regime where the shot noise error is likely to be negligible, implying a simple, uniform weight to be nearly optimal, unless the catalog is very diluted. Discreteness errors are less of a concern with modern surveys under construction, such as the 2dFGRS or the SDSS.

Furthermore, it was noticed in [152] that the traditional volume-limited sample method does almost as good as a single optimized measurement extracting all the information from the catalog, if the depth of the subsample, R_i , is chosen such that for the scale considered signal to noise is approximately maximal. Of course, estimating the cosmic error is still a problem, but the advantage of the volume-limited approach is that prior determination of the selection function is not necessary, which simplifies considerably the analysis.

6.10. Cosmic distribution function and cross-correlations

6.10.1. Cosmic distribution function and likelihood

For a set of (possibly biased) estimators, $\mathbf{f} = \{\hat{f}_k\}_{k=1,K}$, let us define the covariance matrix as $C_{kl} = \text{Cov}\{\hat{f}_k, \hat{f}_l\}$. The extra-diagonal terms can be correlations between a given estimator (e.g. of the power spectrum) at different scales (as in Section 6.5.4), between different estimators at the same scale (e.g. factorial moments, see Section 6.10.2), or in general different estimators at different scales. Knowledge of these cross-correlations can in fact help to better constrain theories with observations because they bring more information on the shape of the cosmic distribution function.

As mentioned in Section 6.2, the cosmic distribution function Υ is the probability distribution for an estimator *given a theory* (or class of theories parametrized in some convenient form), i.e. $\Upsilon = \Upsilon(\hat{\mathbf{f}}|\text{theory})$ is the probability of measuring $\hat{\mathbf{f}}$ in a finite galaxy catalog given a theory. Knowledge of $\Upsilon(\hat{\mathbf{f}}|\text{theory})$ allows one to extract constraint on cosmological parameters from the data through maximum likelihood analysis, where the *likelihood function* is given by the cosmic distribution function thought as a function of the parameters that characterize the theory (with $\hat{\mathbf{f}}$ replaced in terms of the observed data).

In particular, if the cosmic distribution function Υ is Gaussian, it is entirely determined once the covariance matrix \mathbf{C} is known:

$$\Upsilon(\hat{\mathbf{f}}|\mathbf{C}, \mathbf{f}, \mathbf{b}) = \frac{1}{\sqrt{(2\pi)^K |\mathbf{C}|}} \exp \left[-\frac{1}{2} \sum_{k,l} \delta \hat{f}_k C_{kl}^{-1} \delta \hat{f}_l \right], \quad (485)$$

where \mathbf{C}^{-1} and $|\mathbf{C}|$ are, respectively, the inverse and the determinant of the covariance matrix, \mathbf{f} is the true value of the statistics in question ($\mathbf{f} = \langle \hat{\mathbf{f}} \rangle$ for unbiased estimators) and \mathbf{b} a vector accounting for possible cosmic bias. Both \mathbf{C} and \mathbf{f} (and \mathbf{b} if non-zero) are calculated from theoretical predictions as a function of cosmological parameters.

It is very important to note that the Gaussian assumption for Υ is, in general, different from assuming that the density field is Gaussian unless the estimator $\hat{\mathbf{f}}$ corresponds to the density contrast.⁷⁸ For this reason, Eq. (485) is not necessarily a good approximation for estimators that are not linear in the density contrast even if the underlying statistic of the density field is Gaussian. We shall come back to this point in Section 6.10.3.

Why is it useful to take as \mathbf{f} non-linear functions of the density contrast? The problem is that the assumption of Gaussianity for the density field itself is very restrictive to deal with galaxy clustering: it does not include information on higher-order moments which arise due to e.g. non-linear evolution, non-linear galaxy bias or primordial non-Gaussianity. Since there is no general expression for the multi-point PDF of the density field which describes its non-Gaussian shape,⁷⁹ one must resort to a different approach. The key idea is that taking \mathbf{f} to be a statistic⁸⁰ of the density field, it is possible to work in a totally different regime. Indeed, when the cosmic error is sufficiently small, there must be many independent contributions to $\hat{\mathbf{f}}$ so that, by the central limit theorem, its cosmic distribution function should approach Gaussianity.⁸¹ On the other hand, the cosmic error becomes large when probing large scales, where there are not many independent samples; in this case, assumption of a Gaussian density field plus the non-linear transformation involved in $\hat{\mathbf{f}}$ leads to a useful guess about the asymptotic behavior of Υ . In practice, the specific shape of Υ must be *computed* for a given set of theories, and the limit of validity of the asymptotic forms discussed above should be carefully checked, as discussed further in Section 6.10.3.

The remainder of this section is organized as follows. In Section 6.10.2 we discuss about correlations between different statistics. As an example, we show how knowledge of the number of objects

⁷⁸ In this case Υ is proportional to the density PDF.

⁷⁹ The Edgeworth expansion, Eq. (144), in principle provides a way to accomplish this [5]. In practice, however, its regime of validity is very restricted.

⁸⁰ These are non-linear functions of the data, e.g. the power spectrum is quadratic.

⁸¹ Note that, in contrast to the PDF of the density field, this limit is usually approached *at small scales*, we shall discuss examples below in Section 6.10.3.

in a galaxy catalog can be used to reduce the error bar on the measurement of the two-point correlation function. Then, in Section 6.10.3, we address the problem of non-Gaussianity of the cosmic distribution function.

6.10.2. Cross-correlations between different statistics

An important kind of cross-correlation is given by that between statistics of different kind. For example, the calculation leading to Eq. (400) is a *conditional* average with the constraint that the average number density is equal to the observed one:

$$(\Delta \hat{\xi} | \bar{n}_g)^2 \equiv \langle \hat{\xi}^2 | \bar{n}_g \rangle - \langle \hat{\xi} | \bar{n}_g \rangle^2 \quad (486)$$

$$= \frac{\int \xi^2 \Upsilon(\xi, \bar{n}_g) d\xi}{\int \Upsilon(\xi, \bar{n}_g) d\xi} - \left[\frac{\int \xi \Upsilon(\xi, \bar{n}_g) d\xi}{\int \Upsilon(\xi, \bar{n}_g) d\xi} \right]^2. \quad (487)$$

The knowledge of this supplementary information decreases the expected error on the measurement of $\xi(r)$ and provides better constraints on the models. The calculation of Bernstein leading to Eq. (395) does not make use of the fact that \bar{n}_g can be measured separately:

$$(\Delta \hat{\xi})^2 = \langle \hat{\xi}^2 \rangle - \langle \hat{\xi} \rangle^2 \quad (488)$$

$$= \int \xi^2 \Upsilon(\xi, \bar{n}_g) d\xi d\bar{n}_g - \left[\int \xi \Upsilon(\xi, \bar{n}_g) d\xi d\bar{n}_g \right]^2 \quad (489)$$

and therefore slightly overestimates the error on $\xi(r)$ as emphasized in [393]. For example, if the function Υ is Gaussian, we have

$$(\Delta \hat{\xi} | \bar{n}_g)^2 = (\Delta \hat{\xi})^2 [1 - \rho_{\xi, \bar{n}_g}^2], \quad (490)$$

where the correlation coefficient ρ_{AB} is defined for estimators \hat{A} and \hat{B} as

$$\rho_{AB} \equiv \frac{\langle (\hat{A} - \langle \hat{A} \rangle)(\hat{B} - \langle \hat{B} \rangle) \rangle}{\Delta \hat{A} \Delta \hat{B}}. \quad (491)$$

From this simple result, we see that joint measurement of (theoretically) *more correlated or anti-correlated* statistics brings better constraints on the underlying theory.

In [630] and as described in Section 6.7.4, cross-correlations between factorial moments are computed analytically at fixed scale. From the theory of propagation of errors, it is straightforward to compute cross-correlations between other count-in-cells statistics of physical interest, such as average count \bar{N} , variance $\bar{\xi}$ and cumulants S_p . Theoretical calculations and measurements in numerical simulations [630,153] show that, for realistic galaxy catalogs such as the SDSS, \bar{N} and $\bar{\xi}$ are not, in general, strongly correlated, and similarly for correlations between \bar{N} and higher-order statistics. Interestingly, $\bar{\xi}$ and S_3 are not very strongly correlated, but S_3 and S_4 are. Actually, in general and as expected, the degree of correlation between two statistics of orders k and l decreases with $|k - l|$.

6.10.3. Validity of the Gaussian approximation

We now discuss the validity of the Gaussian approximation, Eq. (485), for the cosmic distribution function. To illustrate the point, we take two examples, the first one about count-in-cells statistics, the second one about the power spectrum and bispectrum.

Exhaustive measurements in one of the Hubble volume simulations [631] show that for count-in-cell statistics, $\Upsilon(\hat{A})$ is approximately Gaussian if $\Delta\hat{A}/A \lesssim 0.2$. Therefore, at least for count-in-cells, Gaussianity is warranted only if the errors are small enough. When the cosmic errors become significant, the cosmic distribution function becomes increasingly skewed, developing a tail at large values of \hat{A} [631]. This result applies to most counts-in-cells estimators ($\hat{P}_N, \hat{F}_k, \hat{\xi}, \hat{S}_p$). One consequence is that the most likely value is below the average, resulting in an *effective cosmic bias*, even for unbiased statistics such as factorial moments: typically, the measurement of a statistic \hat{A} in a finite catalog is likely to underestimate the real value, except in some rare case where it will overestimate it by a larger amount.⁸² To take into account the asymmetry in the shape, it was proposed in [631] to use a generalized version of the lognormal distribution, which describes very well the shape of function $\Upsilon(\hat{A})$ for a single statistic, as illustrated in Fig. 41:

$$\Upsilon(\hat{A}) = \frac{s}{\Delta A [s(\hat{A} - A)/\Delta A + 1] \sqrt{2\pi\eta}} \exp\left(-\frac{\{\ln[s(\hat{A} - A)/\Delta A + 1] + \eta/2\}^2}{2\eta}\right), \quad (492)$$

$$\eta = \ln(1 + s^2), \quad (493)$$

where s is an adjustable parameter. It is fixed by the requirement that the analytical function, Eq. (492), have identical average, variance and skewness $S_3 = 3 + s^2$, as the measured $\Upsilon(\hat{A})$.

However, the generalization of Eq. (492) to multivariate cosmic distribution functions is not easy, although feasible at least in some restricted cases (e.g. see [585]). An alternate approach would employ a multivariate Edgeworth expansion [5].

Since the Gaussianity of the cosmic distribution function mainly depends on the variance of the statistic under consideration, it is expected that for surveys where errors are not negligible, Gaussianity is not a good approximation. Fig. 42 illustrates this for IRAS surveys in the case of the power spectrum and bispectrum [565], as a function of normalized variables, $\delta A/\Delta A \equiv (\hat{A} - A)/\langle(\hat{A} - A)^2\rangle^{1/2}$. For the bispectrum, this choice of variable makes the cosmic distribution function approximately independent of scale and configuration.

The left panel of Fig. 42 shows the power spectrum cosmic distribution function as a function of scale, from least to most non-Gaussian, scales are $k/k_f = 1-10$, $k/k_f = 11-20$, $k/k_f = 21-30$, $k/k_f = 31-40$, where $k_f = 0.005 h/\text{Mpc}$. As expected, non-Gaussianity is significant at large scales, as there are only a few independent modes (due to the finite volume of the survey), and thus the power spectrum PDF is chi-squared distributed. As smaller scales are considered, averaging over more modes leads to a more Gaussian distribution, although the convergence is slow since the contributing modes are strongly correlated due to shot noise.

The right panel in Fig. 42 shows a similar plot for the bispectrum. In sparsely sampled surveys such as QDOT, deviation from Gaussianity can be very significant. In a large volume-limited sample of $600 \text{ Mpc}/h$ radius with many galaxies (dotted curve), Gaussianity becomes an excellent approximation, as expected. The cosmic distribution function for χ^2 initial conditions was also calculated in [565]; in this case, non-Gaussianity is significant even for large-volume surveys, and thus must be taken into consideration in order to properly constrain primordial non-Gaussianity [567,211].

⁸² This is, of course, analogous to non-Gaussianity in the density PDF. Positive skewness means that the most likely value is to *underestimate* the mean, see Eq. (230). To compensate for this there is a rare tail at large values compared to the mean, see e.g. Fig. 20.

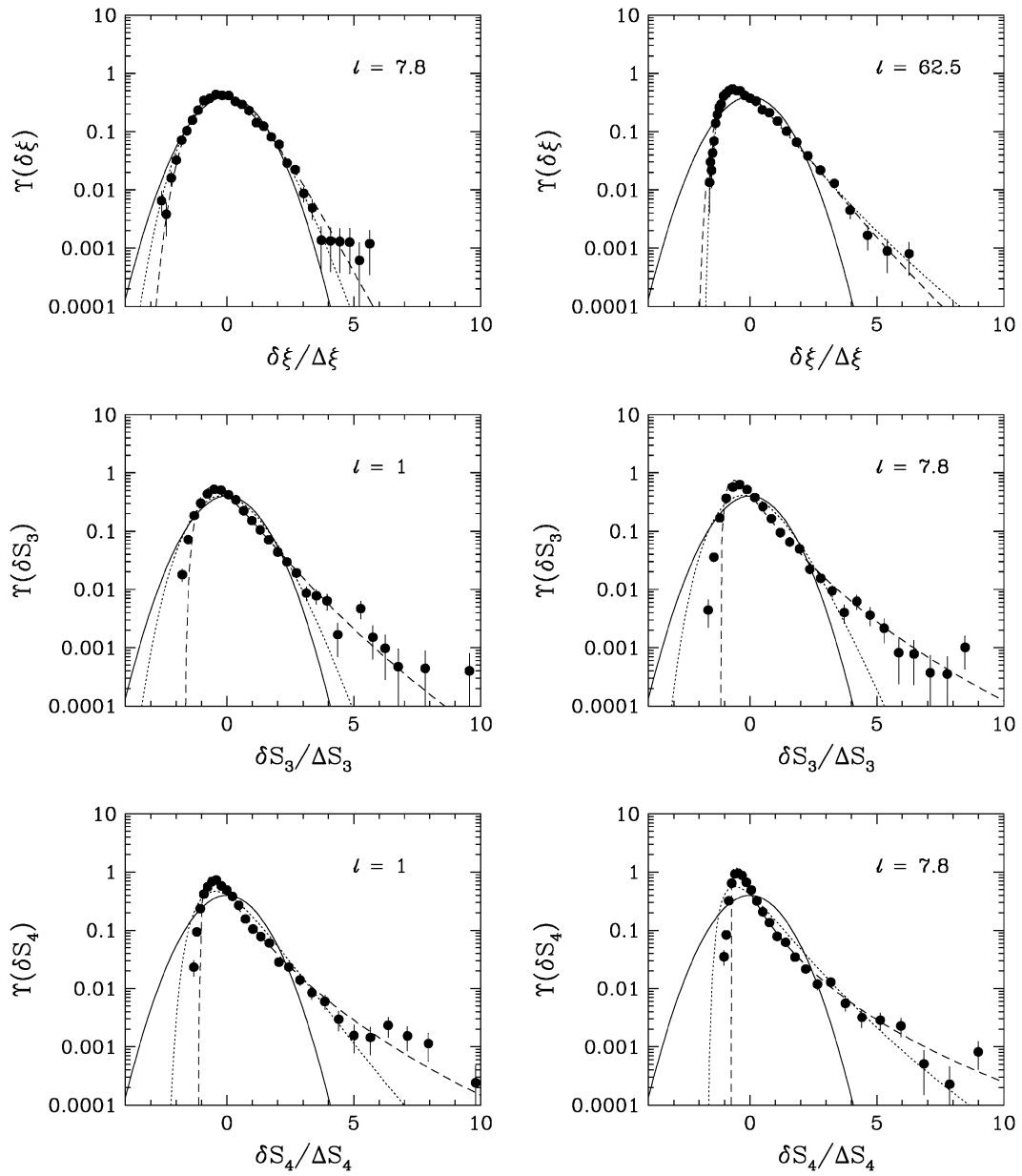


Fig. 41. The cosmic distribution function of measurements $\Upsilon(\hat{\xi})$ (upper line of panels), $\Upsilon(\hat{S}_3)$ (middle line of panels) and $\Upsilon(\hat{S}_4)$ (lower line of panels) measured from a distribution of subsamples extracted from a Hubble volume simulation (see end of Section 6.7.4 for more details). The scale of the measurements, either $R = 1, 7.8$ or $62.5 h^{-1}$ Mpc, is indicated on each panel. The solid, dotted and dash curves correspond to the Gaussian, lognormal and generalized lognormal [Eq. (492)] distributions, respectively. With the choice of the coordinate system, the magnitude of the cosmic error does not appear directly, but is reflected indirectly by the amount of skewness of the lognormal distribution.

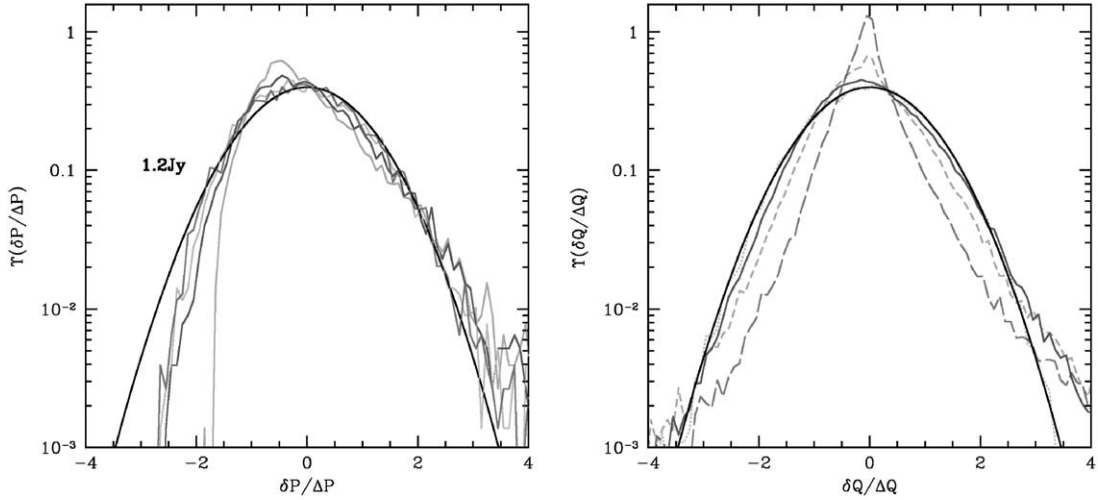


Fig. 42. Left panel: Power spectrum cosmic distribution function in an IRAS 1.2Jy-like survey as a function of scale in logarithmic scale, smooth solid line denotes a Gaussian distribution. From least to most non-Gaussian, scales are $k/k_f = 1-10$, $k/k_f = 11-20$, $k/k_f = 21-30$, $k/k_f = 31-40$, where $k_f = 0.005 h/\text{Mpc}$. Right panel: Cosmic distribution function of $\delta Q/\Delta Q \equiv (Q - \bar{Q})/\Delta Q$ for different surveys in models with Gaussian initial conditions: second-order Lagrangian PT with 256^3 objects in a volume of $600 \text{ Mpc}/h$ radius (dotted), IRAS 1.2Jy (solid), IRAS 2Jy (dashed), IRAS QDOT (long-dashed). The smooth solid curve is a Gaussian distribution.

6.11. Optimal techniques for Gaussian random fields

Up to now, we have restricted our discussion to a particular subset of estimators used commonly in the literature, which apply equally well to two-point and higher-order statistics. To give account of recent developments, we now reinvestigate the search for optimal estimators in the framework of Gaussian random fields. That is, the cosmic distribution function, with estimators $\hat{\mathbf{f}}$ that *will* be taken as density contrasts (measured in pixels or their equivalent in some space of functions, such as spherical harmonics), will be assumed to be Gaussian. As discussed above, this approach is only justifiable to obtain estimates of the power spectrum (or two-point correlation function) at the largest scales, where Gaussianity becomes a good approximation.

First we recall basic mathematical results about minimum variance and maximum likelihood estimators (Section 6.11.1). In Section 6.11.2, we discuss optimal weighting for two-point statistics taking into account the full covariance matrix (compare to Section 6.9), and in Section 6.11.3 we briefly address techniques for obtaining uncorrelated estimates of the power spectrum, comparing with results discussed in previous sections when relevant. Finally, we briefly describe the Karhunen–Loève transform, useful for compressing large amounts of data expected in current and forthcoming surveys (Section 6.11.4).

6.11.1. Maximum likelihood estimates

The basic results given here are well known in statistical theory [610,690]. For more details and applications to optimal measurements of the power spectrum in cosmological data sets see e.g. [646,647,80,293].

Let us assume that we have at our disposal some data $\hat{\mathbf{x}}$, say, a vector of dimension N with the cosmic distribution function $\Upsilon(\hat{\mathbf{x}})$, which is Gaussian and can be expressed explicitly as a function of $\hat{\mathbf{x}}$ and a set of unknown parameters \mathbf{f} , which we aim to estimate, given our data. When thought as a function of the parameters \mathbf{f} , $\Upsilon(\mathbf{f})$ is usually known as the *likelihood function*.⁸³ The corresponding estimators, $\hat{\mathbf{f}} = (\hat{f}_1, \dots, \hat{f}_K)$, $K \leq N$, are sought in the space of functions of the data $\hat{\mathbf{x}}$. The problem of finding an optimal estimator $\hat{\mathbf{f}}$ can be formally approached at least in two ways, the first one consisting in minimizing the cosmic error on $\hat{\mathbf{f}}$, the second one consisting in maximizing the likelihood.

We restrict ourselves to unbiased estimators,

$$\langle \hat{\mathbf{f}} \rangle \equiv \int d^N \hat{\mathbf{x}} \Upsilon(\hat{\mathbf{x}}|\mathbf{f}) \hat{\mathbf{f}}(\hat{\mathbf{x}}) = \mathbf{f} . \quad (494)$$

The search for the first kind of optimal estimator, already discussed in Section 6.9, consists in minimizing the cosmic error

$$\Delta^2 f_k = \langle (\hat{f}_k - f_k)^2 \rangle , \quad (495)$$

given constraint (494). It is useful at this point to assume that the likelihood function is sufficiently smooth and to introduce the so-called Fisher information matrix

$$F_{kl} \equiv \left\langle \frac{\partial^2 [-\log \Upsilon(\mathbf{f})]}{\partial f_k \partial f_l} \right\rangle \left\langle \frac{\partial \log \Upsilon(\mathbf{f})}{\partial f_k} \frac{\partial \log \Upsilon(\mathbf{f})}{\partial f_l} \right\rangle . \quad (496)$$

Let us assume that the matrices \mathbf{F} and the covariance matrix \mathbf{C} defined by $C_{kl} \equiv \text{Cov}(f_k, f_l) = \langle \delta \hat{f}_k \delta \hat{f}_l \rangle$ are positive definite. From the Cauchy–Schwarz inequality one gets the so-called Cramér–Rao inequality

$$(\Delta f_k)^2 F_{kk} \geq 1 \quad (497)$$

so that the inverse of the Fisher matrix can be thought as the minimum errors that one can achieve. Through a change of variable this inequality can be generalized in

$$(\mathbf{a}^t \cdot \mathbf{C} \cdot \mathbf{a})(\mathbf{b}^t \cdot \mathbf{F} \cdot \mathbf{b}) \geq (\mathbf{a}^t \cdot \mathbf{b})^2 , \quad (498)$$

where \mathbf{a} and \mathbf{b} are two sets of constants. It implies

$$|\mathbf{C}| \geq \frac{1}{|\mathbf{F}|} . \quad (499)$$

An estimator $\hat{\mathbf{f}}$ which obeys the equality in Eq. (498) or (499) is called *minimum variance bound* (MVB). This can happen if and only if the estimator $\hat{\mathbf{f}}$ can be expressed as a linear function of the derivative of log-likelihood function with respect to the parameters:

$$\left(\frac{\partial \log \Upsilon}{\partial \mathbf{f}} \right)^t \cdot \mathbf{b} = g(\mathbf{f})(\hat{\mathbf{f}} - \mathbf{f})^t \cdot \mathbf{a} , \quad (500)$$

where the constant of proportionality $g(\mathbf{f})$ might depend on the parameters but not on the data $\hat{\mathbf{x}}$. As a result, for an arbitrary choice of the parameters \mathbf{f} , minimum variance unbiased estimators are not

⁸³ Therefore, the assumption of a Gaussian density field means $\Upsilon(\hat{\mathbf{x}})$ as a function of $\hat{\mathbf{x}}$ is Gaussian, whereas in the limit that a large number of uncorrelated data contribute, $\Upsilon(\mathbf{f})$ becomes Gaussian.

necessarily MVB. The second way of seeking an optimal estimator consists in maximizing directly the likelihood function in the space of parameters, $\mathbf{f} \rightarrow \hat{\mathbf{f}}$. The goal is to find $\hat{\mathbf{f}}_{\text{ML}}$ such that

$$\Upsilon(\hat{\mathbf{x}})|_{\mathbf{f}=\hat{\mathbf{f}}_{\text{ML}}(\hat{\mathbf{x}})} \geq \Upsilon(\hat{\mathbf{x}})|_{\mathbf{f}} \quad (501)$$

for any possible value of \mathbf{f} . A practical, sufficient but not necessary condition is given by the solution of the two sets of equations

$$\frac{\partial \log \Upsilon}{\partial \mathbf{f}} = 0, \quad (502)$$

$$\frac{\partial^2 \log \Upsilon}{\partial f_k \partial f_l} < 0. \quad (503)$$

The solution of Eq. (501), if it exists, does not lead necessarily to an unbiased estimator nor a minimum variance estimator. But if by chance the obtained ML estimator is unbiased, then it minimizes the cosmic error. Moreover, if there is an MVB unbiased estimator, it is given by the ML method. Note that in the limit that large number of uncorrelated data contributes, the cosmic distribution function tends to a Gaussian and the ML estimator is asymptotically unbiased and MVB. In that regime, the cosmic cross-correlation matrix of the ML estimator is very well approximated by the inverse of the Fisher information matrix

$$C_{kl} = \text{Cov}(f_k, f_l) = \langle \delta \hat{f}_k \delta \hat{f}_l \rangle \simeq (\mathbf{F}^{-1})_{k,l}. \quad (504)$$

On the other hand, from the Gaussian assumption for $\Upsilon(\hat{\mathbf{x}})$, it follows that the ML estimator for the power spectrum [$\hat{P}(k_\alpha) \equiv \hat{f}_\alpha$] is the solution of

$$\hat{f}_\alpha = \frac{1}{2} F_{\alpha\beta}^{-1} \frac{\partial C_{ij}}{\partial f_\beta} [C^{-1}]_{ik} [C^{-1}]_{jl} (\delta_k \delta_l - N_{kl}) \quad (505)$$

(where δ_k denotes the density contrast at \mathbf{r}_k) for which the estimate is equal to the prior, $\hat{\mathbf{f}} = \mathbf{f}$. That is, in order to obtain the ML estimator, one starts with some prior power spectrum \mathbf{f} , then finds the estimate $\hat{\mathbf{f}}$, puts this back into the prior, and iterates until convergence. In Eq. (505), the Fisher matrix is obtained from Eq. (496),

$$F_{\alpha\beta} = \frac{1}{2} \frac{\partial C_{ij}}{\partial f_\alpha} [C^{-1}]_{ik} [C^{-1}]_{jl} \frac{\partial C_{kl}}{\partial f_\beta}, \quad (506)$$

the covariance matrix $C_{ij} = \xi_{ij} + N_{ij}$ contains a term due to clustering (given by the two-point correlation function at separation $|\mathbf{r}_i - \mathbf{r}_j|$, ξ_{ij}), and a shot noise term $N_{ij} \equiv \bar{n}_i \delta_D(\mathbf{r}_i - \mathbf{r}_j)$. Applications of the ML estimator to measurements of the 2D galaxy power spectrum was recently done for the APM [203] and EDSGC [331] surveys (see Section 8.2.2).

6.11.2. Quadratic estimators

In reality it is in general difficult to express explicitly the likelihood function in terms of the parameters. In addition, even if we restrict to the case where the parameters are given by the power spectrum as a function of scale as discussed in the previous section, one must iterate numerically to obtain the ML estimates, and their probability distribution also must be computed numerically in order to provide error bars.⁸⁴ As a result, a useful approach is to seek an optimal estimator,

⁸⁴ However, see [81] for an analytic approximation in the case of the 2D power spectrum using an offset lognormal.

unbiased and having minimum variance, by restricting the optimization to a subspace of estimators, as discussed in Section 6.9. Of course, this method is not restricted to the assumption of Gaussianity, provided that the variance is calculated including non-Gaussian contributions. It turns out there is an elegant solution to the problem [293,296], which in its exact form is unfortunately difficult to implement in practice, but it does illustrate the connection to ML estimate (505) in the Gaussian limit, and also provides a generalization of the standard optimal weighting results, Eqs. (474) and (476), to include non-Gaussian (and non-diagonal) elements of the covariance matrix.

Since the power spectrum is by definition a quadratic quantity in the overdensities, it is natural to restrict the search to quadratic functions of the data. In this framework, the unbiased estimator⁸⁵ of the power spectrum having minimum variance reads [293,296]

$$\hat{f}_\alpha = F_{\alpha\beta}^{-1} \frac{\partial C_{ij}}{\partial f_\beta} [\tilde{C}^{-1}]_{ijkl} (\delta_k \delta_l - \hat{N}_{kl}), \quad (507)$$

where the variance is given by Eq. (504) and the Fisher matrix by Eq. (506) replacing $\frac{1}{2}[C^{-1}]_{ik}[C^{-1}]_{jl}$ with $[\tilde{C}^{-1}]_{ijkl}$, where

$$\tilde{C}_{ijkl} = \langle (\delta_i \delta_j - \hat{N}_{ij} - \xi_{ij})(\delta_k \delta_l - \hat{N}_{kl} - \xi_{kl}) \rangle \quad (508)$$

is the (shot noise subtracted) power spectrum covariance matrix. Here \hat{N}_{ij} denotes the “actual” shot noise, meaning that the self-pairs contributions to ξ_{ij} are not included, see [296] for details. In the Gaussian limit, $[\tilde{C}^{-1}]_{ijkl} \rightarrow \frac{1}{2}[C^{-1}]_{ik}[C^{-1}]_{jl}$ (symmetrized over indices k and l) and the minimum variance estimator, Eq. (507), reduces to ML estimator, Eq. (505), assuming iteration to convergence is carried out as discussed above. If the iteration is not done, the estimator remains quadratic in the data, and it corresponds to using Eq. (505) with a fixed prior; this should be already a good approximation to the full ML estimator, otherwise it would indicate that the result depends sensitively on the prior and thus there is no significant information coming from the data. The use of such quadratic estimators in the Gaussian limit to measure the galaxy power spectrum is discussed in detail in [648], see also [647,646,80]. Extension to minimum variance cubic estimators for the angular bispectrum in the Gaussian limit is considered in [307,245].

Note that, the full minimum variance estimator involves inverting a rank 4 matrix, a very demanding computational task, which however simplifies significantly in the Gaussian limit where \tilde{C} factorizes. Another case in which the result becomes simpler is the so-called FKP limit [212], where the selection function $\bar{n}_g(\mathbf{r})$ can be taken as locally constant, compared to the scale under consideration. This becomes a good approximation at scales much smaller than the characteristic size of the survey, which for present surveys is where non-Gaussian contributions become important, so it is a useful approximation. In this case the minimum variance pair weighting for a pair ij is only a function of the separation α of the pair, not their position or orientation, since \bar{n}_i and \bar{n}_j are assumed to be constants locally. As a result, the power spectrum covariance matrix can be written in terms of a two by two reduced covariance matrix, which although not diagonal due to non-Gaussian contributions, becomes so in the Gaussian limit, leading to the standard result, Eq. (476). We refer the reader to [296] for more details.

⁸⁵ This is assuming that the mean density is perfectly known.

6.11.3. Uncorrelated error bars

Clearly, minimum variance estimates can be deceptive if correlations between them are substantial. Ideally one would like to obtain not only an optimal estimator (with minimum error bars), but also estimates which are uncorrelated (with diagonal covariance matrix), like in the case of the power spectrum of a Gaussian field in the infinite volume limit. Once the optimal (or best possible) estimator $\hat{\mathbf{f}}$ is found, it is possible to work in a representation where the cosmic covariance matrix \mathbf{C} becomes diagonal,

$$\mathbf{C} \cdot \Psi_j = \lambda_j \Psi_j, \quad (509)$$

where the eigenvectors Ψ_j form an orthonormal basis. A new set of estimators can be defined as

$$\hat{\mathbf{g}} \equiv \Psi^{-1} \cdot \hat{\mathbf{f}}, \quad (510)$$

which are statistically orthogonal

$$\langle \delta \hat{g}_i \delta \hat{g}_j \rangle = \lambda_i \delta_{ij} = \Psi_i^\dagger \cdot \mathbf{C} \cdot \Psi_i \delta_{ij}. \quad (511)$$

These new estimators can in principle be completely different from the original set, but if by chance the diagonal terms of \mathbf{C} are dominant, then we have $\hat{\mathbf{g}} \simeq \hat{\mathbf{f}}$. In fact, if one takes the example of the two-point correlation function (or higher order) in case the galaxy number density is known, using the new estimator $\hat{\mathbf{g}}$ is equivalent to changing the binning function Θ defined previously to a more complicated form. Among those estimators which are uncorrelated, it is however important to find the set $\hat{\mathbf{g}}$ such that the equivalent binning function is positive and compact in Fourier space and $\langle \hat{\mathbf{g}} \rangle \simeq \mathbf{f}$, in order to keep the interpretation of the power in this new representation as giving the power centered about some well-defined scale [294,296].

The above line of thoughts can in fact be pushed even further by applying the so-called “pre-whitening” technique to $\hat{\mathbf{f}}$: if $\hat{\mathbf{f}}$ is decomposed in terms of signal plus noise, pre-whitening basically consists in multiplying $\hat{\mathbf{f}}$ by a function h such that the noise becomes white or constant. If the noise is uncorrelated, this method allows one to diagonalize simultaneously the covariance matrix of the signal and the noise. When non-Gaussian contributions to the power spectrum covariance matrix are included, however, such a diagonalization is not possible anymore. However, in the FKP approximation, as described in the previous section, it was shown that an approximate diagonalization (where two of the contributions coming from two- and four-point functions are exactly diagonal, whereas the third coming from the three-point function is not) works extremely well, at least when non-Gaussianity is modeled by the hierarchical ansatz [296]. The quantity whose covariance matrix has these properties corresponds to the so-called pre-whitened power spectrum, which is most easily written in real space [296]:

$$\hat{\xi}(r) \rightarrow \frac{2\hat{\xi}(r)}{1 + [1 + \xi(r)]^{1/2}}. \quad (512)$$

Note that in the linear regime, $\hat{\xi}(k)$ reduces to the linear power spectrum; however, unlike the non-linear power spectrum, $\hat{\xi}(k)$ has almost diagonal cosmic covariance matrix even for non-linear modes. More details on the theory and applications to observations can be found in e.g. [296,297] and [298,487,299], respectively.

6.11.4. Data compression and the Karhunen–Loève transform

A problem faced with modern surveys such as the 2dFGRS and SDSS, is that the data set $\hat{\mathbf{x}}$ becomes quite large for “brute force” application of estimation techniques. Before statistical treatment of the data as discussed in the previous sections, it might be necessary to find a way to reduce their size, but keeping as much information as possible. The (discrete) Karhunen–Loève transform (KL) provides a fairly simple method to do that (see e.g. [680,646] and references therein for more technical details and e.g. [487,443] for practical applications to observations). Basically, the idea is to work in the space of eigenvectors Ψ_j of the cross-correlation matrix $\mathbf{M} \equiv \langle \delta \hat{\mathbf{x}} \cdot \delta \hat{\mathbf{x}}^t \rangle$, i.e. to diagonalize the cosmic covariance matrix of the data,

$$\mathbf{M} \cdot \Psi_j = \lambda_j \Psi_j, \quad (513)$$

where the matrix Ψ is unitary, $\Psi^{-1} = \Psi^t$. A new set of data, $\hat{\mathbf{y}}$, can be defined

$$\hat{\mathbf{y}} \equiv \Psi^t \cdot \hat{\mathbf{x}}, \quad (514)$$

which is statistically orthogonal

$$\langle \delta \hat{y}_i \delta \hat{y}_j \rangle = \lambda_i \delta_{ij} = {}^t \Psi_i \cdot \mathbf{M} \cdot \Psi_j \delta_{ij}. \quad (515)$$

The idea is to sort the new data from highest to lowest value of λ_i . Data compression will consist in ignoring data \hat{y}_i with λ_i lower than some threshold.

An interesting particular case of the KL transform is when the data can be decomposed in signal plus noise uncorrelated with each other [79]:

$$\hat{\mathbf{x}} = \hat{\mathbf{s}} + \hat{\mathbf{n}}. \quad (516)$$

The signal and the noise covariance matrices read

$$\mathbf{S} \equiv \langle \delta \hat{\mathbf{s}} \cdot \delta \hat{\mathbf{s}}^t \rangle, \quad \mathbf{N} \equiv \langle \delta \hat{\mathbf{n}} \cdot \delta \hat{\mathbf{n}}^t \rangle. \quad (517)$$

Then, instead of diagonalizing the cosmic covariance matrix of the data, one solves the generalized eigenvalue problem

$$\mathbf{S} \cdot \Psi_j = \lambda_j \mathbf{N} \cdot \Psi_j, \quad \Psi_j^t \cdot \mathbf{N} \cdot \Psi_j = 1. \quad (518)$$

The new data vector given by Eq. (514) is statistically orthogonal and verifies⁸⁶

$$\langle \delta \hat{y}_i \delta \hat{y}_j \rangle = (1 + \lambda_i) \delta_{ij}. \quad (519)$$

One can be easily convinced that this new transform is equivalent to a KL transform applied on the “prewhitened” data, $(\mathbf{N}^t)^{-1/2} \cdot \hat{\mathbf{x}}$, where

$$\mathbf{N} \equiv (\mathbf{N}^t)^{1/2} \cdot \mathbf{N}^{1/2}. \quad (520)$$

The advantage of this rewriting is that the quantity λ_i can be now considered as a signal to noise ratio $1 + \lambda_i = 1 + S/N$. Data compression on the prewhitened data makes now full physical sense, even if the noise is inhomogeneous or correlated.

The KL compression is generally used as a first step to reduce the size of the data set keeping as much information as possible, which can then be processed by the methods of ML estimation

⁸⁶ In the approximation that the distribution of $\hat{\mathbf{x}}$ is Gaussian, this also implies statistical independence.

or quadratic estimation which otherwise would not be computationally feasible. The final results should be checked against the number of KL modes kept in the analysis, to show that significant information has not been discarded. Note that in addition, since the methods generally used after KL compression assume Gaussianity, one must check as well that modes which probe the weakly non-linear regime are *not* included in the analysis to avoid having undesired biases in the final results.

6.12. Measurements in N -body simulations

Measurements of statistics in N -body simulations are of course subject to the cosmic error problem, but can be contaminated by other spurious effects related to limitations of the numerical approach used to solve the equations of motion. Transients, related to the way initial conditions are usually set up were already discussed in Section 5.7. Here, we first consider the cosmic error and the cosmic bias problems, which in practice are slightly different from the case of galaxy catalogs. Second, we briefly mention problems due to N -body relaxation and short-range softening of the gravitational force.

6.12.1. Cosmic error and cosmic bias in simulations

Here we restrict to the case of N -body simulations of self-gravitating collisionless dark matter. Most of simulations are done in a cubic box with periodic boundaries. The first important consequence is that the average number density of particles, \bar{n}_g , is perfectly determined.

The second consequence as mentioned earlier is that edge effects are inexistent. The only sources of errors are finite volume and shot noise. With the new generation of simulations, discreteness effects are in general quite small except at small scales or if a sparse synthetic catalog of “galaxies” is extracted from the dark matter distribution. Finite-volume effects in simulations have been extensively studied in [147,149,150]. For these effects to be insignificant in measured moments or correlation functions of the density distribution, the simulation box size L has to be large compared to the typical size of a large cluster, the correlation length R_0 . Typically it is required that $R_0 \lesssim L/20$. Even if this condition is fulfilled, the sampling scales (or separations) R must be small fractions of the box size in order to achieve fair measurements, typically $R \lesssim L/10$. Indeed, because of finite-volume effects, moments of the density distribution, cumulants and N -point correlation functions tend to be systematically underestimated, increasingly with scale. This is a consequence of cosmic bias and *effective bias* due to the skewness of the cosmic distribution function, as discussed in Section 6.10.

The estimation of cosmic bias was addressed quantitatively at large scales in [580] using PT where it was found that although moments can be affected by as much as 80% at smoothing scales one-tenth of the size of the box (for $n = -2$), the skewness S_3 was affected by at most 15% at the same scale. Finite-volume effects for velocity statistics are much more severe, as they are typically dominated by long-wavelength fluctuations, e.g. see [342].

The most obvious consequence of finite-volume effects is the fact that the high-density tail of the PDF develops a cutoff due to the finite number of particles. A method was proposed in [145,147,149] and exploited in other works [150,472] to correct the PDF for finite-volume effects, by smoothing and extending to infinity its large- δ tail. Another way to bypass finite-volume effects consists in doing several simulations and taking the average value (see, e.g. [356,251,28]) of the moments or cumulants, with the appropriate procedure for cumulants to avoid possible biases. This is however, by

itself not necessarily sufficient, because in each realization, large-scale fluctuations are still missing due to the periodic boundaries (e.g. [580]). In other words, doing a number of random realizations of given size L with periodic boundaries is not equivalent to extracting subsamples of size L from a very large volume. With many realizations one can reduce arbitrarily the effect of the skewness of the distribution, but not the influence of large-scale waves not present due to the finite volume of the simulations.

6.12.2. N -body relaxation and force softening

Due to the discrete nature of numerical simulations, there are some dynamical effects owing to interactions between small number of particles. To reduce these relaxation effects it is necessary to bound forces at small interparticle separation, thus a softening ϵ is introduced as discussed in Section 2.9. However, this softening does not guarantee the fluid limit. The latter is achieved locally only when the number of particles in a softening volume ϵ^3 is large. Typically, the softening parameter is of order the mean interparticle distance λ in low-resolution simulations, or of order $\lambda/20$ in high-resolution simulations (Section 2.9). At early stages of simulations, where the particles are almost homogeneously distributed, relaxation effects are thus expected to be significant. Later, when the system reached a sufficient degree of non-linearity, these effects occur only in underdense regions.⁸⁷ It is, therefore, important to wait long enough so that the simulation has reached a stage where typical non-linear structures contain many particles.

Statistically, this is equivalent to say that the correlation length should be much larger than the mean interparticle distance, $R_0 \gg \lambda$ [150]. This criterion is valid for most statistics but there are exceptions. For example, it was shown that the void PDF can be contaminated by the initial pattern of particles (such as a grid) even at late stages [149]. Indeed, underdense regions tend to expand and to keep the main features of this initial pattern. Another consequence is that the local Poisson approximation is not valid if this initial pattern presents significant correlations or anticorrelations (such as a grid or a “glass” [28,688]).

Finally, short-range softening of the forces itself can contaminate the measurement of statistics at small scales. With a careful choice of the timestep (see, e.g. [199]) the effects of the softening parameter are negligible for scales sufficiently large compared to ϵ , a practical criterion being that the considered scale R verifies $R = \alpha\epsilon$ with α of order a few [150].

7. Applications to observations

7.1. The problem of galaxy biasing

Application to galaxy surveys of the results that have been obtained for the clustering of dark matter is not trivial, because in principle there is no guarantee that galaxies are faithful tracers of the dark matter field. In other words, the galaxy distribution may be a *biased* realization of the underlying dark matter density field.

⁸⁷ In fact, in these regions, small but rare groups of particles experiencing strong collisions can be found even at late stages of the simulations.

A simplified view of biasing often encountered in the literature is that the two fields, galaxy and matter density fields, are simply proportional to each other:

$$\delta_g(\mathbf{x}) = b\delta(\mathbf{x}) . \quad (521)$$

It implies in particular that the power spectra obey $P_g(k) = b^2 P(k)$. As long as one considers two-point statistics this might be a reasonable prescription; however, when one wants to address non-Gaussian properties, this is no more sufficient: the connection between dark matter fluctuations and galaxies, or clusters of galaxies, should be given in more detail.

In principle, this relation should be obtained as a prediction of a given cosmological model. However, although significant progress has been done recently to study galaxy formation from “first principles” via hydrodynamic numerical simulations [122,369,72,498], they still suffer from limited dynamical range and rely on simplified descriptions of star formation and supernova feedback, which are poorly understood. This fundamental problem implies that when dealing with galaxies, one must usually include additional (non-cosmological) parameters to describe the relation between galaxies and dark matter. These parameters, known generally as *bias parameters*, must be determined from the data themselves. In fact, the situation turns out to be more complicated than that: since there is no generally accepted framework for galaxy biasing yet, one needs to test the parameterization itself against the data in addition to obtaining the best-fit parameter set.

The complexity of galaxy biasing is reflected in the literature, where many different approaches have emerged in the last decade or so. In addition to the hydrodynamic simulations, two other major lines of investigations can be identified in studies of galaxy biasing. The simplest one, involves a phenomenological mapping from the dark matter density field to galaxies, which is reviewed in the next section. Another approach, that has become popular in recent years, is to split the problem of galaxy biasing into two different steps [686]. First, the formation and clustering of dark matter halos, which can be modeled neglecting non-gravitational effects, this is the subject of Sections 7.1.2 and 7.1.3. This step is thought to be sufficient to describe the spatial distribution of galaxy clusters. The second step, discussed in Section 7.1.4, is the distribution of galaxies within halos, which is described by a number of simplifying assumptions about the complex non-gravitational physics. It is generally believed that such processes are likely to be very important in determining the properties of galaxies while having little effects on the formation and clustering of dark matter halos.

Note that observational constraints on biasing (from higher-order correlations) are discussed in the next section (see Sections 8.2.6 and 8.3.5).

7.1.1. Some general results

The first theoretical approach to galaxy biasing was put forward by Kaiser [360], who showed that if rich galaxy clusters were rare density peaks in a Gaussian random field, they will be more strongly clustered than the mass, as observed [503,15]. These calculations were further extended in [491,21]. In particular, it was found that rare peaks were correlated in such a way that

$$\langle \delta_{\text{peak}}^2 \rangle = b_{\text{peak}}^2 \langle \delta^2 \rangle , \quad (522)$$

where δ_{peak} is the local density contrast in the number density of peaks with a *bias parameter*

$$b_{\text{peak}}(v) = \frac{v}{\sigma} , \quad (523)$$

where σ is the variance at the peak scale and v the intrinsic density contrast of the selected peaks in units of σ . These results led to studies of biasing in CDM numerical simulations [173,685], which indeed showed that massive dark matter halos are more strongly clustered than the mass. However, numerical simulations also showed later that dark matter halos are not always well identified with peaks in the linear density field [368].

An alternative description of biasing, which does not rely on the initial density field, is the local Eulerian bias model. In this case, the assumption is that at scales R large enough compared to those where non-gravitational physics operates, the *smoothed* (over scale R) galaxy density at a given point is a function of the underlying *smoothed* density field at *the same point*,

$$\hat{\delta}_g(\mathbf{x}) = \mathcal{F}[\hat{\delta}(\mathbf{x})], \quad \hat{A}(\mathbf{x}) \equiv \int_{|\mathbf{x}'| < R} d^3\mathbf{x}' A(\mathbf{x} - \mathbf{x}') W(\mathbf{x}'), \quad (524)$$

where W denotes some smoothing filter. For large R , where $\hat{\delta} \ll 1$, it is possible to perturbatively expand the function \mathcal{F} in Taylor series and compute the galaxy correlation hierarchy [235]. Indeed, one can write

$$\hat{\delta}_g = \sum_{k=0}^{\infty} \frac{b_k}{k!} \hat{\delta}^k, \quad (525)$$

where the linear term b_1 corresponds to the standard linear bias factor. In this large-scale limit, such a local transformation preserves the hierarchical properties of the matter distribution, although the values of the hierarchical amplitudes may change arbitrarily. In particular [235],

$$\begin{aligned} \sigma_g^2 &= b_1^2 \sigma^2, \\ S_{g,3} &= b_1^{-1} (S_3 + 3c_2), \\ S_{g,4} &= b_1^{-2} (S_4 + 12c_2 S_3 + 4c_3 + 12c_2^2), \\ S_{g,5} &= b_1^{-3} [S_5 + 20c_2 S_4 + 15c_2 S_3^2 + (30c_3 + 120c_2^2) S_3 + 5c_4 + 60c_2 c_3 + 60c_2^3], \end{aligned} \quad (526)$$

where $c_k \equiv b_k/b_1$. As pointed out in [235], this framework encompasses the model of bias as a sharp threshold clipping [360,523,21,615], where $\delta_g = 1$ for $\delta > v\sigma$ and $\delta_g = 0$ otherwise. Although it does not have a series representation around $\delta = 0$, such a clipping applied to a Gaussian background produces a hierarchical result with $S_{g,p} = p^{p-2}$ in the limit $v \gg 1$, $\sigma \ll 1$. This is the same result as we obtain from Eq. (526) for an exponential biasing of a Gaussian matter distribution, $\delta_g = \exp(\alpha\delta/\sigma)$, which is equivalent to the sharp threshold when the threshold is large and fluctuations are weak [21,615]. The exponential bias function has an expansion $\mathcal{F} = \sum_k (\alpha\delta/\sigma)^k/k!$ and thus $b_k = b_1^k$, independent of α and σ . With $S_p = 0$, the terms induced in Eq. (526) by b_k alone also give $S_{g,p} = p^{p-2}$.

As a result of Eq. (526), it is clear that for higher-order correlations, $p > 2$, a linear bias assumption cannot be a consistent approximation even at very large scales, since non-linear biasing can generate higher-order correlations. To draw any conclusions from the galaxy distribution about matter correlations of order p , properties of biasing must be included to order $p - 1$.

Let us make at this stage a general remark. From Eq. (526) it follows that in the simplest case, when the bias is linear, a value $b_1 > 1$ reduces the S_p parameters and it may suggest that this changes how the distribution deviates from a Gaussian (e.g. the galaxy field would be “more Gaussian” than the underlying density field, given that S_3 is smaller). However, this is obviously an incorrect conclusion, a linear scaling of the density field cannot alter the degree of non-Gaussianity. The reason is that the actual measure of non-Gaussianity is encoded not by the hierarchical amplitudes S_p but rather by the dimensionless skewness $B_3 = S_3\sigma$, kurtosis $B_4 = S_4\sigma^2$, and so on, which remain invariant under linear biasing. These dimensionless quantities are indeed what characterize the PDF, as it clearly appears in an Edgeworth expansion, Eq. (144).

Since Fourier transforms are effectively a smoothing operation, similar results to those above hold for Fourier-space statistics at low wavenumbers. In this regime, the galaxy density power spectrum $P_g(k)$ is given by

$$P_g(k) = b_1^2 P(k) \quad (527)$$

and the galaxy (reduced) bispectrum obeys [recall Eq. (154)]

$$Q_g(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = \frac{1}{b_1} Q(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) + \frac{b_2}{b_1^2}. \quad (528)$$

As discussed in Section 4.1.3, Q given by Eq. (155), is very insensitive to cosmological parameters and depends mostly on triangle configuration and the power spectrum spectral index. Since the latter is not affected by bias in the large-scale limit, Eq. (527), it can be measured from the galaxy power spectrum and used to predict $Q(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)$ as a function of triangle configuration. As first proposed in [224,236], a measurement of Q_g as a function of triangle shape can be used to determine $1/b_1$ and b_2/b_1^2 . So far, this technique has only been applied to IRAS galaxies [567,211], as will be reviewed in the next section (see Section 8.3.3).⁸⁸

The results above suggest that local biasing does not change the *shape* of the correlation function or power spectrum in the large-scale limit, just scaling them by a constant factor b_1^2 independent of scale. This derivation [235] assumes that the smoothing scale is large enough so that $\hat{\delta} \ll 1$, but in fact, it can be shown that this continues to hold in more general situations. For example, an arbitrary local transformation of a *Gaussian* field, leads to a bias that cannot be an increasing function of scale and that becomes constant in the large-scale limit, irrespective of the amplitude of the rms fluctuations [140].⁸⁹ However, it is easy to show that if the underlying density field is hierarchical (in the sense that the C_{pq} parameters in Eq. (348) are independent of scale), a local mapping such as that in Eq. (524) does lead to a bias independent of scale in the large-scale limit even if $\hat{\delta} \gg 1$ [41,553].

Recent studies of galaxy biasing [553,180,72,440] have focused on the fact that Eq. (524) assumes not only that the bias is local but also deterministic; that is, the galaxy distribution is completely determined by the underlying mass distribution. In practice, however, it is likely that galaxy formation depends on other variables besides the density field, and that consequently the relation between $\hat{\delta}_g(\mathbf{x})$ and $\hat{\delta}(\mathbf{x})$ is not deterministic but rather stochastic,

$$\hat{\delta}_g(\mathbf{x}) = \mathcal{F}[\hat{\delta}(\mathbf{x})] + \varepsilon_\delta(\mathbf{x}), \quad (529)$$

⁸⁸ Similar relations to Eqs. (526) and (528) can be obtained for cumulant correlators, see [626].

⁸⁹ But this is an unrealistic situation since Gaussianity breaks down when the rms fluctuations are larger than unity.

where the random field $\varepsilon_\delta(\mathbf{x})$ denotes the scatter in the biasing relation *at a given* δ due to the fact that $\hat{\delta}(\mathbf{x})$ does not completely determine $\hat{\delta}_g(\mathbf{x})$. Clearly for an arbitrary scatter, the effects of $\varepsilon_\delta(\mathbf{x})$ on clustering statistics can be arbitrarily strong. However, under the assumption that the scatter is local, in the sense that the correlation functions of $\varepsilon_\delta(\mathbf{x})$ vanish sufficiently fast at large separations (i.e. faster than the correlations in the density field), the deterministic bias results hold for the two-point correlation function in the large-scale limit [553]. For the power spectrum, on the other hand, in addition to a constant large-scale bias, stochasticity leads to a constant offset (given by the rms scatter) similar to Poisson fluctuations due to shot noise [553,180].

Another interesting aspect of stochasticity was studied in [440], in connection with non-local biasing. A simple result can be obtained as follows. Suppose that biasing is non-local but linear, then we can write

$$\delta_g(\mathbf{x}) = \int \delta(\mathbf{x}') K(\mathbf{x} - \mathbf{x}') d^3 \mathbf{x}', \quad (530)$$

where the kernel K specifies how the galaxy field at position \mathbf{x} depends on the density field at arbitrary locations \mathbf{x}' . This convolution of the density field leads to stochasticity in real space, i.e. the cross-correlation coefficient r ,

$$r(s) \equiv \frac{\langle \delta(\mathbf{x}) \delta_g(\mathbf{x}') \rangle}{\sqrt{\xi_g(s) \xi(s)}}, \quad (531)$$

where $s \equiv |\mathbf{x} - \mathbf{x}'|$, is not necessarily unity. However, due to the convolution theorem, the cross-correlation coefficient in *Fourier space* will be exactly unity, thus

$$\langle \delta_g(\mathbf{k}) \delta(\mathbf{k}') \rangle = \delta_D(\mathbf{k} + \mathbf{k}') b(k) P(k) \quad (532)$$

and

$$\langle \delta_g(\mathbf{k}) \delta_g(\mathbf{k}') \rangle = \delta_D(\mathbf{k} + \mathbf{k}') b^2(k) P(k), \quad (533)$$

where the bias $b(k)$ is the Fourier transform of the kernel K . The study in [440] showed, on the other hand, that the real-space stochasticity (in the sense that $r < 1$) at large scales was weak for some class of models. At small scales, however, significant deviations from $r < 1$ cannot be excluded, for example due to non-linear couplings in Eq. (530). However, without specifying more about the details of the biasing scheme, it is very difficult to go much beyond these results.

Most of the general results discussed so far have been observed in hydrodynamical simulations of galaxy formation. For example, in [72] it has been obtained that at large scale ($R \gtrsim 15 \text{ Mpc}/h$) the bias parameter tends to be constant and the cross-correlation coefficient r reaches unity for oldest galaxies. The authors stress that the bias shows a substantial scale dependence at smaller scales, which they attribute to the dependence of galaxy formation on the temperature of the gas (which governs its ability to cool). In addition, they observe a substantial amount of stochasticity for young galaxies ($r \approx 0.5$), even at large scales. However, these results are in disagreement with observations of the LCRS survey, where it was found that after correcting for errors in the selection function the cross-correlation between early and late-type galaxies is $r \approx 0.95$ [71].

Another assumption that enters into the local Eulerian biasing model discussed above is that the galaxy field depends on the underlying density field *at the same time*. In practice, it is expected that to some extent the merging and tidal effects *histories* affect the final light distribution. This can lead to non-trivial time evolution of biasing. For instance, as shown in [241], if galaxy formation was

very active in the past but after some time it becomes subdominant, then *in the absence of merging* the galaxy density contrast is expected to follow the continuity equation,

$$a \frac{\partial \delta_g}{\partial a} + \mathbf{u} \cdot \nabla \delta_g + (1 + \delta_g) \nabla \cdot \mathbf{u} = 0, \quad (534)$$

where \mathbf{u} is the peculiar velocity field of the dark matter field: galaxies are simple test particles that follow the large-scale flows. Formally this equation can be rewritten as

$$\frac{d \log(1 + \delta_g)}{a d\tau} = \frac{d \log(1 + \delta)}{d\tau}, \quad (535)$$

where $d/d\tau$ is the convective derivative. As a consequence, the galaxy density field is expected to resemble more and more the density field in terms of correlation properties: both the bias parameters, b_k , and the cross-correlation coefficient, r , are expected to approach unity, galaxies “de-bias” when they just follow the gravitational field [482,241,649]. The higher-order moments characterized by S_p are also expected to get closer to those for the dark matter field. These calculations have been illustrated in [241,641].

One obvious limitation of these “galaxy conserving” schemes is the assumption that there is no merging, which is expected to play a central role in hierarchical structure formation. In addition, ongoing galaxy formation leads to galaxies formed at different redshifts with different “bias at birth”. Indeed, models based on the continuity equation predict a *slower* time evolution of bias than observed in simulations [73,599], i.e. galaxies become unbiased *faster* than when these effects are neglected.

An interesting consequence of Eq. (535) has been unveiled in [119] where they remark that the solution is

$$1 + \delta_g(\mathbf{x}, z) = [1 + \delta_g^L(\mathbf{q})][1 + \delta(\mathbf{x}, z)], \quad (536)$$

where the galaxy field at the *Lagrangian* position \mathbf{q} is obtained from the linear density field at $\mathbf{q} = \mathbf{x} - \Psi(\mathbf{q}, z)$ by $\delta_g^L(\mathbf{q}) = \sum b_k^L/k! \delta^L(\mathbf{q})$. That is, in this model, the bias is assumed to be local in *Lagrangian* space rather than Eulerian space. In this particular case, unlike in peaks biasing mentioned above, once the galaxy field is identified in the initial conditions, its subsequent evolution is incorporated by Lagrangian perturbation theory to account for displacement effects due to the gravitational dynamics. In this case, the tree-level bispectrum amplitude becomes [120]

$$Q_g = \frac{1}{b_1} Q + \frac{b_2^L}{b_1^2} + \frac{4b_1^L}{7b_1^2} \times \frac{\Delta Q_{12} P_g(k_1) P_g(k_2) + \text{cyc.}}{P_g(k_1) P_g(k_2) + \text{cyc.}}, \quad (537)$$

where $\Delta Q_{12} \equiv 1 - (\mathbf{k}_1 \cdot \mathbf{k}_2)^2 / (k_1 k_2)^2$ and $b_1 \equiv 1 + b_0^L + b_1^L$. Note that the last term in this expression gives a different prediction than Eq. (528) for the dependence of the galaxy bispectrum as a function of triangle configuration that can be tested against observations; application to the PSCz survey bispectrum [211] suggests that the model in Eq. (528) fits better the observations than Eq. (537).

Finally, we should also mention that a number of phenomenological (more complicated) mappings from dark matter to galaxies have been studied in detail in the literature [431,133,474,38]. The results are consistent with expectations based on the simpler models discussed in this section.

7.1.2. Halo clustering in the tree hierarchical model

As mentioned previously, the validity of prescription (524) is subject to the assumption that the mass density contrast is small. For biasing at small scales this cannot be a valid assumption. Insights into the functional relation between the halo field and the matter field then demand for a precise modeling of the matter fields. The *tree hierarchical model*, Eq. (222), has been shown to provide a solid ground to undertake such an investigation [41,57]. In these papers the *connected part* of joint density distribution have been computed for an arbitrary number of cells, $p_c(\delta_1, \dots, \delta_p)$ and showed to be of the form

$$p_c(\delta_1(\mathbf{x}_1), \dots, \delta_p(\mathbf{x}_p)) = \sum_{a=1}^{t_p} Q_{p,a}(\delta_1, \dots, \delta_N) \sum_{\text{labelings edges}} \prod_{i=1}^{p-1} \xi_2(\mathbf{x}_i, \mathbf{x}_j) \quad (538)$$

with

$$Q_{p,a}(\delta_1, \dots, \delta_p) = \Pi_i p(\delta_i) v_q(\delta_i), \quad (539)$$

where $v_q(\delta)$ is a function of the local density contrast that depends on the number q of lines it is connected to in the graph. This form implies, for instance, that

$$p(\delta_1, \delta_2) = p(\delta_1)p(\delta_2)[1 + \xi_2(\mathbf{x}_1, \mathbf{x}_2)v_1(\delta_1)v_1(\delta_2)]. \quad (540)$$

At small scales, when the variance is large, the density contrast of dark matter halos is much larger than unity, and should be reliably given by a simple threshold condition, $\delta_i \gtrsim \delta_{\text{thres}}$. Therefore, the function v_1 describes the halo bias, and higher-order connected (two-point) joint moments follow directly from this *bias function* and the two-point correlation function of the mass. In this framework a number of important properties and results have been derived:

- (i) the correlation functions of the halo population follow a tree structure similar to the one of the matter field in the large separation limit (e.g. when the distances between the halos are much larger than their size);
- (ii) the values of the vertices depend only on the internal properties of the halos, namely on the reduced variable,

$$x = \frac{\rho}{\bar{\rho}\sigma^2}; \quad (541)$$

- (iii) all vertices are growing functions of x and have a specific large x asymptotic behavior,

$$v_1(x) \equiv b(x) \sim x, \quad (542)$$

$$v_p(x) \sim b^p(x). \quad (543)$$

The large x limit that has been found for the high-threshold clipping limit is once again recovered since we expect in such a model that $S_{h,p} \rightarrow p^{p-2}$ when $x \rightarrow \infty$. Property (iii), together with (ii), also holds for halos in the framework of the Press–Schechter approach, as we shall see in the next section [see discussion below Eq. (556)].

In addition, it is possible to derive the functions $v_p(x)$ in terms of the vertex generating function $\zeta(\tau)$. These results read

$$v_p(x) = \frac{\int_{-\infty}^{i\infty} dy \varphi^{(p)}(y) \exp(xy)}{\int_{-\infty}^{i\infty} dy \varphi(y) \exp(xy)}, \quad (544)$$

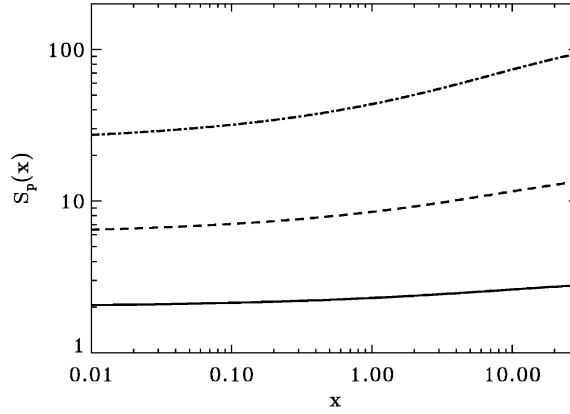


Fig. 43. Example of a computation of the S_3 , S_4 and S_5 parameters in the tree hierarchical model for dark matter halos selected with a varying threshold in x , defined by Eq. (541). Calculations have been made with the vertex generating function, $\zeta(\tau) = (1 - \tau/\kappa)^{-\kappa}$ with $\kappa = 1.3$. For large values of x one explicitly sees the $S_p \rightarrow p^{p-2}$ behavior expected in the high-threshold limit.

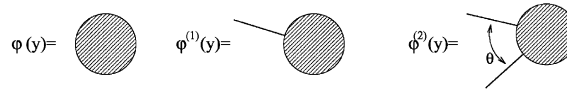


Fig. 44. The functions $\varphi(y)$, $\varphi^{(1)}(y)$ and $\varphi^{(2)}(y)$ are the generating functions of trees with respectively 0, 1 and 2 external lines. For orders above 2 a possible angular dependence with the outgoing lines cannot be excluded.

where the function $\varphi^{(p)}(y)$ can be expressed in terms of ζ and its derivatives (see [57] for details). In case of the minimal tree model where all vertices are pure numbers, we have

$$\varphi(y) = y\zeta(\tau) + \tau^2/2, \quad \tau/\zeta'(\tau) = -y, \tag{545}$$

$$\varphi^{(1)}(y) = \tau(y), \tag{546}$$

$$\varphi^{(2)}(y) = -\frac{y\zeta''(\tau)}{1 + y\zeta'''(\tau)}, \tag{547}$$

$$\varphi^{(3)}(y) = -\frac{y\zeta'''(\tau)}{[1 + y\zeta'''(\tau)]^3}, \tag{548}$$

⋮

These results provide potentially a complete model for dark matter halo biasing. The explicit dependence of the skewness and kurtosis parameters has been computed in these hierarchical models in [57], see Fig. 43.

Although initially undertaken in the strongly non-linear regime, these results a priori extend to weakly non-linear scales, that is, to scales where halo separations are in the weakly non-linear regime. Indeed only the tree structure, in a quite general sense (see [48,57] for details), is required to get these results. In this case the vertex $v_2(x)$ might bear a non-trivial angular dependence originating

from the expression of $\varphi^{(2)}(y)$, see Fig. 44. There is, therefore a priori no reason to recover the result in Eq. (528) for the halo bispectrum. The connection, if any, with simple relations such as Eq. (524) is thus still to be understood. Stochasticity emerging due to non-linear effects is in particular likely to limit the validity of Eq. (524).

7.1.3. Halo clustering in the extended Press–Schechter approach

The results obtained in the previous section correspond to the correlations properties of dense halos detected in a snapshot of the non-linear density field. This approach does not give any insights into the merging history of the halos that is likely to be important for the galaxy properties. And because dark matter halos are highly non-linear objects, their formation and evolution has traditionally been studied using numerical simulations.

However, a number of analytical models [460,459,119,589], based on the so-called Press–Schechter (PS) formalism [531] and extensions [78,95,388,370], revealed a good description of the numerical simulation results.

The PS formalism aims at giving the comoving number density of halos as a function of their mass m ,

$$\frac{m^2 n(m)}{\bar{\rho}} = \sqrt{\frac{2y^2}{\pi}} \exp\left(-\frac{y^2}{2}\right) \frac{d \ln y}{d \ln m}, \quad (549)$$

where $\bar{\rho}$ denotes the average density of the universe, and $y \equiv \delta_c/\sigma(m)$, with $\delta_c \approx 1.68$ the collapse threshold given by the spherical collapse model and $\sigma^2(m)$ is the variance of the linearly extrapolated density field smoothed at scale $R = (3m/4\pi\bar{\rho})^{1/3}$. The average number of halos in a spherical region of comoving radius R_0 and over-density δ_0 is

$$\mathcal{N}(m|\delta_0) dm = \frac{m_0}{m} f(\sigma, \delta_c|\sigma_0, \delta_0) \frac{d\sigma^2}{dm} dm, \quad (550)$$

where

$$f(\sigma, \delta_c|\sigma_0, \delta_0) = \frac{1}{\sqrt{2\pi}} \frac{\delta_c - \delta_0}{(\sigma^2 - \sigma_0^2)^{3/2}} \exp\left[-\frac{(\delta_c - \delta_0)^2}{2(\sigma^2 - \sigma_0^2)}\right] \quad (551)$$

is the fraction of the mass in a region of initial radius R_0 and linear over-density δ_0 that is at present in halos of mass m [78,95]. The Lagrangian halo density contrast is then [460]

$$\delta_h^L(m|\delta_0) = \frac{\mathcal{N}(m|\delta_0)}{n(m)V_0} - 1, \quad (552)$$

where $V_0 = 4\pi R_0^3/3$. When $R_0 \gg R$ so that $\sigma_0 \ll \sigma$ and $|\delta_0| \ll \delta_c$, this gives

$$\delta_h^L(m|\delta_0) = \frac{y^2 - 1}{\delta_c} \delta_0. \quad (553)$$

On the other hand, the Eulerian halo density contrast is [460]

$$\delta_h(m|\delta_0) = \frac{\mathcal{N}(m|\delta_0)}{n(m)V} - 1, \quad (554)$$

where the volume $V = 4\pi R^3/3$ is related to the initial volume by $R_0 = R(1 + \delta)^{1/3}$ with $\delta(\delta_0) = \sum_{m=1}^{\infty} v_m \delta_0^m$ given by the spherical collapse model. When considered as a function of δ , Eq. (554)

gives a bias relation similar to Eq. (525) with bias parameters [459]

$$\begin{aligned} b_1(m) &= 1 + \epsilon_1, & b_2(m) &= 2(1 - v_2)\epsilon_1 + \epsilon_2, \\ b_3(m) &= 6(v_3 - v_2)\epsilon_1 + 3(1 - 2v_2)\epsilon_2 + \epsilon_3, \end{aligned} \quad (555)$$

with

$$\epsilon_1 = \frac{y^2 - 1}{\delta_c}, \quad \epsilon_2 = \frac{y^2(y^2 - 3)}{\delta_c^2}, \quad \epsilon_3 = \frac{y^2(y^4 - 6y^2 + 3)}{\delta_c^3}. \quad (556)$$

This framework has been extended to give halo biasing beyond the spherical collapse approximation, in particular Catelan et al. [119] discuss the use of the ZA, the frozen-flow approximation and second-order Eulerian PT. In addition, Sheth et al. [593] study the effects of ellipsoidal collapse on both the mass function and the biasing of dark matter halos. They show that tidal effects change the threshold condition for collapse to become a function of mass, $\delta_c(m)$, and that the resulting halo bias and mass function are in better agreement with numerical simulations than the PS ones. In particular, less massive halos are more strongly clustered than in PS calculations as summarized by fitting formulae derived from N -body simulations [350,527], and low (high) mass halos are less (more) abundant than predicted in PS [590,343].

The higher-order moments for dark matter halos can be calculated from the expansion in Eqs. (555) and (526), as first done in [459]. For instance, in the rare peak limit $b_1 \sim y^2/\delta_c \gg 1$ and $b_2 \sim b_1^2$ so that the three-point function obeys the hierarchical model with $Q_3 = 1$ (or equivalently $S_3 = 3$). This actually extends to any order to give $Q_N = 1$, i.e. $S_p = p^{p-2}$ in this limit [459].

The fact that dark matter halos are spatially exclusive induces non-trivial features on their correlation functions at small scales, which cannot be modeled simply as a biasing factor acting on the mass correlation functions. In particular, the variance becomes significantly less than the Poisson value at small scales [460]. A detailed discussion of exclusion effects can be found in [589].

7.1.4. Galaxy clustering

Since galaxy formation cannot yet be described from first principles, a number of prescriptions based on reasonable recipes for approximating the complicated physics have been proposed for incorporating galaxy formation into numerical simulations of dark matter gravitational clustering [371,598,134]. These “semi-analytic galaxy formation” schemes can provide detailed predictions for galaxy properties in hierarchical structure formation models, which can then be compared with observations.

The basic assumption in the semi-analytic approach is that the distribution of galaxies within halos can be described by a number of simplifying assumptions regarding gas cooling and feedback effects from supernova. For the purposes of large-scale structure predictions, the main outcome of this procedure is the number of galaxies that populate a halo of a given mass, $N_{\text{gal}}(m)$. Typically, at large mass $\langle N_{\text{gal}}(m) \rangle \sim m^\alpha$ with $\alpha < 1$, and below some cutoff mass $N_{\text{gal}}(m) = 0$. The physical basis for this behavior is that for large masses the gas cooling time becomes larger than the Hubble time, so galaxy formation is suppressed in large-mass halos (therefore $\langle N_{\text{gal}}(m) \rangle$ increases less rapidly than the mass). On the other hand, in small-mass halos effects such as supernova winds can blow away the gas from halos, also suppressing galaxy formation.

A useful analytical model has been recently developed, generally known as “the halo model”, which can be easily modified to provide a description of galaxy clustering using knowledge of the

$N_{\text{gal}}(m)$ relation and the clustering of dark matter halos described in Section 7.1.3. The starting point is a description of the dark matter distribution in terms of halos with masses, profiles and correlations consistent with those obtained in numerical simulations. This is a particular realization of the formalism first worked out in [552] for general distribution of seed masses, although precursors which did not include halo–halo correlations were studied long before [477,502,446].

Let $u_m(\mathbf{r})$ be the profile of dark matter halos of mass m (for example, as given in [475,463]), normalized so that $\int d^3x' u_m(\mathbf{x}-\mathbf{x}')=1$, and $n(m)$ be the mass function, with $\int n(m)m dm=\bar{\rho}$ and $\bar{\rho}$ the mean background density. The power spectrum in this model is written as [587,495,579,419,158,570]

$$\begin{aligned} \bar{\rho}^2 P(k) &= (2\pi)^3 \int n(m)m^2 dm |u_m(\mathbf{k})|^2 + (2\pi)^6 \int u_{m_1}(k)n(m_1)m_1 dm_1 \\ &\quad \times \int u_{m_2}(k)n(m_2)m_2 dm_2 P(k; m_1, m_2), \end{aligned} \quad (557)$$

where $P(k; m_1, m_2)$ represents the power spectrum of halos of mass m_1 and m_2 . The first term denotes the power spectrum coming from pairs inside the same halo (“1-halo” term), whereas the second contribution comes from pairs in different halos (“2-halo” term). Similarly, the bispectrum is given by

$$\begin{aligned} \bar{\rho}^3 B_{123} &= (2\pi)^3 \int n(m)m^3 dm \prod_{i=1}^3 u_m(\mathbf{k}_i) + (2\pi)^6 \int u_{m_1}(k_1)n(m_1)m_1 dm_1 \\ &\quad \times \int u_{m_2}(k_2)u_{m_2}(k_3)n(m_2)m_2^2 dm_2 P(k_1; m_1, m_2) + \text{cyc.} \\ &\quad + (2\pi)^9 \left(\prod_{i=1}^3 \int u_{m_i}(k_i)n(m_i)m_i dm_i \right) B_{123}(m_1, m_2, m_3), \end{aligned} \quad (558)$$

where $B_{123}(m_1, m_2, m_3)$ denotes the bispectrum of halos of mass m_1, m_2, m_3 . Again, contributions in Eq. (558) can be classified according to the spatial location of triplets, from “1-halo” (first term) to “3-halo” (last term). The halo–halo correlations, encoded in $P(k; m_1, m_2)$, $B_{123}(m_1, m_2, m_3)$ and so on, are described by non-linear PT plus the halo-biasing prescription discussed in Section 7.1.3, Eq. (555), plus Eqs. (526)–(528) with mass correlation functions obtained from perturbation theory.

To describe *galaxy* clustering, one needs to specify the distribution (mean and the higher-order moments) of the number of galaxies which can inhabit a halo of mass m . This is an output of the semi-analytic galaxy formation schemes, e.g. [371,36], or some parameterization can be implemented (see e.g. [570,39,37]) which is used to fit the clustering statistics. Assuming that galaxies follow the dark matter profile, the galaxy power spectrum reads [579,570]

$$\begin{aligned} \bar{n}_g^2 P_g(k) &= (2\pi)^3 \int n(m) \langle N_{\text{gal}}^2(m) \rangle dm |u_m(\mathbf{k})|^2 \\ &\quad + (2\pi)^6 \left[\int u_m(k)n(m) dm b_1(m) \langle N_{\text{gal}}(m) \rangle \right]^2 P_L(k) \end{aligned} \quad (559)$$

and similarly for the bispectrum, where the mean number density of galaxies is

$$\bar{n}_g = \int n(m) \langle N_{\text{gal}}(m) \rangle dm. \quad (560)$$

Thus, knowledge of the number of galaxies per halo moments $\langle N_{\text{gal}}^n(m) \rangle$ as a function of halo mass gives a complete description of the galaxy clustering statistics within this framework. Note that in the large-scale limit, the galaxy bias parameters reduce to $[u_m(k) \rightarrow 1]$

$$b_i \approx \frac{1}{\bar{n}_g} \int n(m) dm b_i(m) \langle N_{\text{gal}}(m) \rangle. \quad (561)$$

Therefore, in this prescription the large-scale bias parameters are not independent, the whole hierarchy of b_i 's is a result of Eqs. (555) for $b_i(m)$ and the $\langle N_{\text{gal}}(m) \rangle$ relation, which can be described by only a few parameters. In addition, the higher-order moments $\langle N_{\text{gal}}^n(m) \rangle$ with $n > 1$, determine the small-scale behavior of galaxy correlations; however, relations can be obtained between these moments and the mean which, if robust to details,⁹⁰ means that the parametrization of the mean relation is the main ingredient of galaxy biasing. In this sense, this framework promises to be a very powerful way of constraining galaxy biasing.

The weighing introduced by $\langle N_{\text{gal}}^n(m) \rangle$ on clustering statistics has many desirable properties. In particular, the suppression of galaxy formation in high-mass halos leads to a galaxy power spectrum that displays power-law-like behavior⁹¹ [36,579,495,570] and higher-order correlations show smaller amplitudes at small scales than their dark matter counterparts [570] (see Fig. 45), as observed in galaxy catalogs. A very important additional consideration is that this high-mass suppression also leads to velocity dispersion of galaxies in agreement with galaxy surveys such as LCRS [349].

7.2. Projection effects

This section is devoted to the particular case of angular surveys. These surveys constitute a large part of the available data and allow to probe the statistical properties of the cosmic density field at large scales, as we shall discuss in the next section, and furthermore they do not suffer from redshift-space distortions. Although they do not really probe new aspects of gravitational dynamics, the filtering scheme deserves a specific treatment. It is also worth noting here, as we shall briefly discuss in the next section, that this filtering directly applies to weak lensing observations that are now emerging, see e.g. [453] for a review.

In the following we first review the general aspects of projection effects, and quickly turn to the widely used small-angle approximation, where most applications have been done. We then show how the 3D hierarchical model projects into a 2D hierarchy, where the 3D and 2D hierarchical coefficients are simply related. In Sections 7.2.4 and 7.2.5 we go beyond the hierarchical assumption to present predictions for the projected density in PT. Finally, in Section 7.2.6, we discuss the reconstruction of the one-point PDF of the projected density.

⁹⁰ The simplest of such relations assumes Poisson statistics, where $\langle N_{\text{gal}}(N_{\text{gal}} - 1) \cdots (N_{\text{gal}} - j) \rangle = \langle N_{\text{gal}} \rangle^{j+1}$, but it is known to fail for low-mass halos which have sub-Poisson dispersions [371,36]. A simple fix assumes a binomial distribution [570], with two free parameters that reproduce the mean and second moment, and automatically predict the $n > 2$ moments. However, it is not known yet how well this model does predict the $n > 2$ moments. Other prescriptions are given in [36,39,37]; in particular, Berlind and Weinberg [39] study in detail the sensitivity of galaxy clustering to the underlying distribution.

⁹¹ In addition, note that a power-law behavior has also been obtained in numerical simulations by selecting “galaxies” as halos of specific circular velocities [141].

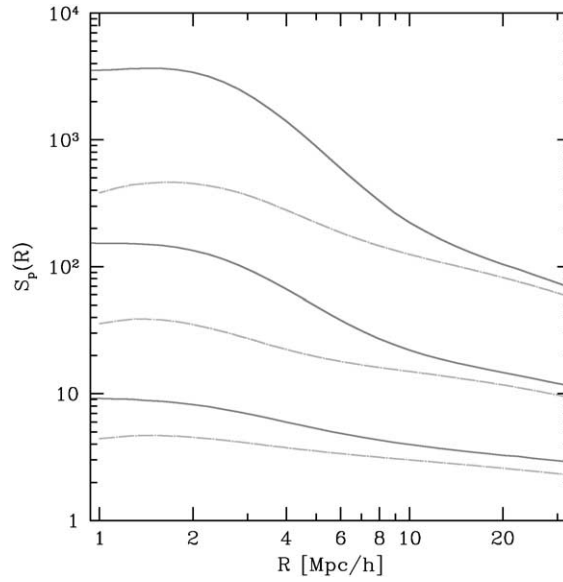


Fig. 45. The S_p parameters for $p=3,4,5$ (from bottom to top) for dark matter (solid) and galaxies (dot-dashed) as a function of smoothing scale R . These predictions correspond to those of the halo model, for galaxies they assume that $\langle N_{\text{gal}} \rangle = (m/m_0)^{0.8}$ for $m > m_0 = 8 \times 10^{11} M_{\odot} h^{-1}$, $\langle N_{\text{gal}} \rangle = (m/m_0)$ for $m_c < m < m_0$ and $\langle N_{\text{gal}} \rangle = 0$ for $m < m_c = 4 \times 10^9 M_{\odot} h^{-1}$.

7.2.1. The projected density contrast

Let us describe the comoving position \mathbf{x} in terms of the radial distance χ and angular distance \mathcal{D} so that $\mathbf{x} = (\chi, \mathcal{D}\theta)$.⁹² The radial distance is defined by⁹³

$$d\chi = \frac{c dz/H_0}{\sqrt{\Omega_A + (1 - \Omega_m - \Omega_A)(1+z)^2 + \Omega_m(1+z)^3}} \quad (562)$$

with H_0 Hubble's constant⁹⁴ and c the speed of light, while the angular distance is defined by

$$\mathcal{D}(\chi) = \frac{c/H_0}{\sqrt{1 - \Omega_m - \Omega_A}} \sinh\left(\sqrt{1 - \Omega_m - \Omega_A} \frac{H_0 \chi}{c}\right). \quad (563)$$

In general, for angular surveys, the measured density contrast of galaxy counts at angular direction θ is related to the 3D density contrast through

$$\delta_{2D}(\theta) = \int d\chi \chi^2 \psi(\chi) \delta_{3D}(\chi, \mathcal{D}\theta), \quad (564)$$

where $\psi(\chi)$ is the selection function (normalized such that $\int d\chi \chi^2 \psi(\chi) = 1$); it is the normalized probability that a point (galaxy) at a distance χ is included in the catalog.

⁹² See cosmology textbooks, e.g. [511], or the pedagogical summary in [316] for a detailed presentation of these aspects.

⁹³ Note that the Ω parameters refer here to those evaluated at redshift $z=0$.

⁹⁴ Throughout this work we use $H_0 = 100h$ km/s/Mpc.

In practice, the depth of the projection is finite due to the rapid decrease of the selection function $\psi(\chi)$ with χ at finite distance. The selection function $\psi(\chi)$ for a sample limited by apparent magnitudes between m_1 and m_2 is typically given by

$$\psi(\mathcal{D}) = \psi^* \int_{q_1(\mathcal{D})}^{q_2(\mathcal{D})} dq \phi^* q^\alpha e^{-q}, \quad q_i(\mathcal{D}) = 10^{-(5/2)(M_i(\mathcal{D})-M^*)}, \quad i = 1, 2 \quad (565)$$

with

$$M_i(\mathcal{D}) = m_i - 5 \log_{10} \mathcal{D}(1+z) - 25, \quad (566)$$

where ψ^* is a normalization constant and $\phi(q) = \phi^* q^\alpha e^{-q}$ is the luminosity function, i.e. the number density of galaxies of a given luminosity. M^* and α might be expressed as a function of redshift z to account for k -corrections and luminosity evolution. When redshift information is available, one can also rewrite the selection function in terms of the sample redshift number counts $N(z)$ alone.

7.2.2. The small-angle approximation

The cumulants of the projected density can obviously be related to those of the 3D density fields. Formally, they correspond to the ones of the 3D field filtered by a conical-shaped window. From Eq. (564) we obtain

$$w_N(\theta_1, \dots, \theta_N) = \int \prod_{i=1}^N d\chi_i \chi_i^2 \psi(\chi_i) \langle \delta(\chi_1, \mathcal{D}_1 \theta_1) \cdots \delta(\chi_N, \mathcal{D}_N \theta_N) \rangle_c. \quad (567)$$

The computation of such quantities can be easily carried out in the small-angle approximation. Such approximation is valid when the transverse distances $\mathcal{D}|\vec{\theta}_i - \vec{\theta}_j|$ are much smaller than the radial distances χ_i . In this case, integral (567) is dominated by configurations where $\chi_i - \chi_j \sim \mathcal{D}_i |\vec{\theta}_i - \vec{\theta}_j| \sim \mathcal{D}_j |\vec{\theta}_i - \vec{\theta}_j|$. This allows to make the change of variables $\chi_i \rightarrow r_i$ with $\chi_i = \chi_1 + r_i \mathcal{D}_1 (\theta_i - \theta_1)$. Then, since the correlation length (beyond which the multi-point correlation functions are negligible) is much smaller than the Hubble scale $c/H(z)$ (where $H(z)$ is the Hubble constant at redshift z), the integral over r_i converges over a small distance of the order of $\mathcal{D}_1 |\vec{\theta}_i - \vec{\theta}_1|$ and expression (567) can be simplified to read

$$w_N(\theta_1, \dots, \theta_N) = \int d\chi_1 \chi_1^{2N} \mathcal{D}^{N-1} \psi(\chi_1)^N \times \int_{-\infty}^{\infty} \prod_{i=2}^N (\theta_i - \theta_1) dr_i \xi_N[(\chi_1, \mathcal{D}_1 \theta_1), \dots, (\chi_N, \mathcal{D}_1 \theta_N)]. \quad (568)$$

This equation constitutes the small-angle approximation for the correlation functions. If these behave as power laws, Eq. (568) can be further simplified. For instance, the two-point function is then given by the *Limber equation* [402]

$$w_2(\theta) = \theta^{1-\gamma} r_0^\gamma \int d\chi \chi^4 \mathcal{D}^{1-\gamma} \psi^2(\chi) \int_{-\infty}^{\infty} dr (1+r^2)^{-\gamma/2}, \quad (569)$$

if the 3D correlation function is $\xi_2(r) = (r/r_0)^{-\gamma}$. The fact that the last integral that appears in this expression converges⁹⁵ justifies the use of the small-angle approximation. It means that the projected

⁹⁵ It is given by $\int_{-\infty}^{\infty} dr (1+r^2)^{-\gamma/2} = \sqrt{\pi} \Gamma((-1+\gamma)/2) / \Gamma(\gamma/2)$, which converges for $\gamma > 1$.

correlation functions are dominated by intrinsic 3D structures, that is, the major contributions come from finite values of r_i which corresponds to points that are close together in 3D space.

The small-angle approximation seems to be an excellent approximation both for w_2 and for w_3 up to $\theta \simeq 2^\circ$. This can be easily checked by numerical integration of a given model for ξ_2 and ξ_3 , see e.g. [508,48,254].

An equivalent way of looking at the small-angle approximation is to write the corresponding relations in Fourier space. The angular two-point correlation function can be written in terms of the 3D power spectrum as [364]

$$w_2(\theta) = 2\pi \int d\chi \chi^4 \psi^2(\chi) \int d^2\mathbf{k}_\perp P(k_\perp) e^{i\mathcal{D}\mathbf{k}_\perp\theta} . \quad (570)$$

Expression (570) shows that in Fourier space the small-angle approximation consists in neglecting the radial component of \mathbf{k} (to be of the order of the inverse of the depth of the survey) compared to \mathbf{k}_\perp (of the order of the inverse of the transverse size of the survey). Thus, in the small-angle approximation, the power spectrum of the projected density field is

$$P_{2D}(l) = 2\pi \int d\chi \frac{\chi^4 \psi^2(\chi)}{\mathcal{D}^2} P\left(\frac{l}{\mathcal{D}}\right) . \quad (571)$$

This can be easily generalized to higher-order correlations in Fourier space:

$$\langle \delta_{2D}(\mathbf{l}_1) \cdots \delta_{2D}(\mathbf{l}_N) \rangle_c = (2\pi)^{N-1} \delta_D(\mathbf{l}_1 + \cdots + \mathbf{l}_N) \int d\chi \frac{\chi^{2N} \psi^N(\chi)}{\mathcal{D}^{2N-2}} P_N\left(\frac{\mathbf{l}_1}{\mathcal{D}}, \dots, \frac{\mathbf{l}_N}{\mathcal{D}}\right) . \quad (572)$$

Note that the Fourier-space expression given above assumes in fact not only the small-angle approximation, but also the flat-sky approximation which neglects the curvature of the celestial sphere. General expressions for the power spectrum and higher-order correlations beyond the small-angle (and flat-sky) approximation can be derived from Eq. (567) by Legendre transforms, see e.g. [242,671].

7.2.3. Projection in the hierarchical model

The inversion of Eq. (568), to relate ξ_N in terms of w_N is still not trivial in general because the projection effects mix different scales. As in the case of the two-point correlation function, i.e. Limber's equation, it is much easier to obtain a simple relation between 3D and 2D statistics for models of ξ_N that have simple scale dependence. In the *Hierarchical model* introduced in Section 4.5.5,

$$\xi_N(\mathbf{r}_1, \dots, \mathbf{r}_N) = \sum_{a=1}^{t_N} Q_{N,a} \sum_{\text{labelings}} \prod_{\text{edges}}^{N-1} \xi_2(\mathbf{r}_A, \mathbf{r}_B) , \quad (573)$$

and, remarkably, it follows that the projected angular correlations obey a similar relation

$$w_N(\theta_1, \dots, \theta_N) = \sum_{a=1}^{t_N} q_{N,a} \sum_{\text{labelings}} \prod_{\text{edges}}^{N-1} w_2(\theta_A, \theta_B) , \quad (574)$$

where $q_{N,a}$ is simply proportional to $Q_{N,a}$. Moreover, the relation between $q_{N,a}$ and $Q_{N,a}$ depends only on the order N and is independent of the particular tree topology. To express q_N in terms of

Table 13

Projection factors for different slopes γ and parameters M_0^* and α_0 in the luminosity function

γ	M_0^*	α_0	r_3	r_4	r_5	r_6	r_7	r_8	r_9
1.7	−19.8	−1.0	1.19	1.52	2.00	2.71	3.72	5.17	7.25
1.7	−19.3	−1.2	1.21	1.57	2.12	2.93	4.13	5.88	8.44
1.7	−20.3	−0.8	1.18	1.48	1.93	2.56	3.46	4.73	6.51
1.8	−19.8	−1.0	1.20	1.55	2.08	2.85	3.98	5.62	8.00
3.0	−19.8	−1.0	1.54	2.85	5.78	12.4	27.8	63.9	150

Q_N we can use a power-law model for the two-point correlation: $\xi_2(r) = (r/r_0)^{-\gamma}$. For small angles, we thus have

$$q_N = r_N Q_N,$$

$$r_N = \frac{I_1^{N-2} I_N}{I_2^{N-1}} \quad \text{with } I_k = \int_0^\infty d\chi \chi^{2k} \mathcal{D}^{k-1} \psi^k(\chi) \mathcal{D}^{-\gamma(k-1)} (1+z)^{-3(k-1)}, \quad (575)$$

where we have taken into account redshift evolution of the two-point correlation function in the non-linear regime assuming stable clustering (see Section 4.5.2), $\xi_2(r, z) = \xi_2(r)(1+z)^{-3}$. The integrals I_k are just numerical values that depend on the selection function and γ . The values of ψ^* and ϕ^* in Eq. (565) are thus irrelevant for q_N . The only relevant parameters in the luminosity function are M^* and α .

The resulting values of r_N increase with γ and M^* and decrease with α , but do not change much within the uncertainties in the shape of the luminosity function (see Section 56 in [508], and [249]). This is illustrated in Table 13 where values of r_N are plotted for different parameters in the selection function. In the analysis of the APM, variations of γ are only important for very large scales, $\theta > 3^\circ$, where γ changes from 1.8 to 3. In this case r_N displays a considerable variation and Eq. (575) is not a good approximation.

As an example we can consider the selection function given by the characteristic “bell shape” in a magnitude-limited sample:

$$\psi(r) \propto r^{-b} \exp[-r^2/\mathcal{D}^2], \quad (576)$$

where \mathcal{D} is related to the effective sample depth and b is typically a small number (e.g. for the APM $b \simeq 0.1$ and $\mathcal{D} \simeq 350$ Mpc/h). For this selection function and a power law $P(k) \propto k^n$ (e.g. $\gamma = -(n+3)$) we can calculate r_3 explicitly.

$$r_3 = \frac{8}{3\sqrt{3}} \left(\frac{\sqrt{27}}{4} \right)^b \frac{\Gamma[3/2 - b/2] \Gamma[3/2 - n - 3/2b]}{\Gamma[3/2 - n/2 - b]^2} \left(\frac{3}{2} \right)^n. \quad (577)$$

For $b = 0$ and $n = 0$ we find $r_3 = 8/3\sqrt{3} \simeq 1.54$, while for $b = 0$ and $n = -1$, closer to the APM case, $r_3 = 2\pi/3\sqrt{3} \simeq 1.21$, comparable to the values given in [249].

It is important to note that although r_N are unaffected by changes in ψ , the overall normalization of I_k can change significantly. Because of this, while the amplitude of ξ_2 is uncertain by 40% for $\Delta M^* = 1.0$ and $\Delta\alpha = 0.4$, the corresponding uncertainty in r_3 is only 2%. This is an excellent motivation for using the hierarchical ratios q_N as measures of clustering.

Note that the above hierarchical prediction could only provide a good approximation to clustering observations at small scales, where the hierarchical model in Eq. (573) might be a good approximation (see Sections 4.5.5 and 8.2.4). On larger scales, accurate predictions require projection using the PT hierarchy, which is different from Eq. (573), as the N -point correlation functions have a significant shape dependence (see Section 4.1). Despite this ambiguity on how to compare angular observations to theoretical predictions, note that these two approaches give results that agree within 20% (e.g. see Fig. 47).

7.2.4. The correlation hierarchy for the projected density

We can define the area-averaged angular correlations $\bar{\omega}_p(\theta)$ in terms of the angular correlation functions $w_N(\theta_1, \dots, \theta_N)$:

$$\bar{\omega}_p(\theta) \equiv \frac{1}{A^p} \int_A dA_1 \cdots dA_p w_p(\theta_1, \dots, \theta_p) = \langle \delta_{2D}^p(\theta) \rangle_c, \quad (578)$$

where $A = 2\pi(1 - \cos\theta)$ is the solid angle of the cone, $dA_p = \sin\theta_p d\theta_p d\varphi_p$ and $\delta_{2D}(\theta)$ is the density contrast inside the cone. Thus $\bar{\omega}_p(\theta)$ only depends on the size of the cone, θ , analogous to smoothed moments in the 3D case. The use of Eq. (568) leads to

$$\begin{aligned} \bar{\omega}_p(\theta) &= \frac{1}{A^p} \int d\chi_1 \chi_1^{3p-1} \psi_1^p \int \prod_{i=1}^p dA_i \\ &\times \int_{-\infty}^{\infty} \prod_{i=2}^p dr_i \xi_p[(\chi_1, \mathcal{D}_1\theta_1), \dots, (\chi_p, \mathcal{D}_1\theta_p)]. \end{aligned} \quad (579)$$

One can see that the cumulants of the projected density are thus line-of-sight averages of the density cumulants in *cylindrical* window function,

$$\langle \delta_{\text{proj}, \theta}^p \rangle_c = \int d\chi \chi^{2p} \psi^p(\chi) \langle \delta_{\mathcal{D}\theta, \text{cyl}}^p \rangle_c L^{p-1}, \quad (580)$$

where $\delta_{\mathcal{D}\theta, \text{cyl}}^p$ is the filtered 3D density with a cylindrical filter of transverse size $\mathcal{D}\theta$ and depth L . For instance, written in terms of the power spectrum, the second moment reads

$$\bar{\omega}_2(\theta) = 2\pi \int d\chi \chi^4 \psi^2(\chi) \int d^2\mathbf{k}_\perp P(k_\perp) W_{2D}^2(\mathcal{D}\theta k_\perp), \quad (581)$$

where W_{2D} is the top-hat 2D window function,

$$W_{2D}(l\theta) = 2 \frac{J_1(l\theta)}{l\theta}. \quad (582)$$

Relation (580) shows that the cumulant hierarchy is preserved. If we define the s_p parameters in angular space,

$$s_p(\theta) \equiv \frac{\bar{\omega}_p(\theta)}{[\bar{\omega}_2(\theta)]^{p-1}}, \quad (583)$$

it follows that they are all finite and independent of L .

In the weakly non-linear regime, we can compute exactly the hierarchy for the projected density because the density cumulants for a cylindrical window are those obtained for the 2D dynamics (see Section 5.9). In case of a power-law spectrum the s_p are independent of the filtering scale. The line-of-sight integrations can then be performed explicitly.⁹⁶ Using Eq. (580) and the results of Section 5.9 gives⁹⁷

$$s_p = r_p S_p^{2D}, \quad (584)$$

$$r_p = \frac{I_1^{p-2} I_p}{I_2^{p-1}} \quad \text{with } I_k = \int_0^\infty d\chi \chi^{2k} \psi^k(\chi) \mathcal{D}^{-(n+3)(k-1)} D_1^{2k-2}(z). \quad (585)$$

Note that the r_p coefficients are very similar to those in the non-linear case except that the redshift evolution of the fluctuation is taken here to be given by the linear growth rate. This is actually relevant only when the redshift under consideration is comparable to unity.

An interesting point is that it may seem inconsistent to use both tree-level PT predictions and the small-angle approximation, as a priori it is not clear whether their regimes of validity overlap. As shown in [254] for characteristic depths comparable to APM there is at least a factor of five in scale where both approximations are consistent, depending on the 3D power spectrum shape. For deeper surveys, of course, the consistency range is increased, so this is a meaningful approach.

As expected, similar results hold for the hierarchy of correlation functions in the weakly non-linear regime. The results for the angular three-point function and bispectrum have been studied with most detail [242,225,101,671]. From Eqs. (571) and (572) and for power-law spectra, it follows that the configuration dependence of the bispectrum is conserved by projection, only the amplitude is changed by the projection factor r_3 , as in Eq. (585) [275,242,225,101]. However, as soon as the spectral index changes significantly on scales comparable to those sampled by the selection function, this simple result does not hold anymore [242]. A number of additional results regarding the shape dependence of projected correlations include, (i) a study of the dependence on configuration shape as a function of depth [81], that also includes redshift-dependent galaxy biasing; (ii) the power of angular surveys to determine bias parameters from the projected bispectrum in spherical harmonics [671] and (iii) comparisons of PT predictions and numerical simulations in angular space [225], as we summarize in the next section.

⁹⁶ For CDM models a semianalytic result can be obtained for the skewness, see [524] for details.

⁹⁷ It is important to note that in Eq. (584) the coefficients S_p^{2D} need to be used and not those corresponding to 3D top-hat filtering as suggested by the tree hierarchical model.

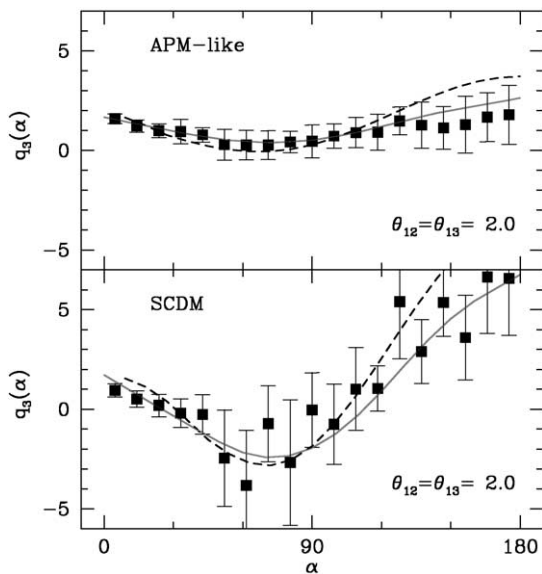


Fig. 46. Projected leading order PT predictions (solid curves) and N -body results (points with sampling errors) for the angular three-point amplitude $q_3(\alpha)$ at fixed $\theta_{12} = \theta_{13} = 2^\circ$ for a survey with the APM selection function. N -body results correspond to the average and variance of five realizations of the APM-like model (top) and the SCDM model (bottom). The dashed lines show the corresponding PT predictions for $r_{12} = r_{13} = 15 \text{ Mpc}/h$ projected with the hierarchical model.

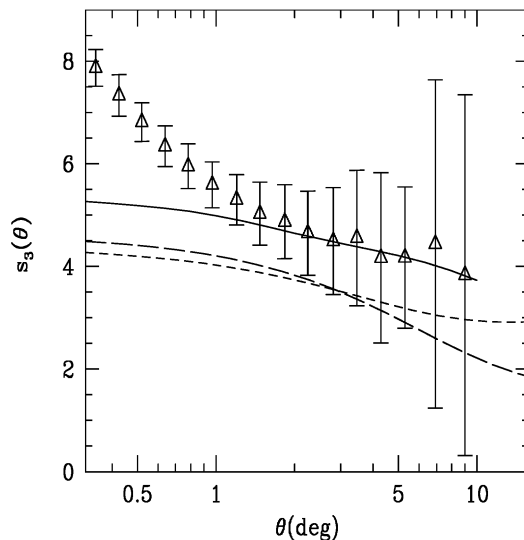


Fig. 47. Tree-level PT predictions for the APM-like power spectrum (solid curves) and corresponding N -body results (points with sampling errors) for the projected smoothed skewness $s_3(\theta)$ as a function of the radius θ (in degrees) of the cells in the sky. The short- and long-dashed line show the hierarchical prediction $s_3 \simeq r_3 S_3$, see text for details.

7.2.5. Comparison with numerical simulations

We now illustrate the results described in the previous section and compare their regime of validity against numerical simulations.

Fig. 46 shows the angular three-point correlation function for APM-like and SCDM spectrum projected to the depth of the APM survey, see [225] for more details. As discussed before, the configuration dependence of the three-point amplitude is quite sensitive to the shape of the power spectrum. Both the shape and amplitude of $q_3(\alpha)$ predicted by PT (solid curves) are reproduced by the N -body results (points) even on these moderately small scales.⁹⁸ The error bars in the simulation results are estimated from the variance between 5 maps from different N -body realizations and have been scaled to $1 - \sigma$ uncertainties for a single observer. The dashed lines correspond to the results of the 3D Q_3 for $r_1 = r_2 = 15h^{-1} \text{ Mpc}$ multiplied by the hierarchical projection factor in Eq. (575), e.g. $q_3 = Q_3 r_3$. The model seems to work well for small α , but there are significant deviations for large α , which illustrate that this projection model does not work well, as discussed above.

⁹⁸ At the mean depth of the APM, two degrees corresponds to $\simeq 15h^{-1} \text{ Mpc}$.

In the weakly non-linear regime, the third moment of smoothed angular fluctuations, defined in (579), can be explicitly written in terms of the power spectrum using PT. It is given by

$$\begin{aligned} \bar{\omega}_3 = & 6(2\pi)^2 \int d\chi \chi^6 \psi^3(\chi) \left[\frac{6}{7} \left(\int k dk W_{2D}^2(k\mathcal{D}\theta) P(k) \right)^2 \right. \\ & \left. \times \frac{1}{2} \int k dk W_{2D}^2(k\mathcal{D}\theta) P(k) \int k^2 dk \mathcal{D}\theta W_{2D}(k\mathcal{D}\theta) W'_{2D}(k\mathcal{D}\theta) P(k) \right], \end{aligned} \quad (586)$$

where W'_{2D} is the derivative of the top-hat window W_{2D} defined in Eq. (582). Therefore, in case of a power-law spectrum $P(k) \sim k^n$, we have [48],

$$s_3 = r_3 \left(\frac{36}{7} - \frac{3}{2}(n+2) \right), \quad (587)$$

with r_3 given in general by Eq. (585). The coefficient r_3 is found in practice to be of order unity and to be very weakly dependent on the adopted shape for the selection function.

It is worth to note that the hierarchical model in Section 7.2.3 yields a different prediction for s_3 than the above tree-level value. In the hierarchical case, $s_3 \simeq r_3 S_3$ [249,250] with $S_3 = \frac{34}{7} - (n+3)$. For example, for $n \simeq -1$, the hierarchical model yields $s_3 \simeq 3.43$ while the tree-level prediction yields $s_3 \simeq 4.38$. This difference becomes smaller as we move toward larger n (e.g. larger scales), being zero at $n = \frac{4}{7}$, but it is significant for the range of scales probed with current observations, even after taking uncertainties into account.

Fig. 47 compares the predictions for the angular skewness s_3 by tree-level PT (solid lines) for a power spectrum that matches the APM catalogue and the APM measurements (triangles). These predictions correspond to a numerical integration of PT predictions in Eq. (587) [254]. The dashed lines show the “naive” hierarchical prediction $s_3 \simeq r_3 S_3$ at the angular scale $\theta \simeq R/\mathcal{D}$ given by the depth, \mathcal{D} , of the survey. The long-dashed line uses a fixed value of $r_3 = 1.2$, while the dashed line corresponds to $r_3 = r_3(n)$ given by Eq. (577) with $n = -(3 + \gamma)$ given by the logarithmic slope of the variance of the APM-like $P(k)$ at the angular scale $\theta \simeq R/\mathcal{D}$. These results are compared with the mean of 20 all sky simulations described in [254] (error bars correspond to the variance in 20 observations). As can be seen in the figure, the hierarchical model gives a poor approximation, while the projected tree-level results match well the simulations for scales $\theta \gtrsim 1^\circ$, which correspond to the weakly non-linear regime where $\xi_2 \lesssim 1$. On small scales the discrepancies between the tree-level results and the simulations are due to 3D non-linear effects and also projection: on small scales the simulations follow the hierarchical model in Eq. (573), rather than the PT predictions, and therefore $s_3 \simeq r_3 S_3$ gives a good approximation, but S_3 should be the non-linear 3D value (for example, as given by HEPT or EPT, see Sections 4.5.6 and 5.13, respectively). Further comparisons with numerical simulations for s_3 and s_4 are presented in Fig. 54 together with observational results.

7.2.6. Reconstructing the PDF of the projected density

It is interesting to note that it is possible to write down a functional relation between the cumulant generating function defined in Eq. (141) for the projected density, $\varphi_{\text{proj}}(y)$, and the one corresponding

to cylindrical filtered cumulants, $\varphi_{\text{cyl}}(y)$ [659,468,58]. This can be done from relation (580) which straightforwardly leads to

$$\varphi_{\text{proj}}(y) = \int \frac{d\chi}{\Xi_{\theta}(\chi)} \varphi_{\text{cyl}}[y\chi^2\psi(\chi)\Xi_{\theta}(\chi)] \quad (588)$$

with

$$\Xi_{\theta}(\chi) = \frac{\langle \delta_{\mathcal{D}\theta, \text{cyl}}^2 \rangle}{\langle \delta_{2\text{D}}^2 \rangle} L, \quad (589)$$

which can be rewritten in terms of the matter fluctuation power spectrum,

$$\Xi_{\theta}(\chi) = \frac{\int d^2\mathbf{k} P(k, z) W^2(k\mathcal{D}\theta)}{\int d\chi' \chi'^4 \psi^2(\chi') \int d^2\mathbf{k} P(k, z') W^2(k\mathcal{D}'\theta)}. \quad (590)$$

In this expression we have explicitly written the redshift dependence of the power spectrum. In the case of a power-law spectrum,

$$P(k, z) = P_0(z) \left(\frac{k}{k_0} \right)^n \quad (591)$$

and it takes a much simpler form given by

$$\Xi_{\theta}(\chi) = \frac{P_0(z) \mathcal{D}^{-n-2}}{\int d\chi' \chi'^4 \psi^2(\chi') P_0(z') \mathcal{D}'^{-n-2}}. \quad (592)$$

Together with Eq. (588) this result provides the necessary ingredients to reconstruct the one-point PDF of the projected density with an inverse Laplace transform of $\varphi_{\text{proj}}(y)$. Note that projection effects alter the shape of the singularity in $\varphi(y)$ though it preserves the large-density exponential cutoff [659,58]. So far this approach has only been used in the literature to study the reconstruction of the one-point PDF of the local convergence field in the context of weak lensing observations [659,468].

We now turn to a brief summary of the basics of weak lensing and its connections to projection effects.

7.3. Weak gravitational lensing

The first theoretical investigations on the possibility of mapping the large-scale structure of the universe with weak gravitational lensing date back to the early 1990's [69,70,456,364]. It was then shown that the number of background galaxies was large enough to serve as tracers of the deformation field induced by the intervening large-scale structures. In this context the observation of a coherent shear pattern in the orientation of background galaxies is interpreted as due to lensing effects caused by the mass concentration along the line of sight. The potential interest of such observations has led to further theoretical investigations such as the determination of the dependence of the results on cosmological parameters [676,53,339,665], and to extensive observational efforts. The latter have recently led to the first reliable detections of the so-called ‘‘cosmic shear’’ [666,11,693,365].

Although in nature totally different from galaxy counts, it is worth pointing out that such observations eventually aim at mapping the line-of-sight mass fluctuations so that techniques developed for studying galaxy angular catalogues can be applied. Here we briefly introduce the physics of lensing

with emphasis on connections to angular clustering. More comprehensive presentations can be found in [25,453].

7.3.1. The convergence field as a projected mass map

The physical mechanism at play in weak lensing surveys is the deflection of photon paths in gravitational potential fields. The deflection angle per unit distance, $\delta\vec{\alpha}/\delta s$, can be obtained from simple computations of the geodesic equation in the weak field limit.⁹⁹ When the metric fluctuations are purely scalar, the deflection angle reads

$$\frac{\delta\vec{\alpha}}{\delta s} = -2\vec{\nabla}_x \frac{\phi}{c^2}, \quad (593)$$

where the spatial derivative is taken in a plane that is orthogonal to the photon trajectory.

The direct consequence of this bending is a displacement of the apparent position of the background objects. This depends on the distance of the source plane, D_{OS} , and on the distance between the lens plane and the source plane D_{LS} . It is given by

$$\vec{\gamma}^S = \vec{\gamma}^I - \frac{2}{c^2} \frac{D_{LS}}{D_{OS} D_{OL}} \vec{\nabla}_\gamma \left(\int ds \phi(s, \gamma) \right), \quad (594)$$

where $\vec{\gamma}^I$ is the position in the image plane and $\vec{\gamma}^S$ is the position in the source plane. The gradient is taken here with respect to the angular position (this is why a D_{OL} factor, distance to the lens plane, appears). The total deflection is obtained by an integration along the line of sight, assuming the lens is thin compared to its angular distance. Calculations are also usually done in the so-called Born approximation for which the potential is computed along the unperturbed photon trajectory.

The observable effect which is aimed at, however, is the induced deformation of background objects. Such an effect is due to the *variations* of the displacement field with respect to the apparent position. These variations induce a change in both size and shape of the background objects which are encoded in the amplification matrix, \mathcal{A} , describing the linear change between the source plane and the image plane:

$$\mathcal{A} = \begin{pmatrix} \frac{\partial \gamma_i^I}{\partial \gamma_j^S} \end{pmatrix}. \quad (595)$$

Its inverse, \mathcal{A}^{-1} , is actually directly calculable in terms of the gravitational potential. It is given by the derivatives of the displacement with respect to the apparent position

$$\mathcal{A}^{-1} \equiv \frac{\partial \gamma_i^S}{\partial \gamma_j^I} = \delta_{ij} - 2 \frac{D_{LS}}{D_{OS} D_{OL}} \varphi_{,ij}, \quad (596)$$

where φ is the projected potential. Usually its components are written as

$$\mathcal{A}^{-1} = \begin{pmatrix} 1 - \kappa - \gamma_1 & -\gamma_2 \\ -\gamma_2 & 1 - \kappa + \gamma_1 \end{pmatrix}, \quad (597)$$

taking advantage of the fact that it is a symmetric matrix. The components of this matrix are expressed in terms of the convergence, κ (a scalar field) and the shear, γ (a pseudo-vector field).

⁹⁹ See e.g. [457,544] for a comprehensive presentation of these calculations.

The key idea for weak lensing observations is then that collection of tiny deformation of background galaxies can be used to measure the local *shear* field from which the projected potential, and therefore the convergence field, can be reconstructed [364]. The latter has a simple cosmological interpretation: from the trace of Eq. (596) one obtains the convergence¹⁰⁰

$$\kappa(\gamma) = \frac{3}{2} \Omega_m \int dz_s n(z_s) \int d\chi \frac{\mathcal{D}(\chi_s, \chi) \mathcal{D}(\chi)}{\mathcal{D}(\chi_s)} \delta(\chi, \gamma) (1+z) \quad (598)$$

as the integrated line-of-sight density contrast. In Eq. (598), χ is the distance along the line of sight and \mathcal{D} are the angular distances. In this relation sources are assumed to be located at various redshifts with a distribution $n(z_s)$ normalized to unity, and all the distances are expressed in units of c/H_0 . Relation (598) is then entirely dimensionless. Note that in general the relation between the redshift and the distances depends on cosmological parameters, see Eq. (563).

7.3.2. Statistical properties

To gain insight into the expected statistical properties of the convergence field, it is important to keep in mind that in Eq. (598) the convergence κ is not normalized as would be the local projected density contrast. The projected density contrast is actually given by

$$\delta_{2D}(\gamma) = \frac{\kappa}{\bar{\omega}}, \quad (599)$$

where $\bar{\omega}$ is the mean lens efficiency,

$$\bar{\omega} = \frac{3}{2} \Omega_m \int dz_s n(z_s) \int d\chi \frac{\mathcal{D}(\chi_s, \chi) \mathcal{D}(\chi)}{\mathcal{D}(\chi_s)} (1+z). \quad (600)$$

This implies that the skewness of the convergence field is then given by

$$s_3^\kappa = \frac{s_3^{\text{proj}}}{\bar{\omega}}, \quad (601)$$

where s_3^{proj} is the skewness of the projected density contrast given by Eq. (587). As a consequence, the skewness of κ is expected to display a strong Ω_m dependence. This property has indeed been found in [53] where it has been shown using PT that

$$s_3^\kappa \approx 40 \Omega_m^{-0.75} \quad (602)$$

for sources at redshift unity.¹⁰¹ This result has been subsequently extended to the non-linear regime [339,326,467,469,666,158], higher-order moments, the bispectrum [158], and to more complex quantities such as the shape of the one-point PDF of the convergence field [658,659,468].

7.3.3. Next to leading order effects

Contrary to the previous cases, corrections to the previous leading order PT results, e.g. Eq. (602), do not involve only next to leading order terms due to the non-linear dynamics but also other couplings that appear specifically in the weak lensing context. Let us list and comment the most significant of them:

¹⁰⁰ In these sections, Ω_m is understood to be at $z=0$.

¹⁰¹ For the same reasons that s_3^κ has a strong Ω_m dependence, it also depends significantly on the source redshift distribution.

- (i) An exact integration of the lens equations leads to lens–lens coupling and departures from the Born approximation. This induces extra couplings that have been found to be in all cases negligible for a source population at redshift of about unity [53,667].
- (ii) The source population clustering properties can also induce non-trivial effects as described in [55]. This is due to the fact that the source plane is by itself a random media which introduces further couplings due to either intrinsic galaxy number fluctuations or due to overlapping of lens and source populations. These effects have been found to be small if the redshift distribution of the sources is narrow enough [55,284] which might indeed put severe constraints on the observations.
- (iii) The magnification effect (when κ is large, galaxies are enlarged and can thus be more efficiently detected) could also induce extra couplings. Although it is difficult to estimate the extent of such an effect, it appears to have only modest effects on the high-order statistical properties of the convergence field [285].

Finally, it is important to note that the first reports of cosmic shear detections have been challenged by suggestions that part of the signal at small scales might be due to intrinsic galaxy shape correlations [308,166,121]. This is a point that should be clarified by further investigations.

7.3.4. Biasing from weak gravitational lensing

With the arrival of wide surveys dedicated to weak lensing observations,¹⁰² a very powerful new window to large-scale structure properties is being opened. Weak lensing observations can indeed be used not only to get statistical properties of the matter density field, but also to map the mass distribution in the universe. In particular, it becomes possible to explore the galaxy–mass local relation [664]. Potentially, galaxy formation models, biasing models, can be directly tested by these observations. It is indeed possible to measure the correlation coefficient r_κ of the convergence field κ with the projected density contrast of the (foreground) galaxy $\delta_{g,2D}$,

$$r_\kappa = \frac{\langle \kappa \delta_{g,2D} \rangle}{\sqrt{\langle \kappa^2 \rangle \langle \delta_{g,2D}^2 \rangle}}, \quad (603)$$

a quantity which, within geometrical factors, is proportional to the r coefficient defined in Eq. (531). What has been measured so far [315] is however $\langle \kappa \delta_{g,2D} \rangle / \langle \delta_{g,2D}^2 \rangle$, that is, a quantity that roughly scales like $\Omega_m r / b$. Pioneering results suggest a scale-independent r/b parameter of about unity for the favored cosmological model ($\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$) [315]. Such observations are bound to become common place in the coming years and will provide valuable tests for galaxy formation models.

7.4. Redshift distortions

In order to probe the 3D distribution of galaxies in the universe, galaxy redshifts are routinely used as an indicator of radial distance from the observer, supplemented by the 2D angular position on the sky. The Hubble expansion law tells us that the recession velocity of an object is proportional to its distance from us; however, the observed velocity has also a contribution from peculiar velocities, which are generated due to the dynamics of clustering and are unrelated to the Hubble expansion

¹⁰² See for example, <http://terapix.iap.fr/Descart/>.

and thus contaminate the distance information. Therefore, the clustering pattern in “redshift space” is somewhat different from the actual real-space distribution. This is generically known as “redshift distortions”.

At large scales, the main effect of peculiar velocities is due to galaxies infall into clusters. Galaxies between us and the cluster have their infall velocities added to the Hubble flow and thus appear farther away in redshift space, whereas those galaxies falling into the cluster from the far side have their peculiar velocities subtracted from the Hubble flow, and thus appear closer to us than in real space. As a consequence of this, large-scale structures in redshift space appear flattened or “squashed” along the line of sight. On the other hand, at small scales (smaller than the typical cluster size) the main effect of peculiar velocities is due to the velocity dispersion from virialization. This causes an elongation along the line of sight of structures in redshift space relative to those in real space, the so-called “finger of God” effect (which points to the observer’s location).

7.4.1. The density field in redshift space

We now discuss the effects of redshift distortions on clustering statistics such as the power spectrum, the bispectrum and higher-order moments of the smoothed density field. See the exhaustive review [295] for details on theoretical description of linear redshift distortions and observational results. In redshift space, the radial coordinate \mathbf{s} of a galaxy is given by its observed radial velocity, a combination of its Hubble flow plus “distortions” due to peculiar velocities. Here we restrict to the “plane-parallel” approximation so that the line of sight is taken as a fixed direction, denoted by \hat{z} . Plane-parallel distortions maintain statistical homogeneity, so Fourier modes are still the natural basis in redshift space. On the other hand, statistical isotropy is now broken because clustering along the line of sight is different from that in the perpendicular directions.

However, when the radial character of redshift distortions is taken into account, the picture changes. Radial distortions respect statistical isotropy (about the observer), but break statistical homogeneity (since there is a preferred location, the observer’s position). In this case Fourier modes are no longer special, in particular, the power spectrum is no longer diagonal [703]. Alternative approaches to Fourier modes have been suggested in the literature [306,292,616], here we review the simplest case of plane-parallel distortions where most of the results have been obtained. We should note that this is not just of academic interest, it has been checked with N -body simulations that results on monopole averages of different statistics carry over to the radial case with very small corrections [566].

The mapping from real-space position \mathbf{x} to redshift space in the plane-parallel approximation is given by

$$\mathbf{s} = \mathbf{x} - f v_z(\mathbf{x}) \hat{z}, \quad (604)$$

where $f(\Omega_m) \approx \Omega_m^{0.6}$ is the logarithmic growth rate of linear perturbations, and $\mathbf{v}(\mathbf{x}) \equiv -\mathbf{u}(\mathbf{x})/(\mathcal{H}f)$, where $\mathbf{u}(\mathbf{x})$ is the peculiar velocity field, and $\mathcal{H}(\tau) \equiv (1/a)(da/d\tau) = Ha$ is the conformal Hubble parameter (with FRW scale factor $a(\tau)$ and conformal time τ). The density field in redshift space, $\delta_s(\mathbf{s})$, is obtained from the real-space density field $\delta(\mathbf{x})$ by requiring that the redshift-space mapping conserves mass, i.e.

$$(1 + \delta_s) d^3\mathbf{s} = (1 + \delta) d^3\mathbf{x}. \quad (605)$$

Using the fact that $d^3\mathbf{s}=J(\mathbf{x})d^3\mathbf{x}$, where $J(\mathbf{x})=|1-f\nabla_z v_z(\mathbf{x})|$ is the *exact* Jacobian of the mapping in the plane-parallel approximation, it yields

$$\delta_s(\mathbf{s}) = \frac{\delta(\mathbf{x}) + 1 - J(\mathbf{x})}{J(\mathbf{x})}. \quad (606)$$

The zeros of the Jacobian describe caustics in redshift space, the locus of points where the density field is apparently infinite [450]. This surface is characterized in real space by those points which are undergoing turn-around in the gravitational collapse process, so their peculiar velocities exactly cancel the differential Hubble flow. In practice, caustics are smoothed out by sub-clustering, see e.g. the discussion in [330].

An expression for density contrast in redshift space follows from Eq. (606) [562]:

$$\delta_s(\mathbf{k}) = \int \frac{d^3\mathbf{x}}{(2\pi)^3} e^{-i\mathbf{k}\cdot\mathbf{x}} e^{ifk_z v_z(\mathbf{x})} [\delta(\mathbf{x}) + f\nabla_z v_z(\mathbf{x})], \quad (607)$$

where we assumed here that only points where $f\nabla_z v_z(\mathbf{x}) < 1$ contribute. The only other approximation in this expression is the use of the plane-parallel approximation, i.e. this is a fully non-linear expression. To obtain a perturbative expansion, we expand the second exponential in power series,

$$\begin{aligned} \delta_s(\mathbf{k}) &= \sum_{n=1}^{\infty} \int d^3\mathbf{k}_1 \cdots d^3\mathbf{k}_n [\delta_{\text{D}}]_n [\delta(\mathbf{k}_1) + f\mu_1^2 \theta(\mathbf{k}_1)] \frac{(f\mu k)^{n-1}}{(n-1)!} \\ &\quad \times \frac{\mu_2}{k_2} \theta(\mathbf{k}_2) \cdots \frac{\mu_n}{k_n} \theta(\mathbf{k}_n), \end{aligned} \quad (608)$$

where $[\delta_{\text{D}}]_n \equiv \delta_{\text{D}}(\mathbf{k} - \mathbf{k}_1 - \cdots - \mathbf{k}_n)$, the velocity divergence $\theta(\mathbf{x}) \equiv \nabla \cdot \mathbf{v}(\mathbf{x})$ and $\mu_i \equiv \mathbf{k}_i \cdot \hat{z}/k_i$ is the cosine of the angle between the line of sight and the wave vector. In linear PT, only the $n = 1$ term survives, and we recover the well-known formula due to Kaiser [362]:

$$\delta_s(\mathbf{k}) = \delta(\mathbf{k})(1 + f\mu^2). \quad (609)$$

Eq. (608) can be used to obtain the redshift-space density field beyond linear theory. In redshift space we can write

$$\delta_s(\mathbf{k}, \tau) = \sum_{n=1}^{\infty} D_1^n(\tau) \int d^3\mathbf{k}_1 \cdots \int d^3\mathbf{k}_n [\delta_{\text{D}}]_n Z_n(\mathbf{k}_1, \dots, \mathbf{k}_n) \delta_1(\mathbf{k}_1) \cdots \delta_1(\mathbf{k}_n), \quad (610)$$

where $D_1(\tau)$ is the density perturbation growth factor in linear theory, and we have assumed that the n th-order growth factor $D_n \propto D_1^n$, which is an excellent approximation (see [560], Appendix B.3). Since a local deterministic and non-linear bias can be treated in an equal footing as non-linear dynamics, it is possible to obtain the kernels Z_n including biasing and redshift distortions. From Eqs. (525) and (608)–(610), the redshift-space kernels Z_n for the *galaxy* density field read [669,562]¹⁰³

$$Z_1(\mathbf{k}) = (b_1 + f\mu^2), \quad (611)$$

¹⁰³ Detailed expressions for the second-order solutions are given in [313] including the (small) dependences on Ω_m for the unbiased case.

$$Z_2(\mathbf{k}_1, \mathbf{k}_2) = b_1 F_2(\mathbf{k}_1, \mathbf{k}_2) + f \mu^2 G_2(\mathbf{k}_1, \mathbf{k}_2) + \frac{f \mu k}{2} \left[\frac{\mu_1}{k_1} (b_1 + f \mu_2^2) + \frac{\mu_2}{k_2} (b_1 + f \mu_1^2) \right] + \frac{b_2}{2}, \quad (612)$$

$$Z_3(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = b_1 F_3^{(s)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) + f \mu^2 G_3^{(s)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) + f \mu k [b_1 F_2^{(s)}(\mathbf{k}_1, \mathbf{k}_2) + f \mu_{12}^2 G_2^{(s)}(\mathbf{k}_1, \mathbf{k}_2)] \frac{\mu_3}{k_3} + f \mu k (b_1 + f \mu_1^2) \frac{\mu_{23}}{k_{23}} G_2^{(s)}(\mathbf{k}_2, \mathbf{k}_3) + \frac{(f \mu k)^2}{2} (b_1 + f \mu_1^2) \frac{\mu_2}{k_2} \frac{\mu_3}{k_3} + 3 b_2 F_2^{(s)}(\mathbf{k}_1, \mathbf{k}_2) + \frac{b_3}{6}, \quad (613)$$

where we denote $\mu \equiv \mathbf{k} \cdot \hat{z}/k$, with $\mathbf{k} \equiv \mathbf{k}_1 + \dots + \mathbf{k}_n$, and $\mu_i \equiv \mathbf{k}_i \cdot \hat{z}/k_i$. As above, F_2 and G_2 denote the second-order kernels for the real-space density and velocity-divergence fields, and similarly for F_3 and G_3 . Note that the third-order kernel Z_3 must still be symmetrized over its arguments. One can similarly obtain the PT kernels Z_n in redshift space to arbitrary higher order.

We note that there are *two* approximations involved in this procedure: one is the mathematical step of going from Eq. (607) to Eq. (608), which approximates the redshift-space mapping with a power series; the other is the PT expansion itself (i.e. the expansion of $\delta(\mathbf{k})$ and $\theta(\mathbf{k})$ in terms of linear fluctuations $\delta_1(\mathbf{k})$). Therefore, one is not guaranteed that the resulting PT in redshift space will work over the same range of scales as in real space. In fact, in general, *PT in redshift space breaks down at larger scales than in real space*, because the redshift-space mapping is only treated approximately, and it breaks down at larger scales than does the perturbative dynamics. In particular, a calculation of the one-loop power spectrum in redshift space using Eqs. (611)–(613) does not give satisfactory results because expanding the exponential in Eq. (607) is a poor approximation. To extend the leading order calculations, one must treat the redshift-space mapping exactly and only approximate the dynamics using PT [562]. To date, this program has only been carried out using the ZA [220,642,301] and second-order Lagrangian PT [565], as we shall discuss below.

7.4.2. The redshift-space power spectrum

The calculation of redshift-space statistics in Fourier space proceeds along the same lines as in the un-redshifted case. To leading (linear) order, the redshift-space galaxy power spectrum reads [362]

$$P_s(\mathbf{k}) = P_g(k) (1 + \beta \mu^2)^2 = \sum_{\ell=0}^{\infty} a_\ell P_\ell(\mu) P_g(k), \quad (614)$$

where $P_g(k) \equiv b_1^2 P(k)$ is the real-space galaxy power spectrum, $P(k)$ is the linear mass power spectrum and $\beta \equiv f/b_1 \approx \Omega_m^{0.6}/b_1$. Here $P_\ell(\mu)$ denotes the Legendre polynomial of order ℓ , and the multipole coefficients are [290,131]

$$a_0 \equiv 1 + \frac{2}{3} \beta + \frac{1}{5} \beta^2, \quad a_2 \equiv \frac{4}{3} \beta + \frac{4}{7} \beta^2, \quad a_4 \equiv \frac{8}{35} \beta^2; \quad (615)$$

all other multipoles vanish. Eq. (614) is the standard tool for measuring Ω_m from redshift distortions of the power spectrum in the linear regime; in particular, the quadrupole-to-monopole ratio $R_p \equiv a_2/a_0$ should be a constant, independent of wave vector k , as $k \rightarrow 0$. Note, however, that in these expressions Ω_m appears only through the parameter β , so there is a degeneracy between Ω_m and the linear bias factor b_1 . Eq. (615) assumes deterministic bias, for stochastic bias extensions see [517,180].

From Eq. (607), we can write a simple expression for the power spectrum in redshift space, $P_s(\mathbf{k})$:

$$P_s(\mathbf{k}) = \int \frac{d^3\mathbf{r}}{(2\pi)^3} e^{-i\mathbf{k}\cdot\mathbf{r}} \langle e^{i\lambda\Delta v_z} [\delta(\mathbf{x}) + f\nabla_z v_z(\mathbf{x})][\delta(\mathbf{x}') + f\nabla'_z v_z(\mathbf{x}')] \rangle, \quad (616)$$

where $\lambda \equiv fk\mu$, $\Delta v_z \equiv v_z(\mathbf{x}) - v_z(\mathbf{x}')$, $\mathbf{r} \equiv \mathbf{x} - \mathbf{x}'$. This is a fully non-linear expression, no approximation has been made except for the plane-parallel approximation. In fact, Eq. (616) is the Fourier analog of the so-called “streaming model” [508], as modified in [219] to take into account the density–velocity coupling.

The physical interpretation of this result is as follows. The factors in square brackets denote the amplification of the power spectrum in redshift space due to infall (and they constitute the only contribution in linear theory, giving Kaiser’s [362] result). This gives a positive contribution to the quadrupole ($l = 2$) and hexadecapole ($l = 4$) anisotropies. On the other hand, at small scales, as k increases the exponential factor starts to play a role, decreasing the power due to oscillations coming from the pairwise velocity along the line of sight. This leads to a decrease in monopole and quadrupole power with respect to the linear contribution; in particular, the quadrupole changes sign.

In order to describe the non-linear behavior of the redshift-space power spectrum, it has become popular to resort to a phenomenological model to take into account the velocity dispersion effects [493]. In this case, the non-linear distortions of the power spectrum in redshift space are written in terms of the linear squashing factor and a suitable damping factor due to the pairwise-velocity distribution function

$$P_s(\mathbf{k}) = P_g(k) \frac{(1 + \beta\mu^2)^2}{[1 + (k\mu\sigma_v)^2/2]^2}. \quad (617)$$

Here σ_v is a free parameter that characterizes the velocity dispersion along the line of sight. This Lorentzian form of the damping factor is motivated by empirical results showing an exponential one-particle¹⁰⁴ velocity distribution function [489]; comparison with N -body simulations have shown it to be a good approximation [132]; however, these type of phenomenological models tend to approach the linear PT result faster than numerical simulations [301]. In addition, although σ_v can be chosen to fit, say, the quadrupole-to-monopole ratio at some range of scales, the predictions for the monopole or quadrupole by themselves do not work as well as for their ratio.

Accuracy in describing the shape of the quadrupole-to-monopole ratio as a function of scale is important since this statistic gives a direct determination of β from clustering in redshift surveys [290,131,132,302]. An alternative to phenomenological models is to obtain the redshift-space power spectrum using approximations to the dynamics, as we now discuss.

¹⁰⁴ Alternatively, if one assumes the *two*-particle velocity distribution is exponential, the suppression factor is the square root of that in Eq. (617), with σ_v the pairwise velocity dispersion along the line of sight, see e.g. [18]. The observational results regarding velocity distributions and their interpretation are briefly discussed in Section 8.3.2.

In the case of the ZA, it is possible to obtain the redshift-space power spectrum as follows [220,642]. In the ZA, the density field obeys

$$1 + \delta(\mathbf{x}) = \int d^3\mathbf{q} \delta_D[\mathbf{x} - \mathbf{q} - \Psi(\mathbf{q})], \quad (618)$$

where $\Psi(\mathbf{q})$ is the displacement vector at Lagrangian position \mathbf{q} . In the plane-parallel approximation, one can treat redshift distortions in the ZA by noting that it corresponds to amplifying the displacement vector by $f \approx \Omega_m^{0.6}$ along the line of sight; that is, the displacement vector in redshift space is $\Psi_s(\mathbf{q}) = \Psi(\mathbf{q}) + \hat{z}f(\Psi(\mathbf{q}) \cdot \hat{z})$. Fourier transforming the corresponding expression to Eq. (618) in redshift space, the power spectrum gives

$$P(\mathbf{k}) = \int d^3\mathbf{q} \exp(i\mathbf{k} \cdot \mathbf{q}) \langle \exp(i\mathbf{k} \cdot \Delta\Psi) \rangle, \quad (619)$$

where $\Delta\Psi = \Psi(\mathbf{q}_1) - \Psi(\mathbf{q}_2)$ and $\mathbf{q} = \mathbf{q}_1 - \mathbf{q}_2$. For Gaussian initial conditions, the ZA displacement is a Gaussian random field, so Eq. (619) can be evaluated in terms of the two-point correlator of $\Psi(\mathbf{q})$. The results of these calculations show that the ZA leads to a reasonable description of the quadrupole-to-monopole ratio [220,642] provided that the zero-crossing scale is fixed to agree with numerical simulations. In general, the ZA predicts a zero-crossing at wavenumbers larger than found in N -body simulations [301]. Furthermore, although the shape of the quadrupole-to-monopole ratio resembles that in the simulations, the monopole and quadrupole do not agree as well as their ratio. This can be improved by using second-order Lagrangian PT [571], but the calculation cannot be done analytically anymore, instead one has to resort to numerical realizations of the redshift-space density field in 2LPT.

7.4.3. The redshift-space bispectrum

Given the second-order PT kernel in redshift space, the leading order (tree level) galaxy bispectrum in redshift-space reads [313,669,562]

$$B_s(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = 2Z_2(\mathbf{k}_1, \mathbf{k}_2)Z_1(\mathbf{k}_1)Z_1(\mathbf{k}_2)P(k_1)P(k_2) + \text{cyc.}, \quad (620)$$

which can be normalized by the power spectrum monopole to give the reduced bispectrum in redshift space, Q_s ,

$$Q_s(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) \equiv \frac{B_s(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)}{a_0^2(P_g(k_1)P_g(k_2) + \text{cyc.})}, \quad (621)$$

where “cyc.” denotes a sum over permutations of $\{k_1, k_2, k_3\}$. Note that Q_s is independent of power spectrum normalization to leading order in PT. Since, to leading order, Q_s is a *function* of triangle configuration which separately depends on Ω_m , b and b_2 , it allows one in principle to break the degeneracy between Ω_m and b present in measurement of the power spectrum multipoles in redshift space [236,313]. The additional dependence of (the monopole of) Q_s on Ω_m brought by redshift-space distortions is small, typically about 10% [313]. On the other hand, as expected, the quadrupole of Q_s shows a strong Ω_m dependence [562].

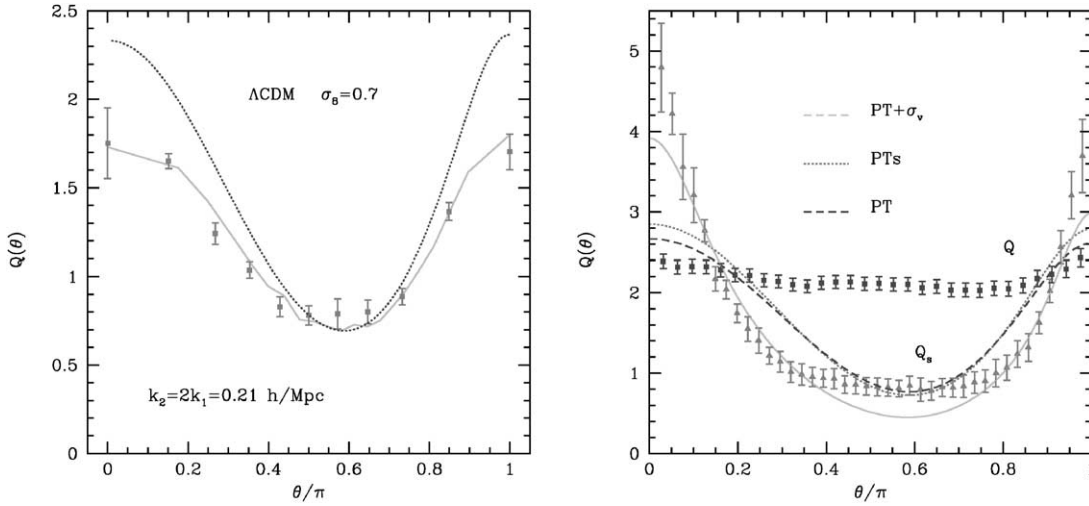


Fig. 48. The left panel shows the bispectrum in redshift space for configurations with $k_2 = 2k_1$ as shown as a function of the angle θ between \mathbf{k}_1 and \mathbf{k}_2 . The dotted line shows the predictions of second-order Eulerian PT, whereas the solid lines correspond to 2LPT. Error bars correspond to the average between four realizations. The right panel shows the bispectrum in redshift space for configurations with $k_2 = 2k_1 = 1.04 h/\text{Mpc}$, i.e. in the non-linear regime. Square symbols denote Q in real space, whereas triangles denote the redshift-space bispectrum. Also shown are the predictions of PT in real space (dashed lines), PT in redshift space (PTs, dotted line) and the phenomenological model with $\sigma_v = 5.5$ ($PT + \sigma_v$, continuous line).

Decomposing into Legendre polynomials, $B_{s \text{ eq}}(\mu) = \sum_{\ell=0}^{\infty} B_{s \text{ eq}}^{(\ell)} P_{\ell}(\mu)$, the redshift-space-reduced bispectrum for equilateral configurations reads [562]

$$Q_{s \text{ eq}}^{(\ell=0)} = \frac{5(2520 + 4410\gamma + 1890\beta + 2940\gamma\beta + 378\beta^2 + 441\gamma\beta^2)}{98b_1(15 + 10\beta + 3\beta^2)^2} + \frac{5(9\beta^3 + 1470b_1\beta + 882b_1\beta^2 - 14b_1\beta^4)}{98b_1(15 + 10\beta + 3\beta^2)^2}, \quad (622)$$

where $\gamma \equiv b_2/b_1$. This result shows that in redshift space, $Q_{s, \text{g}} \neq (Q_s + \gamma)/b_1$ as in Eq. (528), although it is not a bad approximation [562]. In the absence of bias ($b_1 = 1$, $\gamma = 0$), Eq. (622) yields

$$Q_{s \text{ eq}}^{(\ell=0)} = \frac{5(2520 + 3360f + 1260f^2 + 9f^3 - 14f^4)}{98(15 + 10f + 3f^2)^2}, \quad (623)$$

which approaches the real-space result [232] $Q_{\text{eq}} = \frac{4}{7} = 0.57$ in the limit $f \sim \Omega_m^{0.6} \rightarrow 0$. On the other hand, for $f = \Omega_m = 1$, we have $Q_{s \text{ eq}}^{(0)} = 0.464$: for these configurations, the reduced bispectrum is suppressed by redshift distortions.

As discussed before for the power spectrum, leading order calculations in redshift space have a more restricted regime of validity than in real space, due to the rather limited validity of the perturbative expansion for the redshift-space mapping (instead of the perturbative treatment of the dynamics). The same situation holds for the bispectrum. The left panel in Fig. 48 shows the reduced bispectrum Q as a function of angle θ between \mathbf{k}_1 and \mathbf{k}_2 for configurations with $k_2 = 2k_1 = 0.21 h/\text{Mpc}$. The dotted line shows the predictions of tree-level PT in redshift space, Eq. (621), whereas the

symbols correspond to N -body simulations of the Λ CDM model ($\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$, $\sigma_8 = 0.7$) with error bars obtained from four realizations. The disagreement is most serious at colinear configurations. On the other hand, the solid lines obtained using 2LPT [565] agree very well with the N -body measurements. Similarly, good agreement is found for equilateral configurations. The key in the 2LPT predictions is that the redshift-space mapping is done exactly (by displacing the particles from real to redshift space in the numerical realizations of the 2LPT density field), rather than expanded in power series.

At small scales, however, 2LPT breaks down and one must resort to some kind of phenomenological model to account for the redshift distortions induced by the velocity dispersion of clusters. For the bispectrum, this reads [562]

$$B_s(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = \frac{B_s^{\text{PT}}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)}{[1 + \alpha^2[(k_1\mu_1)^2 + (k_2\mu_2)^2 + (k_3\mu_3)^2]\sigma_v^2/2]^2}, \quad (624)$$

where $B_s^{\text{PT}}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)$ is the tree-level redshift-space bispectrum. The assumption is that one can write the triplet velocity dispersion along the line of sight in terms of the pairwise velocity dispersion parameter σ_v and a constant α which reflects the configuration dependence of the triplet velocity dispersion. As noted above, σ_v is determined from simulations solely using the power spectrum ratio; the parameter α is then fitted by comparison with the monopole-to-quadrupole ratio of the equilateral bispectrum measured in the simulations [562]. A somewhat different phenomenological model can be found in [669]; in addition Matsubara [435] studies using a similar phenomenological model the effects of redshift-space distortions in the non-linear regime for the three-point correlation, assuming the validity of the hierarchical model in real space.

The right panel in Fig. 48 shows the redshift-space bispectrum at small scales, to show the effects of non-linear redshift distortions. The square symbols denote Q is real space, which approximately saturates to a constant independent of configuration. On the other hand, the redshift-space Q shows a strong configuration dependence, due to the anisotropy of structures in redshift space caused by cluster velocity dispersion (fingers of God). The phenomenological model (with $\sigma_v = 5.5$ and $\alpha = 3$) in solid lines does quite a good job at describing the shape dependence of Q_s .

Similar studies using numerical simulations have been carried out in terms of the three-point correlation function, rather than the bispectrum, to assess the validity of the hierarchical model in the non-linear regime in redshift space [437,614] and to compare with redshift surveys at small scales [84,267,347]. They obtained analogous results to Fig. 48 for the suppression of Q_s for equilateral configurations compared to Q at small scales due to velocity dispersion. However, studies of the configuration dependence of Q_s in the non-linear regime [437,614,347] find no evidence of the configuration dependence shown in the right panel in Fig. 48. This is surprising, as visual inspection of numerical simulations shows clear signs of fingers of God; this anisotropy should be reflected as a configuration dependence of Q_s . More numerical work is needed to resolve this issue.

7.4.4. The higher-order moments in redshift space

In redshift space, the PT calculation of skewness and higher-order cumulants cannot be done analytically, unlike the case of real space, but can be done by a simple numerical integration for

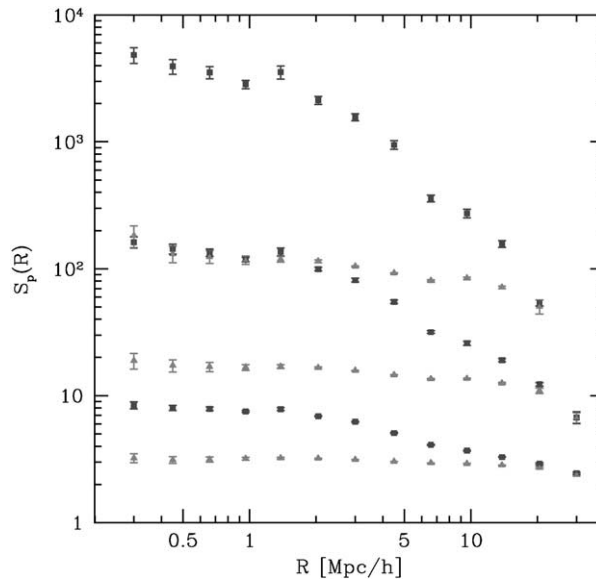


Fig. 49. The S_p parameters for $p=3,4,5$ (from bottom to top) in redshift space for Λ CDM with $\sigma_8=0.9$ as a function of smoothing scale R . Square symbols denote measurements in real-space N -body simulations, whereas triangles correspond to redshift-space values, assuming the plane-parallel approximation.

S_3 [313].¹⁰⁵ The effects of redshift distortions, however, are very small (of order 10%) for the skewness and kurtosis.

On the other hand, at small scales the effect of non-linear redshift distortions is quite strong; since non-linear growth is suppressed in redshift space due to cluster velocity dispersion, the skewness and higher-order moments do not grow much as smaller scales are probed [391,437,614,84,554]. Fig. 49 shows an example for the S_p parameters for top-hat smoothing ($p=3,4,5$) in the Λ CDM model; square symbols denote the real-space values and triangles correspond to redshift-space quantities. Note the close agreement between real and redshift space at the largest scales, and the suppression at small scales for the redshift space case. The latter looks almost scale independent; however, one must keep in mind that correlation functions at small scales should be strongly non-hierarchical, i.e. depend strongly on configuration as shown in the right panel in Fig. 48.

7.4.5. Cosmological distortions

Deep galaxy surveys can probe a large volume down to redshifts where the effects of a cosmological constant, or more generally dark energy, become appreciable. A geometrical effect, as first suggested in [4], arises in galaxy clustering measures because the assumption of an incorrect cosmology leads to an apparent anisotropy of clustering statistics. In particular, structures appear flattened along the line of sight, and thus the power spectrum and correlation functions develop anisotropy, similar to that caused by redshift distortions [18,438,526,181,442]. The challenge to measure this

¹⁰⁵ Using a different approach, Watts and Taylor [682] recently derived a closed form for S_3 in redshift space that does not agree with [313]. This apparent disagreement merits further work.

effect is that redshift distortions are generally larger than cosmological distortions, so a reliable measure of cosmological distortions requires an accurate treatment of redshift distortions.

Recent work along these lines [444], using the approximation of linear PT and that bias is linear, scale and time independent, concludes that the best prospects for measuring cosmological distortions in upcoming surveys is given by the luminous red galaxy (LRG) sample of the SDSS. This sample of about 100 000 galaxies seems to give a good balance between probing structure down to “high” redshift ($z \sim 0.5$) and having a large enough number density so that shot noise is not a limiting factor. Analysis of redshift and cosmological distortions gives a joint $3\text{-}\sigma$ uncertainty on Ω_A and Ω_m of about 15%, assuming $\Omega_A = 0.7$ and $\Omega_m = 0.3$ as the fiducial model. Other surveys, such as the quasar samples in 2dFGRS and SDSS, are predicted to give less stringent constraints due to the sparse sampling [444].

Applications of cosmological distortions to the case of the Lyman- α forest have been proposed in the literature [329,448]. In this case, the distortions are computed by comparing correlations along the line of sight to those by cross-correlating line of sights of nearby quasars. These studies conclude that with only about 25 pairs of quasars at angular separations of $\lesssim 2' - 3'$ it is possible to distinguish an open model from a flat cosmological constant dominated model (with the same $\Omega_m = 0.3$) at the $4\text{-}\sigma$ level. These results, however, assume a linear description of redshift distortions. More recent analysis using numerical simulations [449] suggests that with $13(\theta/1')^2$ pairs at separation less than θ , and including separations $< 10'$, a measurement to 5% can be made if simulations can predict the redshift-space anisotropy with 5% accuracy, or to 10% if the anisotropy must be measured from the data.

Finally, we should mention the effect of clustering evolution along the line of sight, due to observation along the light cone. Estimates of this effect show that for wide surveys such as 2dFGRS and SDSS it amounts to about 10% in the power spectrum and higher-order statistics, while it becomes significantly larger of course for deep surveys, see e.g. [439,283].

8. Results from galaxy surveys

8.1. Galaxies as cosmological tracers

Following the discovery of galaxies as basic objects in our universe [547,322,323], it became clear that their spatial distribution was not uniform but clustered in the sky, e.g. [709]. In fact, the Local Supercluster was recognized early on from 2D maps of the galaxy distribution [184]. The first measurements ever of the angular two-point correlation function $w(\theta)$, done in the Lick survey [653], established already one of the basic results of galaxy clustering that at small scales the angular correlation function $w(\theta)$ has a power-law dependence in θ [see Eq. (625)].

The first systematic study of galaxy clustering was carried out in the 1970s by Peebles and his collaborators. In a truly groundbreaking 12-paper series [500,303,503,501,504,505,275,575,576,226,577,227], galaxies were seen for the first time as tracers of the large-scale mass distribution in the gravitational instability framework.¹⁰⁶ These works confirmed (and extended) the power-law behavior of the angular two-point function, established its scaling with apparent magnitude, and

¹⁰⁶ For an exhaustive review of this and earlier work see [210,508].

measured for the first time the angular power spectrum and the three- and four-point functions which were found to follow the hierarchical scaling $w_N \sim w_2^{N-1}$. The theoretical interpretation of these observations was done in the framework of galaxies that traced the mass distribution in an Einstein de–Sitter universe.¹⁰⁷

These results, however, relied on visual inspections of poorly calibrated photographic plates, i.e. with very crude magnitudes (e.g. Zwicky) or galaxy counts (e.g. Lick) estimated by eye, rather than by some automatic machine. These surveys were the result of adding many different adjacent photographic plates and the uniformity of calibration was a serious issue, since large-scale gradients can be caused by varying exposure time, obscuration by our galaxy and atmospheric extinction. These effects are difficult to disentangle from real clustering, attempts were made to reduce them with smoothing procedures, but this could also result in a removal of real large-scale clustering. More than 20 years after completion of Zwicky and Lick surveys, there were major technological developments in photographic emulsions, computers and automatic scanning machines, such as the Automatic Plate Measuring Machine (APM [374]) and COSMOS [421] micro-densitometers. This allowed a better calibration of wide-field surveys, as measuring machines locate sources on photographic plates and measure brightness, positions and shape parameters for each source [520,519,582,311,384,604,533,574].

In the 1980s large number of redshifts and scanning machines gave rise to a second generation of wide-field surveys, with a much better calibration and a 3D view of the universe.¹⁰⁸ The advent of CCDs revolutionized imaging in astronomy and soon made photographic plate techniques obsolete for large-scale structure studies. Nowadays, photometric surveys are done with large CCD cameras involving millions of pixels and can sample comparable number of galaxies. Furthermore, it is possible with massive multi-fiber or multi-slit spectroscopic techniques to build large redshift surveys of our nearby universe such as the LCRS [584], the 2dFGRS (e.g. see [142]) or the SDSS (e.g. see [699]) as well as of the universe at higher redshifts such as in the VIRMOS (e.g. see [398]) and DEIMOS surveys (e.g. see [177]).

This significant improvement in the quality of surveys and their sampled volume allowed more accurate statistical tests and therefore constrain better theories of large-scale structure. Stringent constraints from upper limits to the CMB anisotropy (e.g. [657]), plus theoretical inputs from the production of light elements (e.g. [696]) and the generation of fluctuations from inflation in the early universe [602,304,280,20] led to the development of CDM models [509,75] where most of the matter in the universe is not in the form of baryons. The 3D mapping of large-scale structures in redshift surveys showed a surprising degree of coherence [378,324,182] which when compared with theoretical predictions of the standard CDM model (e.g. [173]) led to the framework of biased galaxy formation, where galaxies are not faithful tracers of the underlying dark matter distribution (Section 7.1). Subsequent observational challenge from the angular two-point function in the APM survey [422] and counts in cells in the IRAS survey [200,549] led to the demise of standard CDM

¹⁰⁷ In this case self-similarity plus stable clustering leads to hierarchical scaling in the highly non-linear regime, although it does not explain why hierarchical amplitudes are independent of configuration, see Section 4.5. These observations were partially motivated by work on the BBGKY approach to the dynamics of gravitational instability [548] and also generated a significant theoretical activity that led to much of the development of hierarchical models. For a recent historical account of these results and a comparison with current views in the framework of biased galaxy formation in CDM models see [515].

¹⁰⁸ For a review of redshift surveys see e.g. [484,268,607,608].

models in favor of CDM models with more large-scale power, with galaxies still playing the role of (mildly) biased tracers of the mass distribution (e.g. [174]).

The access to the third dimension also allowed analyses of peculiar velocity statistics through redshift distortions [362,290] (Section 7.4, see [295] for a recent review) and measurements of higher-order correlations became more reliable with the hierarchical scaling (Section 4.5.5), $\xi_N \sim Q_N \xi_2^{N-1}$, being established by numerous measurements in 3D catalogs [31,246,92,239]. However, it was not until recently that surveys reached large enough scales to test the weakly non-linear regime and therefore predictions of PT against observations [224,249,248,250,225,567,211]. This is an important step forward, as higher-order statistics encode precious additional information that can be used to break degeneracies present in measurements of two-point statistics, constrain how well galaxies trace the mass distribution, and study the statistics of primordial fluctuations. It is the purpose of the present section to review the observational efforts along these lines.

In this section, we discuss the various results obtained from measurements in galaxy catalogs for traditional statistics such as N -point correlation functions in real and Fourier space and counts-in-cells cumulants (thus leaving out many results on the shape of the CPDF itself, including the void probability function). We do not attempt to provide a comprehensive review of all relevant observations but rather concentrate on a subsample of them. The choice reflects the connections to PT and thus there is a strong emphasis on higher-order statistics. In particular, we do not discuss about cosmic velocity fields, except when redshift distortions are a concern. Also, we do not discuss the spatial distribution of clusters of galaxies since the statistical significance of measurements of higher-order statistics is still somewhat marginal.

The remainder of this section is mainly divided into two large sections, one concerning angular surveys (Section 8.2), the other one concerning redshift surveys (Section 8.3). Finally, Section 8.4 reviews ongoing and future surveys.

8.2. Results from angular galaxy surveys

8.2.1. Angular catalogs

We begin our discussion of angular clustering with a brief description of results from the older generation of catalogs that sets the stage for the more recent results, and then go into a more detailed description of the current state of the subject. Table 14 lists the main angular catalogs that have been extensively analyzed. We show the characteristic parameters of the samples used in the relevant clustering analyses. The information is organized as follows. The second column gives the total area, Ω , of the catalog while the fourth column shows its mean depth, D (associated with the limiting magnitude in the third column). The fifth column gives the volume in terms of a characteristic length, D_E . The sixth column gives the surface density, n_g , which also relates to the mean depth. The three numbers, Ω , D and D_E control volume (area) and edge effects discussed in Section 6. In particular, samples with similar volumes can have quite different sampling biases due to edge effects because of differences in the shape (angular extent) of the survey. The galaxy number density, n_g , relates to discreteness errors (Section 6), which of course are more significant when the total number of objects in the catalog is small. Finally, let us note that some of these catalogs were constructed with different photometric filters (typically blue).

The original *Zwicky catalog* ([710] 1961–1968) contains galaxies to magnitude $m < 15.7$. In the most angular clustering analyses only galaxies brighter than $m = 14.5$ –15 (with ~ 2000 gal/sr) and

Table 14
Angular catalogs

Name	Area Ω (sr)	Magnitudes	Depth D	D_E	No. gal/s	Ref.
Zwicky	1.8	$m_z < 15$	70	73	~ 7000	[710]
Lick	3.3	$m < 19$	220	280	$\sim 10^5$	[581]
Jagellonian	0.01	$m < 21$	400	74	$\sim 10^6$	[543]
ESO/Uppsala	~ 1.8	$d_l > 1'$	60	63	~ 2000	[317]
UGC	~ 1.8	$d_l > 1'$	70	74	~ 2000	[478]
APM	1.3	$b_J = 17-20$	400	380	$\sim 10^6$	[422]
EDSGC	0.3	$b_J = 17-20$	400	230	$\sim 10^6$	[144]
IRAS 1.2Jy	9.5	$f_{60\ \mu\text{m}} > 1.2\text{Jy}$	80	145	480	[218]
DeepRange	0.005	$I_{AB} < 22.5$	2000	150	$\sim 10^8$	[530]
SDSS	$\simeq 3$	$r' < 22$	1000	1300	$\sim 10^7$	[699]

The first five entries correspond to “old” catalogs (1961–1974) based on counts or magnitude/diameters estimates by eye and with poor calibration. The survey area Ω is given in steradians, the depth (mean luminosity distance) and effective size $D_E \equiv (\Omega/4\pi)^{1/3}2D$ are in Mpc/h. The sign \sim reflects the fact that different sub-samples have different values for that quantity.

only in the North galactic cap ($\Omega \sim 1.8$ sr) were used. The mean depth is estimated to be about 50–80 Mpc/h. The base sample used for redshift surveys (see Section 8.3) is a wide survey ($\Omega \simeq 2.7$) with about 20 000 galaxy positions ($m < 15.5$) taken from photographic plates with different calibrations. There have been several studies of systematic errors in Zwicky photometry, showing an important magnitude scale error (see [257] and references therein), however, it is not clear how seriously this affected the clustering properties.

The *Lick catalog* ([581] 1967) consists of 1246 plates of 6×6 square degrees. Counts were done by eye. In the analyses presented by Peebles and collaborators, only 467 plates with $|b_{\text{II}}| > 40^\circ$ were used. These plates have overlapping regions which were used to reduce the counts to a uniform limiting magnitude. Calibration was based on matching the surface density of counts, $\langle n \rangle$, which is much less reliable than calibration based on comparing positions and magnitudes of individual sources. Errors on count estimates were assumed to be independent from cell to cell and to increase the variance by an additive factor proportional to $\langle n \rangle$. In [275], large-scale gradients in the counts were removed by applying a “smoothing factor” which led to some controversy concerning the significance of the analysis [265,183,276,277].

The *Jagellonian field* ([543] 1973) consists of a 6×6 square degrees area with galaxy counts in cells of $3.75' \times 3.75'$, e.g. in a 98×98 grid (higher resolution and deeper than in the Lick catalog). There was no attempt to correct for the lack of uniform optics and plate exposure across the large field of view (e.g. vignetting effects). Although this is quite a deep survey, its angular extent is small and it is clear that it should suffer significantly from the volume and edge effects described in Section 6.

The *ESO/Uppsala* [317] and *UGC* [478] catalogs are based on several hundreds of copies of large (ESO/Palomar) Schmidt plates. Galaxies were found with a limiting visual diameter of about $1'$. There is evidence for the selection function to depend on declination, which has to be taken into account while inverting the angular correlations (e.g. see [345]). Compensation for this effect is likely to produce large-scale artifacts, especially because the sample is relatively small.

The *APM galaxy catalog* ([422] 1990) is based on 185 UK IIIA-J Schmidt photographic plates, each corresponding to 6×6 square degrees on the sky to $b_J = 20.5$ and mean depth of 400 Mpc/h (a factor of two deeper than the Lick catalog) for $b < -40^\circ$ and $\delta < -20^\circ$. These fields were scanned by the APM machine [374]. Galaxy and star magnitudes and positions in the overlapping regions (of 1° /plate) were used to match all plates to a single calibration/exposure. Because there are calibration errors for individual galaxies and positions in a plate, a more careful analysis of vignetting and variable exposure within a plate could be done (as compared to just using the counts). The resulting matching errors can be used to perform a study of the biases induced in the clustering analysis. In the results shown here, an equal-area projection pixel map was used with a resolution of $3.5' \times 3.5'$ cells.

The *EDSGC Survey* ([144] 1992) consists of 60 UK IIIA-J Schmidt photographic plates corresponding to 6×6 square degrees on the sky to $b_J = 20.5$ and mean depth of 400 Mpc/h. In fact, the raw photographic plates are the same in both the APM and EDSGC catalogs, but the later only includes scans of a fraction ($\frac{1}{3}$) of the APM plates, in the central part. The EDSGC database was constructed from COSMOS scans [421], with different calibration and software analyses. Therefore, these two catalogs can be considered as fairly independent realizations of the systematic errors.

The *IRAS 1.2Jy* ([606] 1990) is a redshift subsample of the IRAS Point Source Catalog [123] and is included here because it has also been used to measure angular clustering. This catalog belongs to a newer generation of wide-field surveys, where magnitudes and positions of objects have been obtained by automatic measurements. The CfA [324] and SSRS [168] redshift catalogs have also been used to study angular clustering. More details on redshift samples will be given in Section 8.3.1.

The *Deep Range Catalog* ([530] 1998) consists of 256 overlapping CCD images of 16 arcmin on a side, including 1 arcmin overlap to allow the relative calibration of the entire survey. Images were taken to $I_{AB} < 24$ with a total area extending over a contiguous 4×4 square degrees region. The median redshift for the deeper slices used in the clustering analysis, $I_{AB} = 22\text{--}22.5$ is $z \simeq 0.75$ which corresponds to a depth of approximately $2000h^{-1}$ Mpc. The $I_{AB} = 17\text{--}18$ slice has $z \simeq 0.15$, i.e. a similar depth to the APM catalog. Note the large surface density of this survey. Although this is quite a deep survey its angular extent is rather small and it suffers from the volume and edge effects described in Section 6, especially at the brighter end.

The *Sloan Digital Sky Survey* (SDSS, e.g. see [699]) was still under construction when this review was written and only preliminary results are known at this stage. These results are discussed in a separate section, see Section 8.4 for more details.

Smaller, but otherwise quite similar in design to DeepRange, wide mosaic optical catalogs have been used to study higher-order correlations. For example, the INT-WFC [540] with $\sim 70\,000$ galaxies to $R < 23.5$ over two separated fields of 1.01 and 0.74 square degrees. There are a number of such surveys currently under analyses or in preparation, such as the FIRST radio source survey [33], the NOAO Deep Wide-Field Survey [340], the Canada–France Deep Fields [447], VIRMOS [398], DEIMOS [177] or the NRAO VLA Sky Survey [156].

Most of the catalogs described above have magnitude information, allowing one to study subsets at different limiting magnitudes or depth. This can be used, for instance, to test Limber equation [Eq. (569)] and the homogeneity of the sample [275,422]. Even with the new generation of better calibrated surveys, there has been some concerns about variable sensitivity inside individual plates in the APM and EDSGC catalogs [187] and some questions regarding large-scale gradients in the

APM survey have been raised [221]. Later analysis checked the APM calibration against external CCD measurements over 13 000 galaxies from the Las Campanas Deep Redshift Survey showing an rms error in the range 0.04–0.05 magnitudes [425]. These studies concluded that atmospheric extinction and obscuration by dust in our galaxy have negligible effect on the clustering and also gave convincing evidence for the lack of systematics errors.

8.2.2. The angular correlation function and power spectrum

The angular two-point correlation function in early surveys was estimated from the Zwicky catalog, Jagellonian field and Lick survey in [653,503,505,275,272]. For catalogs with pixel maps (counts in some small cells), such as Lick and Jagellonian, the estimators used were basically factorial moment correlators as described in Section 6.8, whereas for catalogs with individual galaxy positions (such as Zwicky) the estimators were based on pair counts as discussed in Section 6.4.1.

The angular two-point function was found to be consistent between the Zwicky, Lick and Jagellonian samples. For a wide range of angular separations, the estimates were well fitted by a power law:

$$w_2(\theta) \simeq \theta^{1-\gamma}, \quad \gamma \simeq 1.77 \pm 0.04. \quad (625)$$

The resulting 3D two-point function, after using Limber's equation [Eq. (569)] for the deprojection of a power-law model, gives consistent results for all catalogs with

$$\xi_2(r) \simeq \left(\frac{r_0}{r}\right)^\gamma, \quad \gamma \simeq 1.8, \quad r_0 \simeq 5 \pm 0.1 h^{-1} \text{ Mpc} \quad (626)$$

for scales between $0.05 h^{-1} \text{ Mpc} < r < 9 h^{-1} \text{ Mpc}$ [505,275]. On the largest scales, corresponding to $r \gtrsim 10 h^{-1} \text{ Mpc}$, the results are quite uncertain because correlations are small and calibration errors become relevant. The results in [275] suggested a break in $\xi_2(r)$ for $r \gtrsim 10 h^{-1} \text{ Mpc}$. The position of this break, however, depends on the smoothing corrections applied to the Lick catalog (which is the one probing the largest scales) on angles $\theta \gtrsim 3^\circ$ [276,277].

Several other groups have measured small numbers of Schmidt and 4-m plates to produce galaxy surveys of a few hundred square degrees down to $b_J \simeq 20$ and a few square degrees down to $b_J \simeq 23$ [582,311,384,604,533,574]. Most of these studies also show a power-law behavior with consistent values and a sharp break at large scales, the location of the latter depending on the size of the catalog.¹⁰⁹ This sharp break, expected in CDM models, is at least in part caused by finite-volume effects, i.e. the integral constraint discussed in e.g. Section 6.4.2.¹¹⁰ Thus most of these analyses show uncertain estimations for w_2 in the weakly non-linear regime, which is also the case for the ESO/Uppsala and UGC catalogs [345].

¹⁰⁹ More recent studies using CCD cameras find that the power-law form of the small-scale angular correlation function remains in deep samples with amplitude decreasing with fainter magnitudes [161,539,325,530,447], with indications of a less steep power law at the faint end, $I_{AB} \gtrsim 23$ (e.g. [332,530,447]).

¹¹⁰ It is worth pointing out that the cosmic bias caused by the small volume, the boundary or shot noise in the sample typically yields lower amplitudes of w_2 for the smaller (nearby) samples. This has been noticed by several authors (e.g. [175], Fig. 3 in [345]) and sometimes interpreted as a real effect (see also Fig. 8 in [328]).

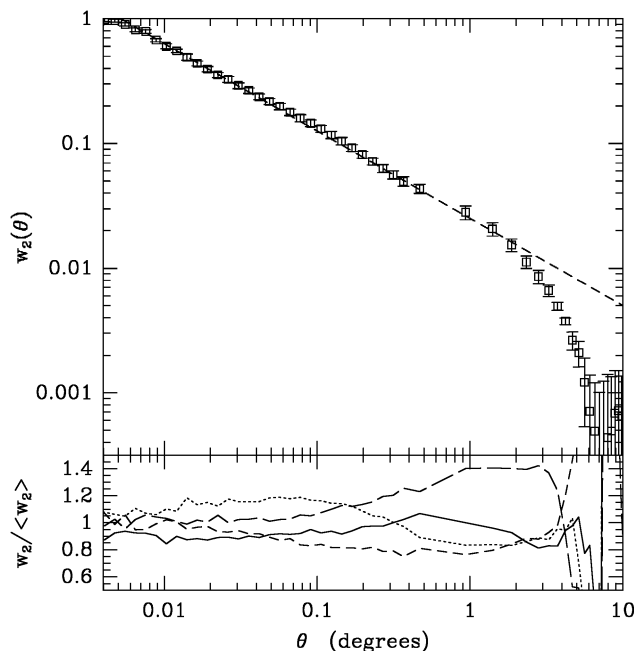


Fig. 50. The two-point angular correlation function $w_2(\theta)$ (squares with error bars), estimated from counts-in-cells and pair counts in the APM map compared with a power law $w_2 \sim \theta^{-0.7}$ (dashed line). Errors are from the dispersion in four disjoint subsamples within the APM. The lower panel shows the ratio of the values in each zone to the average value in the whole sample.

The APM catalog has enough area and depth to probe large scales in the weakly non-linear regime. The first measurements of the angular two-point correlation function [422] led to the discovery of “extra” large-scale power (corresponding to shape parameter¹¹¹ $\Gamma \sim 0.2$), significantly more than in the standard CDM model ($\Gamma = 0.5$). This result has been confirmed by measurement of $w_2(\theta)$ in the EDSGC catalog [144], and subsequent analyses of the inferred 3D power spectrum from inversion of the APM angular correlation function [26] and angular power spectrum [27] and inversion from $w_2(\theta)$ to the 3D two-point function [29] (see Section 8.2.3 for a brief discussion of inversion procedures). Both APM and EDSGC find more power than the Lick catalog on scales $\theta \gtrsim 2^\circ$, suggesting that the Lick data were overcorrected for possible large-scale gradients [422–425].

Fig. 50 shows the two-point angular correlation function $w_2(\theta)$ estimated for $\theta > 1^\circ$ from counts in the pixel maps (i.e. the factorial moment correlator W_{11} , see Section 6.8) and at smaller scales from galaxy pair counts (using the $DD/DR - 1$ estimator, see Section 6.4.1). A fit of the two-point angular correlation with a power law $w_2 \simeq A\theta^{1-\gamma}$, for scales $\theta < 2^\circ$ gives $A \simeq 2.7 \times 10^{-2}$ and $\gamma \simeq 1.7$ (shown as a dashed line). After inverting the Limber equation, the corresponding 3D two-point correlation function is in good agreement with Eq. (626), with a slightly flatter slope $\gamma \simeq 1.7$. The uncertainty in the value of the correlation length r_0 is controlled mainly by the accuracy in the knowledge of the selection function in Eq. (569) and by the cosmic errors that we discuss below.

¹¹¹ The shape parameter when the contribution of baryons is neglected ($\Omega_b \ll \Omega_m$) reads $\Gamma \approx \Omega_m h$, see e.g. [201,76,21]. However, for currently favored cosmological parameters it is more accurate to use $\Gamma = \exp[-\Omega_b(1 + \sqrt{2h/\Omega_m})]\Omega_m h$ [611].

The APM data show good match between several disjoint magnitude slices when scaled according to the Limber equation to the same depth (see Fig. 25 in [425]). The agreement is good up to very large scales $\theta D \gtrsim 40h^{-1}$ Mpc; this indicates that the APM catalog can be used to explore the weakly non-linear regime. Similar conclusions apply to the EDSGC catalog (see [144]), which is compared in terms of $w_2(\theta)$ to APM in [425] (see also [328]): both catalogs agree well for $0.1^\circ < \theta < 0.5^\circ$. At larger angular scales, the EDSGC results differ from APM, essentially because of finite volume and edge effects due to its smaller area. More worrisome is that at smaller scales, $\theta < 0.1^\circ$ there are also discrepancies (presumably related to deblending of galaxies in high-density regions, see [627]) which can be quite significant for higher-order moments as we shall discuss in Section 8.2.5.

The errors shown in Fig. 50 are obtained from the scatter among four disjoint subsamples in the APM, which is often an overestimate of the true cosmic errors at large scales (see end of Section 6.4.3). However, as discussed at length in Section 6, error bars give only a partial view of the real uncertainties (especially in the case of spatial statistics), since measurements at different scales are strongly correlated. This is illustrated in the bottom panel of Fig. 50, where the variations of the measured w_2 from subsample to subsample are coherent (and quite significant at the largest scales where edge effects become important). As a result, the values of w_2 change mostly in amplitude and to a lesser extent in slope from zone to zone. These cross-correlations are not negligible and must be taken into account to properly infer cosmological information since the measurements at different scales are not statistically independent. Only very recently the effect of the covariance between estimates at different scales was included in the analyses of APM [204,203] and EDSGC [331] angular clustering, by focusing on large scales and using the Gaussian approximation to the covariance matrix, similar to Eq. (403). We discuss these results in the next section.

Finally, note that the nearly perfect power-law behavior of the angular correlation function imposes non-trivial constraints on models of galaxy clustering. Since in CDM models the dark matter two-point correlation function is not a power law, this implies that the bias between the galaxy and mass distribution must be scale dependent in a non-trivial way. The current view (see discussion in Section 7.1.4) is that this happens because the number of galaxies available in a given dark matter halo scales with the mass of the halo as a power law with index smaller than unity. In these scenarios, the fact that the galaxy two-point function follows a power law is thus a coincidence. Given the accuracy of the power law behavior (see Fig. 50) this situation is certainly puzzling. It seems unlikely that such a cancellation can take place to such an accuracy.¹¹² On the other hand, these models predict at small scales that galaxy velocity dispersions and S_p parameters are significantly smaller than for the dark matter, as observed. We shall come back to discuss this in more detail below.

8.2.3. Inversion from angular to 3D clustering

The cosmological information contained in the angular correlation function of galaxies can be extracted in basically two different ways. One is to just project theoretical predictions and compare to observations in angular space. It is also useful to carry out the alternative route of an inversion procedure from Eq. (570) to recover the 3D power spectrum, and compare to theoretical predictions in the more familiar 3D space. This has the advantage that it is possible to carry out parameter

¹¹² However, one must keep in mind that features in the spatial correlation function can be significantly washed out due to projection, as first emphasized in [209].

estimation on the scales not affected by non-linear evolution.¹¹³ To successfully apply this method, however, one must be able to propagate uncertainties from angular space to 3D space in a reliable way. Recent work has developed techniques that make this possible.

To go from the angular correlation function to the 3D power spectrum (or two-point function) requires the inversion of an integral equation with a nearly singular kernel, since undoing the projection is unstable to features in the 3D correlations that get smoothed out due to projection. The inverse relation between $\xi_2(r)$ and $w_2(\theta)$ can be written down formally using Mellin transforms [209,490]; however, in practice this result is difficult to implement since it involves differentiation of noisy quantities. Most inversions from $w(\theta)$ [26,29,253] and the angular power spectrum [27] in the APM survey used an iterative deconvolution procedure suggested by Lucy [413] to solve integral equations. However, although Lucy's method can provide a stable inversion, it does not provide a covariance matrix of the recovered 3D power spectrum. Error bars on the reconstructed 3D power spectrum have been estimated by computing the scatter in the spectra recovered from four different zones of the APM survey [26,27]; this can only be considered as a crude estimate and cannot be used to constrain cosmological parameters in terms of rigorous confidence intervals.

A number of methods have emerged in the last couple of years to overcome these limitations. These techniques involve some way of constraining the smoothness of the 3D power spectrum to suppress features in it that lead to minimal effects on the angular clustering and thus make the inversion process unstable. A method using a Bayesian prior on the smoothness of the 3D power spectrum was proposed in [188]. An improved method, based on SVD decomposition [204], identifies and discards those modes that lead to instability. Both methods give the covariance matrix for the estimates of the 3D power spectrum given a covariance matrix of the angular correlations, which can be done beyond the Gaussian approximation. The resulting 3D covariance matrix shows significant anti-correlations between neighboring bins [188,204]; this is expected since oscillatory features in the power spectrum are washed out by projection and thus are not well constrained from angular clustering data. Another technique based on maximum likelihood methods for performing the inversion is presented in [203] (see e.g. discussion in Section 6.11). This has the advantage of being optimal for Gaussian fluctuations; on the other hand, the assumption of Gaussianity means that errors and their covariances are underestimated at scales affected by non-linear evolution where non-Gaussianity becomes important. Including the covariance matrix of angular correlations showed that constraints on the recovered large-scale 3D power spectrum of APM galaxies become less stringent by a factor of two [204,203] compared to some of the previous analyses that assumed a diagonal covariance matrix.

Fig. 51 displays the APM 3D power spectrum $P(k)$ reconstructed from the angular two-point correlation function [26,253] inverting Limber's equation (570) using Lucy's method. The error bars are obtained from the dispersion on $w_2(\theta)$ over four zones as shown in Fig. 50 and should thus be considered as a crude estimate, especially at large scales (see [203] for comparison of errors in different inversion methods). The solid curve corresponds to a reconstruction of the linear part of

¹¹³ In angular space, this distinction is harder to make due to projection, particularly for the two-point correlation function. For example, for APM, $w(\theta)$ at $\theta = 1^\circ, 2^\circ, 3^\circ, 5^\circ$ has contributions from 3D Fourier modes up to $k = 1, 0.4, 0.3, 0.2h/\text{Mpc}$, respectively [188].

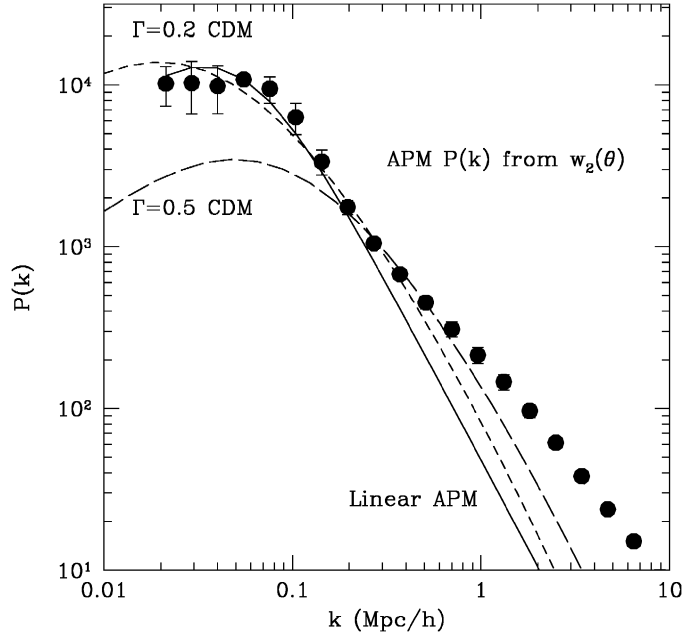


Fig. 51. The APM 3D power spectrum reconstructed from $w_2(\theta)$. The continuous line shows a linear $P(k)$ reconstruction. The short- and long-dashed lines show linear CDM models with $\Gamma=0.2$ and 0.5 , normalized to the data at $k \simeq 0.3h/\text{Mpc}$.

the spectrum, which can be fitted by

$$P_{\text{linear}}^{\text{APM}}(k) \simeq 7 \times 10^5 \frac{k}{[1 + (k/0.05)^2]^{1.6}} \quad (627)$$

for $k < 0.6h/\text{Mpc}$, and $\Omega_m = 1$ [30]. This linearization has been obtained assuming no bias between APM galaxies and dark matter,¹¹⁴ following the linearization first done in [289] and extended in [493] based on the mapping from the linear to non-linear power spectrum (see e.g. Section 4.5.4 for a discussion). Eq. (627) has been obtained by running N -body simulations and agrees well with the mapping prescription of [335]. Note how non-linear effects become important at $k > 0.1h/\text{Mpc}$.¹¹⁵

As can be seen from Fig. 51, a comparison to CDM models on linear scales ($k < 0.3h/\text{Mpc}$) favors low values of power spectrum shape parameter Γ , showing more power on these scales than the standard CDM model with $\Gamma = 0.5$. Indeed, the most recent analysis including the effects of the covariance matrix discussed above concludes using the deprojected data for $k \leq 0.19h/\text{Mpc}$ that $0.05 \leq \Gamma \leq 0.38$ to 95% confidence [203].¹¹⁶ Similar results have been obtained from a similar

¹¹⁴ Unfortunately, as shown in [30], this assumption is inconsistent at small scales: the higher-order moments predicted by evolving the linear spectrum in Eq. (627) are in strong disagreement with the APM measurements at scales $R \lesssim 10 \text{ Mpc}/h$ (see Fig. 54), indicating that galaxy biasing is operating at non-linear scales. On the other hand, the large-scale correlations ($R > 10 \text{ Mpc}/h$) are consistent with no significant biasing, see Section 8.2.6.

¹¹⁵ In fact, it has been demonstrated in [222] that the one-loop PT predictions presented in Section 4.2.2 work very well for this spectrum on scales where the fit in Eq. (627) is valid, $k < 0.6h/\text{Mpc}$.

¹¹⁶ In addition, it was shown that galactic extinction, as traced by the maps in [555], had little effect on the power spectrum over the APM area with $\delta < -20^\circ$.

recent likelihood analysis of the EDSGC survey angular power spectrum [331]. Fig. 51 suggests that on very large scales ($k < 0.05h/\text{Mpc}$), the APM data might show an indication of a break in the power spectrum [253]. From the figure, it might seem as if this is a $3\text{-}\sigma$ detection, but as mentioned above different points are not independent. Analytical studies, using different approximations to account for the covariance matrix between different band powers, indicate that this might be only a $1\text{-}\sigma$ result [204,203].¹¹⁷

The above results on the shape parameter of the power spectrum have been confirmed by analyses of redshift catalogues as will be discussed in Section 8.3.2, and will soon be refined by measurements in large ongoing surveys such as the 2dFGRS or the SDSS (Section 8.4).

On smaller scales, a detailed study [260] of the reconstructed 3D two-point correlation function in the APM [29] shows an inflection point in the shape of $\xi_2(r)$ at the transition to the non-linear scale $r \simeq r_0 \sim 5 \text{ Mpc}/h$, very much as expected from gravitational instability (see Section 4.5.2).

8.2.4. Three-point statistics and higher order

Angular surveys provide at present the best observational constraints on higher-order correlation functions in the non-linear regime. In most cases, however, a detailed exploration of the different configurations available in three-point and higher-order correlations has not been given, due to limitations in signal to noise.¹¹⁸ This will have to await the next generation of photometric surveys (e.g. SDSS [699] and DPOSS [187]).

Table 15 summarizes the measurements achieved in various surveys. As can be seen in the third column of Table 15, the limited size of surveys means that most of the measurements only probed the non-linear regime, except those done in the IRAS and APM catalogs. The first measurements of the three-point angular correlation function w_3 in the Jagellonian field [275], Lick and Zwicky surveys [275] established that at small scales the *hierarchical model* (see Section 4.5.5) gives a good description of the data,

$$w_3(\theta_1, \theta_2, \theta_3) = q_3 [w_2(\theta_1)w_2(\theta_2) + w_2(\theta_2)w_2(\theta_3) + w_2(\theta_3)w_2(\theta_1)] , \quad (628)$$

where q_3 is a constant of order unity with little dependence on scale or configuration (within the large error bars) at the range of scales probed. In addition, the four-point function was found to be consistent in the Lick and Zwicky catalogs with the hierarchical relation,

$$w_4(1, 2, 3, 4) = r_a [w_2(1, 2)w_2(2, 3)w_2(3, 4) + \text{cyc. (12 terms)}] \\ + r_b [w_2(1, 2)w_2(1, 3)w_2(1, 4) + \text{cyc. (4 terms)}] , \quad (629)$$

where $w_2(i, j) \equiv w_2(\theta_{ij})$ with θ_{ij} being the angular separation between points i and j . The amplitudes r_a and r_b correspond to the different topologies of the two types of tree diagrams connecting the four points (see e.g. Fig. 6 and discussion in Section 4.5.5), the so-called snake (r_a , first diagram in Fig. 6) and star diagrams (r_b , second diagram in Fig. 6). The overall amplitude of the four-point

¹¹⁷ However, the initial suggestion by [253] for a break in the APM was confirmed with realistic numerical simulations which show that a mock galaxy catalog as big as the APM can be used to recover such a break when placed at different scales (see Figs. 11 and 12 in [253]). The level of significance for this detection was not studied, so these apparently discrepant analyses require further investigation.

¹¹⁸ In addition, even with the currently available computational power and fast algorithms relying on e.g. KD-tree techniques [461], measuring directly higher-order correlation functions can be very computationally intensive.

Table 15

The angular three- and four-point amplitudes $3q_3$ and $16q_4 \equiv 12r_a + 4r_b$, at physical scales (in Mpc/h) specified in the third column by $\mathcal{D}\theta$. The last five entries correspond to the newer generation of galaxy catalogs (see Table 14). Error bars should be considered only as rough estimates, see text for discussion

$3q_3$	$16q_4$	$\mathcal{D}\theta$	Sample	Year	Ref.	Estimator
1.9 ± 0.3	—	0.4–1.2	Jagellonian	1975	[505]	Cumulant corr.
3.5 ± 0.4	—	0.1–4	Zwicky (-Coma)	1975	[504]	Multiplet counts
5.3 ± 0.9	—	0.1–4	Zwicky	1977	[275]	Multiplet counts
—	100 ± 18	0.1–2	Zwicky	1978	[226]	Multiplet counts
4.7 ± 0.7	—	0.3–10	Lick	1977	[275]	Cumulant corr.
—	77 ± 7	0.5–4	Lick	1978	[226]	Cumulant corr.
4.8 ± 0.1	40 ± 3	0.3–5	Lick	1992	[618]	Cumulant corr.
$\simeq 3$	—	0.3–5 (k)	Lick	1982	[229]	Bispectrum
2.7 ± 0.1	—	0.2–2	ESO-Uppsala	1991	[345]	Multiplet counts
5.4 ± 0.1	—	0.2–2	UGC	1991	[345]	Multiplet counts
3.8 ± 0.3	35 ± 10	4–20	IRAS 1.2Jy	1992	[451]	Cumulant corr.
3.5 ± 0.1	31 ± 1	0.5–50	APM (17–20)	1995	[620]	Cumulant corr.
3.9 ± 0.6	—	4	APM	1999	[225]	Cumulant corr.
2–6	—	4–30	APM	1999	[225]	Cumulant corr.
1.5–3	—	0.2–3	LCRS	1998	[347]	Multiplet counts
8–3	—	0.5–3	DeepRange	2000	[635]	Multiplet counts
2–1	—	3–6	DeepRange	2000	[635]	Multiplet counts
5–1	—	0.5–20	SDSS	2001	[261,262]	Multiplet counts

Table 16

The reduced skewness and kurtosis from counts-in-cells gin angular space. In most cases, only the mean values over a range of scales were published. In cases where measurements of the individual s_p for each smoothing scale are reported in the literature, we quote the actual range and the corresponding range of scales. Error bars should be considered only as rough estimates, see text for discussion

s_3	s_4	$\mathcal{D}\theta$	Sample	Year	Ref.
2.9 ± 0.9	12 ± 4	1–8	Zwicky	1984	[583]
2.4 ± 0.4	9.5 ± 2.4	2–20	CfA	1994	[224]
2.2 ± 0.3	8 ± 3	2–20	SSRS	1994	[224]
2.5 ± 0.4	11 ± 3	2–20	IRAS 1.9Jy	1994	[224]
3.8 ± 0.1	33 ± 4	7–30	APM (17–20)	1994	[249]
5.0 ± 0.1	59 ± 3	0.3–2	APM (17–20)	1994	[249]
7–4	170–40	0.1–14	EDSGC	1996	[622]
3.0 ± 0.3	20 ± 5	0.1	APM (17–20)	1998	[627]
6–2	120–10	0.1–6	DeepRange	2000	[635]
5–2	100–20	0.5–20	SDSS	2001	[261,262,637]

function is thus $16q_4 \equiv 12r_a + 4r_b$, which we quote in Table 15, together with the three-point amplitude $3q_3$. These are useful to compare with the angular skewness and kurtosis in Table 16 discussed in Section 8.2.5 because in the hierarchical model $s_N \simeq N^{N-2}q_N$ to very good accuracy.¹¹⁹ In addition,

¹¹⁹ And similarly in the 3D case, see [85,249] for accurate estimates of the small corrections to this relation.

as discussed in Section 7.2.3, the q_N coefficients are very weakly dependent on details of the survey such as the selection function and its uncertainties, so it is meaningful to compare q_N from different galaxy surveys.

In the first and second columns of Table 15, in addition to the numerical values of q_3 and q_4 , we quote as well the error on the estimate calculated by the authors. Except when noted otherwise, error bars were obtained from the dispersion in different zones of the catalog. Since typically the number of zones used is small (about four in most cases), the estimated errors are very uncertain.¹²⁰ In addition, this method obviously cannot estimate the cosmic variance, which can be a substantial contribution for surveys with small area. Many, if not most, of the differences between the various numerical values given in Table 15 can be explained by statistical fluctuations and systematics due to the finiteness of the catalogs [328] (see Section 6 for a detailed discussion of these issues), as we now briefly discuss.

The results of q_3 in the Zwicky catalog do not seem to be very reliable since the value found in [504] changed by more than 50% due to the omission of only 14 galaxies in the Coma cluster (see [275]). Similar effects have been found in other samples (e.g. ESO-Uppsala [345]). This sensitivity reflects that the sample is not large enough to provide a fair estimate of higher-order statistics. Likewise, the rather low value for q_3 found in the Jagellonian sample is likely strongly affected by finite-volume effects due to the small area covered. Similarly, the values obtained from the projected LCRS in [347] could be partially contaminated by edge effects due to the particular geometry of the catalog (6 strips of $1.5^\circ \times 80^\circ$) and perhaps also by sampling biases due to inhomogeneous sampling around high-density regions.¹²¹

Work has been done as well to study the dependence of q_3 on morphological type, but dividing the data into smaller subsamples tends to produce stronger statistical biases. In the ESO-Uppsala and UGC catalogs, Jing et al. [345] found that spirals have significantly smaller values of q_3 . This could be interpreted through the well-known density–morphology relation [192,529]: spirals avoid rich clusters and groups, an effect that could be more important at smaller scales (this is illustrated to some extent in Fig. 45). The results for the full sample in the ESO-Uppsala and UGC catalogs showed good agreement with the hierarchical scaling (note however that error bars quoted in this case are just due to the dispersion in the fit to the hierarchical model rather than reflecting sample variance).

The measurements of the three-point correlation function in the Lick survey did not show any strong evidence for a dependence of q_3 on the shape of the triangle, although a marginal trend was found that colinear triangles had a higher q_3 than isosceles [275]. The three-point statistics was analyzed in terms of the bispectrum by Fry and Seldner [229], who found the same amplitude for q_3 as in real-space, but some indications of a scale dependence beyond the hierarchical scaling, with q_3 increasing as a function of wavenumber k with a peak corresponding to the angular scale (2.5°) of the break in $w_2(\theta)$, and then decreasing again at large k . A later re-analysis of the large-scale Lick bispectrum [236] showed a marginal indication of dependence on configuration shape, too small compared to the one expected in tree-level PT, and thus in principle an indication of a large galaxy

¹²⁰ However, as discussed in the end of Section 6.4.3 for the two-point correlation function, when the number of subsamples is large, this method tends to overestimate the real cosmic errors.

¹²¹ Due to the fixed number of fibers per field and “fiber collisions”. Using random catalog generation Jing and Borner [347] checked that these effects appeared to be insignificant.

bias [see Eq. (528)]. However, the scales involved were not safely into the weakly non-linear regime and thus this result is likely the effect of non-linear evolution rather than a large galaxy bias [560].

The four-point function measurements in the Lick survey were not able to test the relation in Eq. (629) in much detail, but assuming Eq. (629) measurements for some specific configurations (such as squares and lines) gave a constraint on the amplitudes r_a and r_b which were then translated into a constraint on 3D amplitudes by deprojection (see Section 7.2.3), resulting in $R_a = 2.5 \pm 0.6$ and $R_b = 4.3 \pm 1.2$ [226]. These results on the Lick survey were considerably extended in [618] to higher-order q_N 's up to $N = 8$ in the context of the *degenerate* hierarchical model,¹²² by using two-point moment correlators.¹²³ This confirmed the hierarchical scaling $w_N \sim q_N w_2^{N-1}$ up to $N = 8$, at least for these configurations, with $q_N \approx 1-2$.¹²⁴

The same technique was applied to the IRAS 1.2Jy survey in [451], verifying the hierarchical scaling for $N = 3, 4$ but with q_N 's with $N > 4$ being consistent with zero, and also to the APM survey [620] which showed non-zero amplitudes up to $N = 6$, with a trend of increasing q_N as a function of N , i.e. $q_N = 1.2, 2, 5.3, 10$ for $N = 3, 4, 5, 6$, unlike the case of the Lick catalog. The APM survey was later re-analyzed in terms of cumulant correlators [e.g. see Eq. (348)] in [623], showing hierarchical scaling for $N = 4, 5$ to within a factor of two.¹²⁵ In addition, it showed that at scales $\theta \gtrsim 3.5^\circ$ the factorization property predicted by PT, Eq. (349), starts to hold. By measuring $\langle \delta_1^3 \delta_2 \rangle_c$ and $\langle \delta_1^2 \delta_2^2 \rangle_c$ and assuming the hierarchical model as in Eq. (629) it was possible to constrain (after deprojection) $R_a \simeq 0.8$ and $R_b \simeq 3.7$, in reasonable agreement with the Lick results [226] mentioned in the previous paragraph. These imply an average $q_4 \simeq 2.2$.

The analysis of the three-point function in the DeepRange survey [635] shows a general agreement with the hierarchical model with large errors in q_3 , with a consistent decrease as a function of depth. Indeed, a fit to the hierarchical model, Eq. (628), gives $q_3 = 1.76, 1.39, 2.80, 1.00, 0.34, 0.57$ for I -band magnitudes $I = 17-18, 18-19, 19-20, 20-21, 21-22, 22-22.5$, respectively. This trend is also present in the count-in-cells measurements and, if confirmed in other surveys, has interesting implications for the evolution of galaxy bias (see Fig. 55). Note that in this work errors were estimated using the FORCE code [621,152,630], which is based on the full theory of cosmic errors as described in Section 6.

Some of the analyses above probed the weakly non-linear regime, where the q_N 's are expected to show a characteristic angular dependence predicted by PT, even after projection from 3D to angular space [242,225,101]. Measurements of q_3 in the Lick catalog showed a marginal indication that colinear configurations are preferred compared to isosceles triangles [275,236] (but see [229]).

¹²² In this case all amplitudes corresponding to different tree topologies are assumed to have the same amplitude q_N , and thus $R_a = R_b$, etc., see Section 4.5.5.

¹²³ In the same spirit, it is worth noticing that four-point correlation function estimates for particular configurations can be obtained through measurements of the dispersion of the two-point correlation function over subsamples (or cells) extracted from the catalog (see [83,230]): this is a natural consequence of the theory of cosmic errors on w_2 detailed in Section 6.4.3. This method has the potential defect of being sensitive to possible artificial large-scale gradients in the catalog.

¹²⁴ Note, however, that the errors quoted by the authors come from a fitting procedure, not sampling variance. For $N > 6$, correlations are consistent with zero when using the sampling variance among 12 zones.

¹²⁵ The scales probed in this case, $0.8^\circ < \theta < 4.5^\circ$, are in the transition to the non-linear regime (1° corresponds to about 7 Mpc/h at the APM depth), so it is not expected to show hierarchical scaling. On the other hand, galaxy biasing might help make correlations look more hierarchical, as illustrated in Fig. 45 by the suppression in the growth of S_p parameters as small scales are probed.

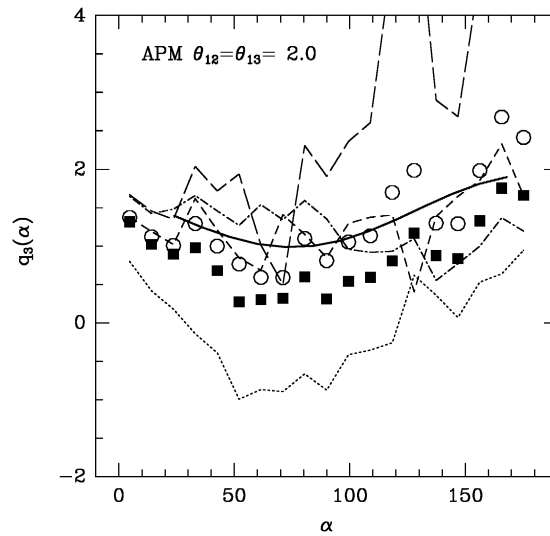


Fig. 52. The angular three-point amplitude $q_3(\alpha)$ from PT prediction (thick continuous line) compared with the APM measurements at $\theta_{12}=\theta_{13}=2^\circ$: closed squares and open circles correspond to the full APM map and to the mean of four disjoint zones. Other curves show results for each of the zones (from [225]).

Projecting the three-point function in redshift space from the LCRS survey, Jing and Borner [347] found a marginal enhancement for colinear triangles, but the scales probed ($r \lesssim 12 \text{ Mpc}/h$) are not safely in the weakly non-linear regime.

For angular catalogs, the APM survey presents the best available sample to check the angular dependence of q_3 predicted by PT [225]. Figs. 52 and 53 show the measurements of $q_3(\alpha)$ in the APM survey at $\theta_{12}=\theta_{13}=0.5\text{--}4.5^\circ$ estimated by counting pairs and triplets of cells of a given angular configuration, see Section 6.8. Closed squares correspond to estimations in the full APM map, while open circles are the mean of q_3 estimated in four disjoint zones. The value of $3q_3 \simeq 3.9 \pm 0.6$ at $\alpha \simeq 0$, shown in Table 15, is in agreement with the cumulant correlators estimated (with 4×4 bigger pixels) in [628]. Furthermore, the average over α is comparable to the values of $s_3/3$ in Table 15 and in particular to the APM and EDSGC estimations [249,451,627].

Fig. 52 shows the results for individual zones in the APM (same as the ones in Fig. 50) for all triangles with $\theta_{12}=\theta_{13}=2^\circ$. These estimations of q_3 are subject to larger finite-volume effects, because each zone is only $\frac{1}{4}$ the size of the full APM.¹²⁶ As in Fig. 50, there is a strong covariance among the estimations in different zones, which results in a large uncertainty for the overall amplitude q_3 . Because the zones cover a range of galactic latitude, the number of systematic errors in the APM catalog (star–galaxy separation, obscuration by the galaxy, plate matching errors) might be expected to vary from zone to zone. No evidence for such systematic variation is found in q_3 : the scatter in individual zone values is compatible with the sampling variance observed in N -body simulations [225]. On larger scales, $\theta \gtrsim 3^\circ$, the individual zone amplitudes exhibit large variance, and in addition

¹²⁶ The fact that the average over the four zones (open circles) is not equal to the measurement in the full APM map is a manifestation of estimator bias [328,630].

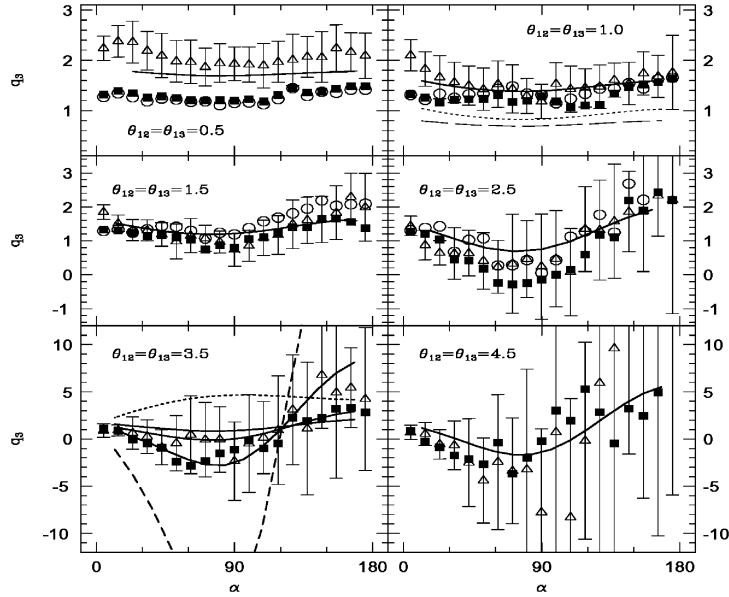


Fig. 53. The projected three-point amplitude q_3 in PT (solid curves) and N -body results (open triangles with error bars) for the APM-like power spectrum are compared with q_3 measured in the APM survey (closed squares and open circles, with same meanings as in Fig. 1). Each panel shows the amplitude at different $\theta_{12} = \theta_{13}$. In upper right panel, dotted and dashed curves correspond to PT predictions with $b_1 = 1$, $b_2 = -0.5$ and $b_1 = 2$, $b_2 = 0$, respectively. In the lower left panel, upper and lower solid curves conservatively bracket the uncertainties in the inferred APM-like power spectrum, long-dashed curve corresponds to SCDM, and the dotted curve shows the leading order prediction for the χ^2 non-Gaussian model.

boundary effects come into play. As seen in Fig. 53 at these scales q_3 is consistent with zero within the errors.

The APM results are compared with the values of q_3 predicted by PT with the linear APM-like spectrum in Eq. (627) (solid curves) and with measurements in N -body simulations (open triangles with error bars) with Gaussian initial conditions corresponding to the same initial spectrum. Since the APM-like model has, by construction, the same $w(\theta)$ as the real APM map, it is assumed that the sampling errors are similar in the APM and in the simulations. This might not be true on the largest scales, where systematics in both the APM survey and the simulations (periodic boundaries) are more important.

At scales $\theta \gtrsim 1^\circ$, the agreement between the APM-like model and the APM survey is quite good; this corresponds roughly to physical scales $r \gtrsim 7h^{-1}$ Mpc, not far from the non-linear scale ($r_0 \simeq 5$, where $\xi_2 \simeq 1$). Also note that the q_3 predicted in the SCDM model (dashed curve in lower-left panel of Fig. 53) clearly disagrees with the APM data; this conclusion is independent of the power spectrum normalization and it is therefore complementary to the evidence presented by two-point statistics [422,200] (see discussion in Sections 8.2.2 and 8.2.3). At smaller angles, $\theta \lesssim 1^\circ$, q_3 in the simulations is larger than in either the real APM or PT (top-left panel in Fig. 53). The discrepancy between simulations and PT on these relatively small scales is due to non-linear evolution. The reason for the discrepancy with the real APM is probably an indication of galaxy biasing at small scales: this will affect the inference of the linear power spectrum from the data [30]

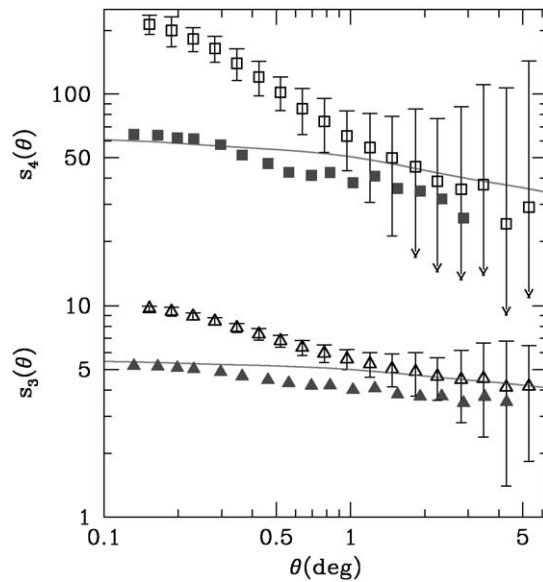


Fig. 54. The angular skewness, s_3 , and kurtosis, s_4 , from the APM catalog (filled triangles and squares) as compared with PT results (continuous line) and APM-like all-sky N -body simulations (open triangles and squares).

and also suppress higher-order correlations compared to the dark matter [570] (see e.g. Fig. 45 and discussion in Section 7.1.3).

8.2.5. Skewness, kurtosis and higher-order cumulants

Table 16 shows the results for the skewness (s_3) and kurtosis (s_4) in several of the angular catalogs described in Section 8.2.1.

The analysis of the Zwicky sample by Sharp et al. [583] used moments of counts in cells to estimate the hierarchical amplitudes q_N , assuming the degenerate hierarchical model in Section 4.5.5. Because counts in cells were used, the measurement is closer to s_3 than to q_3 . As noted in Section 8.2.4, the Zwicky catalog has been shown to be sensitive to a few galaxies in the Coma cluster, a signature that the survey is not large enough to be a fair sample for the estimation of higher-order moments. Indeed, in [583] it was found that the mean over a four-subsample split changed from the values in Table 16 to $s_3 = 4.2 \pm 0.9$ and $s_4 = -7 \pm 12$, a manifestation of the estimation biases discussed in Section 6.

In [224] angular positions from volume-limited subsamples of redshift catalogs (CfA, SSRS and IRAS 1.9Jy) were used to estimate the angular moments.¹²⁷ Note, for example, how the values for s_3 and s_4 in the CfA survey from these smaller samples are lower than in the parent Zwicky sample. This suggests again that there are significant systematic finite-volume effects [249,621,328,630].

Fig. 54 shows s_3 (filled triangles) and s_4 (filled squares) measured in the APM survey [249]. The open figures with error bars correspond to the mean of 20 N -body all-sky simulations presented in [254] with the linear “APM-like” power spectrum in Eq. (627), with $1\text{-}\sigma$ error bars scaled to the size

¹²⁷ The values in Table 16, from Table 8 in [224], have been multiplied by $r_3 \simeq 1.2$ and $r_4 \simeq 1.5$ for a direct comparison in angular space.

of the APM.¹²⁸ The continuous line shows the tree-level PT results of [48] numerically integrated for the APM-like power spectrum, as described in [254].¹²⁹ The uncertainties in the shape of the power spectrum and the evolution of clustering of APM galaxies are comparable or smaller than the simulation error bars [250].

As can be seen in Fig. 54, APM measurements are somewhat below the PT predictions or N -body results at $\theta \gtrsim 1^\circ$ [48], indicating possibly a slight bias for APM galaxies. But note that this difference is not very significant given the errors and the fact that there is a strong covariance and a significant negative bias on these scales (see Section 4.1 in [254]). At smaller angles, $\theta \lesssim 1^\circ$, the N -body results are clearly higher than either PT (due to non-linear evolution) or the real APM results (see also top-left panel in Fig. 53 for the corresponding result for the three-point function). The latter is likely due to galaxy biasing operating at small scales [30], as discussed in the last section.

Estimation of higher-order moments from the EDSGC [622] up to $p = 8$ is in good agreement within the errors with APM on scales $\theta \gtrsim 0.1^\circ$. On smaller scales, $\theta \lesssim 0.1^\circ$, the EDSGC estimates are significantly larger than the APM values, indicating systematic problems in the deblending of crowded fields [627].¹³⁰ The DeepRange results [635] for the corresponding APM slice ($I_{AB} = 17\text{--}18$) give values of S_3 and S_4 which are intermediate between the APM and the EDSGC. This is also the case for the R INT-WFC catalog [540]. At larger scales, on the other hand, they both give slightly smaller results. This is not a very significant deviation but might indicate that the DeepRange survey is not large enough at this bright end and it therefore suffers from the same biases that are apparent when the APM S_p estimations are split in its 6×6 square degree plates. For the fainter slices the DeepRange results are less subject to volume effects and seem to indicate smaller values of S_3 and S_4 [635] as a function of depth (see Fig. 55). Finally, we note also that the skewness has been estimated for radio sources in the FIRST survey [426] (see also [165] for measurements of the angular correlation function), giving values $s_3 = 1\text{--}9$ for a depth corresponding to $1\text{--}50$ Mpc/ h , approximately.

8.2.6. Constraints on biasing and primordial non-Gaussianity

Galaxy biasing and primordial non-Gaussianity can leave significant imprints in the structure of the correlation hierarchy, as discussed in detail in Sections 7.1 and 4.4 and 5.6, respectively. These effects are best understood at large scales, where PT applies and simple arguments such as local galaxy biasing (see e.g. Section 7.1.1) are expected to hold. The APM survey is at present the largest angular survey probing scales in the weakly non-linear regime, thus most constraints on biasing and primordial non-Gaussianity from angular clustering have been derived from it. For constraints derived from galaxy redshift surveys see Section 8.3.5.

The lower-left panel in Fig. 53 shows the linear prediction (dotted lines), corresponding to the projection of Eq. (186) [514], for χ^2 initial conditions (see Section 4.4) with the APM-like initial spectrum [225]. Although the error bars are large and highly correlated, the projected three-point

¹²⁸ These errors should be considered more realistic than those given in the fifth and sixth entry in Table 16, which were derived by combining results at different angular scales assuming they are uncorrelated [249]. These error bars also correspond roughly to a $2\text{-}\sigma$ confidence in a single all-sky map: they are twice as large as those in Fig. 47.

¹²⁹ See e.g. Eq. (587) and Section 7.2.4 for a discussion of projection in the weakly non-linear regime.

¹³⁰ Measurements in this paper were done with an infinite oversampling technique [625]. In general, results without significant oversampling could underestimate S_p (see also [328]) but this does not explain the difference with the APM analysis, where the oversampling was adequate.

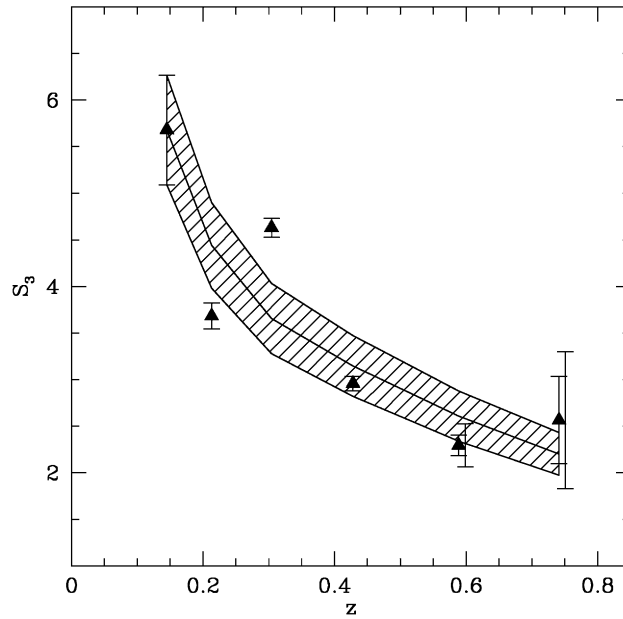


Fig. 55. The solid symbols display s_3 measured at 0.04° for six magnitude slices ($I_{AB} = 17-18, 18-19, 19-20, 20-21, 21-22, 22-22.5$, corresponding to increasing mean redshift) of the DeepRange catalog (from [635]). Each value of s_3 is plotted at the median z of the slice. The shaded band shows the predictions of a model of galaxy bias evolution, see text for details. The right-shifted error bars for the two faintest measurements include errors due to star/galaxy separation [635].

function for this model is substantially larger than that of the APM measurements and the corresponding Gaussian model for intermediate α . This may seem only a qualitative comparison, since as discussed in Section 4.4, non-linear corrections for this model are very significant even at large scales. However, non-linear corrections lead to even more disagreement with the data: although the shape dependence resembles that of the Gaussian case, the amplitude of q_3 when non-linear corrections are included becomes even larger than the linear result, especially at colinear configurations (see Figure 17).

This is also in agreement with Gaztañaga and Mähönen [252], who used the deprojected S_p from the APM survey [249] to constrain non-Gaussian initial conditions from texture topological defects [655] which, as in the case of the χ^2 model, also have dimensional scaling $\xi_N \sim B_N \xi_2^{N/2}$, with $B_3 \approx B_4 \approx 0.5$ (see Fig. 29). In this case it was found that N -body simulations of texture-type initial conditions lead to a significant rise at large scales in the S_p parameters not seen in the APM data, even when including linear and non-linear (local) bias to match the amplitude of S_p at some scale.

Constraints on a non-local biasing model from the APM S_p parameters were considered in [248]. The model of cooperative galaxy formation [96], where galaxy formation is enhanced by the presence of nearby galaxies, was suggested to produce a scale-dependent bias to create additional large-scale power in the standard CDM model and thus match the APM angular correlation function. However, the effect of this scale dependence bias is to imprint a significant scale dependence on the S_p parameters that is ruled out by the APM measurements (see also Fig. 57).

The upper right panel in Fig. 53 shows the PT predictions for the APM-like initial power spectrum, Eq. (627), with linear bias parameter $b_1 = 2$ (dashed curve) and a non-linear (local) bias model [see e.g. Eq. (525)] with $b_1 = 1$, $b_2 = -0.5$ (dotted curve). Even if the errors are 100% correlated, these models are in disagreement with the APM data. A more quantitative statement cannot be made about constraints of bias parameters from the APM higher-order moments since a detailed analysis of the covariance matrix is required. However, for linear bias the measurements imply that APM galaxies are unbiased to within 20–30% [225]. These constraints agree well with the biasing constraints obtained from the inflection point of the reconstructed $\xi_2(r)$ in the APM [260]. On the other hand, consideration of non-linear biasing can open up a wider range of acceptable linear bias parameters [224,248,671].

An alternative to wide surveys which probe the weakly non-linear regime at recent times, deep galaxy surveys can probe the redshift evolution and also reach weakly non-linear scales at high redshift. Although presently this is not possible due to the small size of current deep surveys, it will become so in the near future. An early application along these lines is in Fig. 55, which shows the redshift evolution of s_3 for measurements of [635] in the DeepRange catalog at a fixed angular scale of 0.04° . This corresponds to about $0.3h^{-1}$ Mpc at $z \simeq 0.15$ and $1.5h^{-1}$ Mpc at $z \simeq 0.75$, so the scales involved are in the non-linear regime.¹³¹

The redshift evolution in Fig. 55 is just the opposite of that expected in generic (dimensional) non-Gaussian models, where the skewness s_3 should increase with redshift (see e.g. discussion in Section 5.6). However, since these scales are in the non-linear regime the predictions based on PT cannot be safely used, and galaxy biasing can behave in a more complicated way. In any case, the trend shown in Fig. 55 can be matched by a model, shown in a shaded band, where $S_3(z) = S_3(0)(1+z)^{-0.5}$ [635], which may indicate that galaxy bias is increasing with redshift, as expected in standard scenarios of galaxy formation (see discussion in Section 7.1), and contrary to the evolution expected from strongly non-Gaussian initial conditions. A more quantitative constraint will have to await the completion of future deep surveys that can probe the weakly non-linear regime.

8.3. Results from redshift galaxy surveys

8.3.1. Redshift catalogs

Redshift surveys map the 3D distribution of galaxies in a large volume, and are thus ideally suited to use higher-order statistics to probe galaxy biasing and primordial non-Gaussianity. Table 17 shows a list of the main wide-field redshift catalogs. For a more general review on redshift catalogs see [484,268,607,608].

Redshift surveys require a predefined sample of targets to obtain redshifts; therefore, they are often defined from angular surveys where galaxies are detected photometrically. Below we shortly discuss the main characteristics of the surveys in Table 17, for a brief description of the photometric parent catalogs see Section 8.2.1.

The Center for Astrophysics survey (hereafter CfA [324]) and the Perseus–Pisces redshift Survey (PPS [268]) are both based on the Zwicky catalog. The CfA survey, perhaps the most analyzed redshift survey in the literature, consists of 2417 galaxies with Zwicky magnitudes less than 14.5,

¹³¹ Although a fixed angular scale does not correspond to a fixed spatial scale as a function of z , the comparison is meaningful because the measured $s_3(\theta)$ is scale independent (hierarchical) to a good approximation.

Table 17

Optical and infrared (last four) redshift catalogs. The survey area Ω is given in steradians, the depth and effective size $\mathcal{D}_E \equiv (\Omega/4\pi)^{1/3}2\mathcal{D}$ are in Mpc/h

Name	Area Ω (sr)	Magnitudes	Depth \mathcal{D}	\mathcal{D}_E	No. gal/s	Ref.
CfA	1.8	$m_Z < 14.5$	50	52	~ 1000	[324]
SSRS	1.8	$D(0) > 0.1$	50	52	~ 1000	[168]
PPS	~ 1	$m > 15.5-15$	80	70	~ 3000	[268]
LCRS	0.02	$R < 17.8$	300	70	1.3×10^6	[584]
Stromlo-APM	1.3	$b_J < 17.15$	150	140	1400	[411]
Durham/UKST	0.45	$b_J < 17$	140	90	5500	[535]
2dFGRS	0.6	$b_J < 19.5$	300	220	$\sim 2.5 \times 10^5$	[481]
SDSS	$\simeq 3$	$r' < 18$	275	341	$\sim 10^6$	[704]
QDOT	10	$f_{60\mu\text{m}} > 0.6\text{Jy}$	90	170	245	[200]
IRAS 1.9Jy	9.5	$f_{60\mu\text{m}} > 1.9\text{Jy}$	60	110	220	[606]
IRAS 1.2Jy	9.5	$f_{60\mu\text{m}} > 1.2\text{Jy}$	80	145	480	[218]
PSCz	10.5	$f_{60\mu\text{m}} > 0.6\text{Jy}$	100	188	1470	[550]

covering over 2.67 sr (1.8 sr in the North Galactic cap) with a median redshift corresponding to 3300 km/s. The PPS survey, centered around the Perseus–Pisces supercluster, contains over 3000 galaxies. The Southern Sky Redshift Survey (hereafter SSRS, [168]) is based on the ESO/Uppsala angular sample, and contains about 2000 galaxies. These surveys suffer from the same calibration problems as their parent catalogs, but with redshift information they were aimed to represent a fair sample of the universe. Recent extensions of these surveys to deeper magnitudes ($m < 15.5$, 2000 redshift, $\mathcal{D} \simeq 80$ Mpc/h) are denoted by CfA2 and SSRS2 and have been merged into the Updated Zwicky Catalog (UZC [208]).

The LCRS [584], consists of redshifts selected from a well-calibrated CCD survey of 6 narrow $1.5^\circ \times 80^\circ$ strips in the sky. Although this survey is much deeper and better calibrated than any of the previous ones, it is also potentially subject to important selection and boundary effects: narrow slices, density-dependent sampling (because of a constant number of fibers per field) and the exclusion of galaxies closer than $55''$. All these effects tend to underweight clusters and, even if properly corrected, could introduce important sampling biases in higher-order statistics.¹³²

The Stromlo-APM redshift survey [411] consists of 1790 galaxies with $b_J < 17.15$ selected randomly at a rate of 1 in 20 from APM scans in the south Galactic cap. The Durham/UKST galaxy redshift survey [535] consists of 2500 galaxy redshifts to a limiting apparent magnitude of $b_J = 17$, covering a 1500 sq deg area around the south galactic Pole. The galaxies in this survey were selected from the EDSGC and were sampled, in order of apparent magnitude, at a rate of one galaxy in every three.

The IRAS Point Source Redshift Catalog (hereafter PSCz [550]) is based on the IRAS Point Source Catalog (see [123]), with several small additions applied to achieve the best possible uniformity over the sky. The survey objective was to get a redshift for every IRAS galaxy with 60 μm flux

¹³² For example, it is impossible to recover any lost configuration dependence of correlation functions in the non-linear regime by a correction procedure, since the correcting weight for lost galaxies would have to decide whether they were aligned or isotropically distributed.

$f_{60} > 0.6\text{Jy}$, over as much of the sky as possible. Sky coverage is about 84% with 15 411 galaxies. Earlier subsamples of PSCz include the updated QDOT catalog [200], the IRAS 1.9Jy. [606] and the IRAS 1.2Jy [218] redshift surveys. The QDOT survey chooses at random one in six galaxies from PSCz, leading to 1824 galaxies with galactic latitude $|b| > 10^\circ$. The other subsamples are shallower but denser than QDOT; the 2Jy catalog, complete to a flux limit $f_{60} > 2\text{Jy}$. contains 2072 galaxies, whereas the 1.2Jy. catalog, with $f_{60} > 1.2$ contains 4545 galaxies. IRAS galaxies are mostly biased towards spiral galaxies which tend to undersample rich clusters. Thus IRAS galaxies are both sparser and a biased sample of the whole galaxy population.

The Sloan Digital Sky Survey (SDSS, see e.g. [699]) and the two-degree field 2dF Galaxy Redshift Survey (2dFGRS, see [142]) were still under construction when this review was written and only preliminary results are known at this stage. These results are discussed in Section 8.4.

Other recent redshift surveys for which there is not yet measurements of higher-order statistics include the Canada–France Redshift survey [401], the Century survey [266], the ESO Slice Project [673], the Updated Zwicky Catalog [208] and the CNOC2 Field galaxy survey [113].

8.3.2. Two-point statistics

We now briefly discuss results on two-point statistics from redshift surveys, with emphasis on the power spectrum. We first address optical surveys and then infrared surveys.

The analysis of the redshift-space correlation function in the CfA survey [172] found that, after integration over the parallel direction to project out redshift distortions, the resulting two-point function agreed with that derived from inversion in angular catalogs, Eq. (626), with $\gamma \simeq 1.77$ and $r_0 = 5.4 \pm 0.3 \text{ Mpc}/h$, for projected separations $r_p < 10 \text{ Mpc}/h$. At larger scales, the redshift-space correlation function estimates become steeper and there was marginal evidence for a zero crossing at scales larger than about $20 \text{ Mpc}/h$.¹³³ Modeling the redshift-space correlation function as a convolution of the real-space one with an exponential pairwise velocity distribution function¹³⁴ with velocity dispersion σ_v , Davis and Peebles [172] obtained that $\sigma_v = 340 \pm 40 \text{ km/s}$ at $r_p = 1 \text{ Mpc}/h$, well below the predictions of CDM models.

These results were extended a decade later with the analysis of the power spectrum in the extension of the CfA survey to $m_Z < 15.5$. In [679] it was shown that, in agreement with previous results from the APM survey [422] and IRAS galaxies [200,549], the standard CDM model was inconsistent with the large-scale power spectrum at the 99% confidence level. In addition, Park et al. [489] studied the relation between the real-space and redshift-space power spectrum in CDM simulations, using Eq. (617), and showed that agreement between the small-scale power spectrum and $\Gamma = 0.2$ CDM models required a velocity dispersion parameter $\sigma_v \approx 450 \text{ km/s}$, somewhat larger than the value obtained by modeling the two-point function in [172]. A joint analysis of the CfA/PPS power

¹³³ The measured redshift two-point function will be found to be flatter than the real-space one, with more power on large scales and less power on smaller scales, as expected from theory (see Section 7.4), with evidence for a larger correlation length in redshift space, $s_0 > r_0$, in all CfA, SSRS and IRAS catalogs [239].

¹³⁴ An exponential form was first suggested in [507] and has since been supported by observations, see e.g. [395] for a recent method applied in the LCRS survey. The interpretation of this technique, however, rests on the assumption of a scale-independent velocity dispersion, which seems consistent in LCRS [349], but may not necessarily be true in general, see e.g. [298,351] for the PSCz survey. Theoretically, exponential distributions arise from summing over Gaussian distributions, both in the weakly and highly non-linear regimes, see [357] and [586,186], respectively. These results are also supported by N -body simulations [217,708].

spectrum gave a best fit CDM shape parameter $\Gamma = 0.34 \pm 0.1$ [31]. Similarly, a joint analysis of the CfA/SSRS samples in [169] showed a power spectrum consistent with CDM models with $\Gamma \approx 0.2$ and bias within 20% of unity when normalized to COBE [596,694] CMB fluctuations at the largest scales. A recent analysis [487] of the redshift-space large-scale ($k \lesssim 0.3h/\text{Mpc}$) power spectrum of the Updated Zwicky catalog [208], which includes CfA2 and SSRS, was done using the quadratic estimator and decorrelation techniques (see Sections 6.11.2–6.11.4). The measurements in different subsamples are well fit by a Λ CDM model with normalization $b_1\sigma_8 = 1.2$ –1.4.

The analysis of the LCRS redshift-space power spectrum was done in [403], where they used Lucy’s method [413] to deconvolve the effects of the window of the survey, which are significant given the nearly two-dimensional geometry. They obtained results which were consistent with previous analyses of the CfA2 and SSRS surveys. An alternative approach was carried out in [394], where they estimated the 2D power spectrum, which was found to have a “bump” at $k=0.067$ Mpc/h with amplitude a factor of ≈ 1.8 larger than the smooth best fit $\Gamma = 0.3$ CDM model. This is reminiscent of similar features seen in narrow deep “pencil beams” redshift surveys, e.g. [99].¹³⁵ A recent linear analysis of the LCRS survey [443] using the KL transform methods (see e.g. Section 6.11.4), parameterized the power spectrum in redshift space by a smooth CDM model, and obtained a shape parameter $\Gamma = 0.16 \pm 0.10$, and a normalization $b_1\sigma_8 = 0.79 \pm 0.08$.

The two-point correlation function of LCRS galaxies was measured in [654,349], and integrated along the line of sight to give the projected correlation function in real-space, which was found to agree with Eq. (626), with $\gamma \simeq 1.86 \pm 0.04$ and $r_0 = 5.06 \pm 0.12$ Mpc/h [349]. After modeling the pairwise velocity distribution function by an exponential with dispersion and mean (infall) velocity, the inferred pairwise velocity at 1 Mpc/h was found to be $\sigma_v = 570 \pm 80$ km/s, substantially higher compared to other surveys. In fact, another analysis of the LCRS survey in [395] found a pairwise velocity dispersion of $\sigma_v = 363 \pm 44$ km/s, closer to previous estimates. In this case, the deconvolution of the small-scale redshift distortions was done by a Fourier transform technique, assuming a constant velocity dispersion and no infall [i.e. negligible \mathbf{u}_{12} , see Eq. (198)]. At least part of this disagreement can be traced to the effects of infall, as shown in [348]. For other recent methods and applications to determining the small-scale pairwise velocity dispersion and infall see e.g. [176] and [359], respectively.

Results from the power spectrum of the Stromlo-APM survey [412,639], the Durham/UKST survey [318] and the ESO Slice Project [115] are in agreement with previous results from optically selected galaxies, and show an amplification compared to the power spectrum of IRAS galaxies implying a relative bias factor $b_{\text{opt}}/b_{\text{iras}} \approx 1.2$ –1.3. This is reasonable, since IRAS galaxies are selected in the infrared and are mostly spiral galaxies which, from the observed morphology–density relation [192,529], tend to avoid clusters. We shall come back to this point when discussing higher-order statistics.

The first measurements of counts-in-cells in the QDOT survey [200,549] showed that IRAS galaxies were more highly clustered at scales of 30–40 Mpc/h compared to the predictions of the standard CDM model, in agreement with the angular correlation function from APM [422]. The QDOT power spectrum was later measured in [212] using minimum variance weighting, giving redshift-space values $\sigma_8 = 0.87 \pm 0.07$ and $\Gamma = 0.19 \pm 0.06$. Measurement of the power spectrum of the 1.2Jy survey [215] confirmed and extended this result, although it showed somewhat less power at large scales

¹³⁵ See e.g. [363,488] and the recent analysis in [701] for a discussion of the statistical significance of these features.

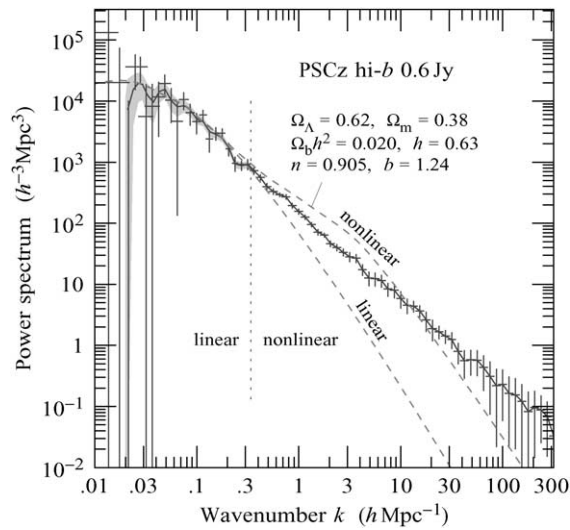


Fig. 56. The *real-space* power spectrum of PSCz galaxies. To the left of the vertical line is the linear measurement of [299] (points with uncorrelated error bars [297]), while to the right is the non-linear measurement of [298] (points with *correlated* error bars). The dashed line corresponds to the flat Λ CDM concordance model power spectrum from [650] with parameters as indicated, non-linearly evolved according to the prescription in [494]. (From [298]).

than QDOT.¹³⁶ The measurement of the two-point function in redshift space for the 1.2Jy survey [216,217] implied a real-space correlation function as in Eq. (626), with $\gamma \simeq 1.66$ and $r_0 = 3.76$ Mpc/h for scales $r \lesssim 20$ Mpc/h, consistent with the fact that IRAS galaxies are less clustered than optically selected galaxies. In addition, the inferred velocity dispersion at 1 Mpc/h was $\sigma_v = 317_{-49}^{+40}$ km/s.

Measurements in the PSCz survey are currently the most accurate estimation of clustering of IRAS galaxies. At large scales, the power spectrum is intermediate between that of QDOT and 1.2Jy surveys, whereas at smaller scales it decreases slightly more steeply [612]. The shape of the large-scale power spectrum is consistent with $\Gamma = 0.2$ CDM models, although it does not strongly rule out other models [612,640]. A comparison with the Stromlo-APM survey shows a relative bias parameter of $b_{\text{Stromlo}}/b_{\text{PSCz}} \approx 1.3$ and a correlation coefficient between optical and IRAS galaxies of $R \geq 0.72$ at the 95% confidence limit on scales of the order of 20 Mpc/h [573]. These results were considerably extended in [298] to obtain the power spectrum in *real-space* by measuring the redshift-space power perpendicular to the line of sight and parameterizing the dependence on non-perpendicular modes to increase signal to noise. The resulting power spectrum is reproduced in Fig. 56. It shows a nearly power-law behavior to the smallest scales measured, with no indication of an inflexion at the non-linear scale, and no sign of turnover at the transition to the stable clustering regime. Compared to the best-fit CDM model (obtained from a joint analysis with CMB fluctuations in [650] and shown as a dashed line), the PSCz requires a significant scale-dependent bias.

¹³⁶ It was later shown that the QDOT measurements were sensitive to a small number of galaxies in the Hercules supercluster [202,638], which was over-represented in the QDOT sample presumably due to a statistical fluctuation in the random numbers used to construct the survey.

Table 18
Some measurements of Q_3 in redshift catalogs

$3Q_3$	Scales	Sample	Year	Ref.
2.4 ± 0.2	—	CfA	1980	[508] (Eq. (57.9))
2.04 ± 0.15	—	[541]	1981	[508]
2.4 ± 0.3	1–2	CfA	1984	[198]
1.8 ± 0.2	1–3	Durham-AAT	1983	[32]
3.9 ± 0.9	1–2	KOSS	1983	[32]
1.5–4.5	1–8	LCRS	1998	[347]
$Q_{\text{eq}} \simeq 0.5$	0.1–1.6	CfA/PPS	1991	[31]
$Q_3 \approx 1$	0.05–0.2	QDOT	2001	[567]
$Q_{\text{eq}} \simeq 0.2$; $Q_{\text{col}} \simeq 0.6$	0.05–0.2	IRAS 1.9Jy	2001	[567]
$Q_{\text{eq}} \simeq 0.4$; $Q_{\text{col}} \simeq 0.8$	0.05–0.2	IRAS 1.2Jy	2001	[567]
$Q_{\text{eq}} \simeq 0.4$; $Q_{\text{col}} \simeq 1.4$	0.05–0.4	PSCz	2001	[211]

In most cases, only the mean values over a range of scales were published. In cases where measurements of the individual values for each scale are reported in the literature, we quote the actual range of estimates over the corresponding range of scales. The top half of the table is in configuration space, the bottom part in Fourier space. Scales are in Mpc/ h and h /Mpc, respectively. When possible, we give estimates for equilateral (eq) and colinear (col) configurations. Error bars should be considered only as rough estimates, see text for discussion.

Finally, we briefly mention results on the parameter $\beta \approx \Omega^{0.6}/b_1$ from measurements of the anisotropy of the power spectrum in redshift space¹³⁷ (see Section 7.4.2). These measurements are complicated by the fact that surveys are not yet large enough to see a clear transition into the linear regime predictions, Eq. (614). In addition, different methods seem to give somewhat different answers [295]; however, the average and standard deviation of reported values are [295] $\beta_{\text{opt}} = 0.52 \pm 0.26$ and $\beta_{\text{iras}} = 0.77 \pm 0.22$ for optically selected and IRAS galaxies, respectively, which is roughly consistent with the relative bias between these two populations. On the other hand, the most recent results from the PSCz survey find $\beta = 0.39 \pm 0.12$ [643], and $\beta = 0.41^{+0.13}_{-0.12}$ [299]. Constraints from the most recent optically selected surveys are considerably noisier, e.g. Stromlo-APM does not even exclude $\beta \sim 1$ [412,639], and LCRS is consistent with no distortions at all, $\beta = 0.30 \pm 0.39$ [443]. Resolution of these issues will have to await results from the full-volume 2dFGRS and SDSS surveys (see also Section 8.4).

8.3.3. Three-point statistics

Determination of three-point statistics from redshift surveys has been carried out mostly in the non-linear regime for optically selected surveys, and mostly in the weakly non-linear regime for IRAS surveys. Table 18 shows different estimates of the three-point function (top list) and the bispectrum (bottom list).

As discussed before, the CfA sample covers a small volume to be a fair estimate of higher-order correlations. Even more so, estimates in the Durham-AAT and KOSS samples are subject to large estimator biases as they have only a few hundred redshifts. Nonetheless, these results roughly agree

¹³⁷ For an exhaustive review of these results up to mid-1997 see [295].

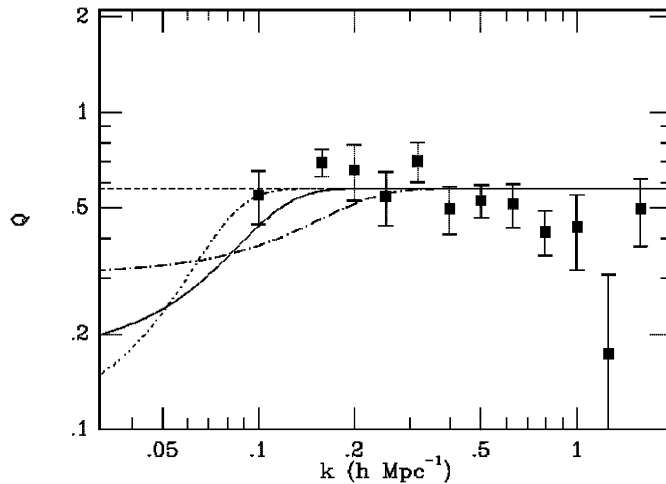


Fig. 57. The redshift-space reduced bispectrum Q_{eq} for equilateral triangles as a function for scale k , for CfA/PPS galaxies (from [31]). The dashed line shows the PT prediction: $Q_{\text{eq}} = \frac{4}{7}$. The other lines show predictions for the cooperative galaxy formation models, see Section 8.3.5.

with each other, although the values of Q_3 are seen to fluctuate significantly. Note that the values in Table 18 are not directly comparable to those inferred from deprojection of angular catalogs (Table 15) as they are affected by redshift distortions (see e.g. Fig. 48).

The LCRS survey provides the best estimate to date of the three-point function at small scales [347]. Estimation of Q_3 in redshift space and in projected space (by integrating along the line of sight) showed values lower by a factor of about 2 than Λ CDM simulations where clusters have been underweighted by $m^{-0.08}$, essentially equivalent to assuming that the number of galaxies as a function of dark matter halo mass m scales as $N_{\text{gal}}(m) \sim m^{0.9}$ in the notation of Section 7.1.3. The authors conclude that the hierarchical model is not a good description of the data, since they see some residual (small) scale and configuration dependence. However, as discussed at the end of Section 7.4.3, one *does not* expect the hierarchical model to be a good description for correlations in redshift space since velocity dispersion creates “fingers of god” along the observer’s direction [562]. The fact that these are clearly seen by visual inspection of the galaxy distribution ought to show up in a clear shape dependence of the three-point function: colinear configurations should be significantly amplified (see Fig. 48). Surprisingly, this is not seen in the LCRS measurements [347].

Measurements of the bispectrum (for equilateral configurations) in redshift space were first carried out for the CfA survey and a sample of redshifts in the Pisces–Perseus super-cluster [31]. This was the first measurement that reached partially into the weakly non-linear regime and compared the bispectrum for equilateral configurations with PT predictions, $Q_{\text{eq}} = \frac{4}{7}$. As shown in Fig. 57 the agreement with PT predictions is very good, even into the non-linear regime.¹³⁸ The errors bars in each bin indicate the variance among different subsamples, three from the CfA and three from the Perseus–Pisces surveys. This result was interpreted as a support for gravitational instability

¹³⁸ This is due to accidental cancellations in redshift space. At larger k ’s, in the absence of redshift distortions, $Q_{\text{eq}}(k)$ increases, see e.g. Fig. 15. However, velocity dispersion suppresses this rise, resulting in approximately the same value as in PT [562]. The same is not true for colinear configurations, see Fig. 48.

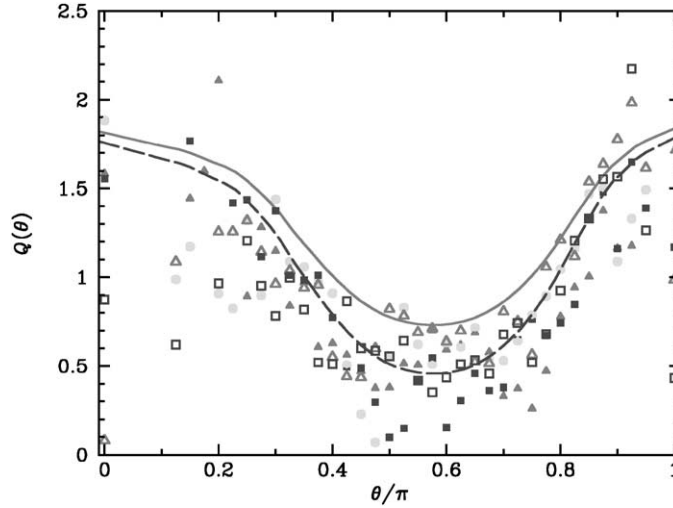


Fig. 58. The bispectrum Q_3 vs. θ for the PSCz catalog for triangles with $0.2 \leq k_1 \leq 0.4$ h/Mpc and with two sides of ratio $k_2/k_1 = 0.4-0.6$ separated by angle θ . The solid curve shows Q_3 in redshift space averaged over many 2LPT realizations of the Λ CDM model. Symbols show results from the PSCz survey for bands in k_1 : (\blacktriangle) $k_1 = 0.20-0.24$ h/Mpc; (\blacksquare) $0.24-0.28$; (\bullet) $0.28-0.32$; (\circ) $0.32-0.36$; (\square) $0.36-0.42$. The dashed curve shows the 2LPT prediction for Λ CDM with the best-fit bias parameters $1/b = 1.20$, $b_2/b^2 = -0.42$. Taken from [211].

from Gaussian initial conditions and in disagreement with models of threshold bias [21,344], which predicted $Q_3 \sim 1$. The results in Fig. 57 were later used in [224] to constrain models of non-local bias that had been proposed to give galaxies extra large-scale power in the standard CDM scenario (see Section 8.3.5 for a discussion). In addition, Baumgart and Fry [31] measured the trispectrum for randomly generated tetrahedral configurations, showing a marginal detection with hierarchical scaling consistent with $Q_4 \sim 1$.

Detailed measurements of the bispectrum in the weakly non-linear regime were not done until a decade later, with the analyses of the IRAS surveys [567,211], which probe a large enough volume of roughly spherical shape. In [567], measurements were done for the QDOT, 1.9Jy and 1.2Jy surveys. In order to constrain galaxy bias and primordial non-Gaussianity, a likelihood method that takes into account the covariance matrix of the bispectrum for different triangles and the non-Gaussian shape of the likelihood function (see e.g. Fig. 42) was used, developed in [566]. This is essential to recover accurate estimates of errors on bias parameters and primordial non-Gaussianity without systematic estimator biases due to the finite volume of the survey.¹³⁹ The results from QDOT were marginal, due to the very sparse sampling (one galaxy every six) Q_3 was only shown to be of order unity without any discernible dependence on configuration. The results from 1.9Jy and 1.2Jy showed a systematic shape dependence similar to that predicted by gravitational instability.

These results were considerably extended with the analysis of the PSCz bispectrum [211]. Fig. 58 shows the PSCz reduced bispectrum Q_3 as a function of the angle θ between \mathbf{k}_1 and \mathbf{k}_2 for triangles

¹³⁹ A likelihood analysis for analysis of the bispectrum was first proposed in [434], based on the Gaussian approximation for the likelihood function and a second-order Eulerian PT calculation of the covariance matrix. Extensions to redshift space are given in [31].

Table 19
Some measurements of S_3 and S_4 in redshift catalogs

S_3	S_4	Scales	Sample	Year	Ref.
$2 \pm 1 - 6 \pm 4$	—	5–20	QDOT	1991	[549]
1.5 ± 0.5	4.4 ± 3.7	0.1–50	IRAS 1.2Jy	1992	[88,92]
1.9 ± 0.1	4.1 ± 0.6	2–22	CfA	1992	[246]
2.0 ± 0.1	5.0 ± 0.9	2–22	SSRS	1992	[246]
2.1 ± 0.3	7.5 ± 2.1	3–10	IRAS 1.9Jy	1994	[224]
2.4 ± 0.3	13 ± 2	2–10	PPS	1996	[267]
2.8 ± 0.1	6.9 ± 0.7	8–32	IRAS 1.2Jy	1998	[377]
1.8 ± 0.1	5.5 ± 1	1–10	SSRS2	1999	[35]
1.9 ± 0.6	7 ± 4	1–30	PSCz	2000	[632]
1.82 ± 0.21	~ 3	12.6	Durham/UKST	2000	[319]
2.24 ± 0.29	~ 8	18.2	Stromlo/APM	2000	[319]

In most cases, only the mean values over a range of scales were published. In cases where measurements of the individual values for each scale are reported in the literature, we quote the actual range of estimates over the corresponding range of scales. In most cases error bars should be considered only as rough estimates, see text for discussion.

with $k_1/k_2 \approx 2$ and different scales as described in the figure caption [211]. The configuration dependence predicted by gravitational instability [232,313] (solid lines for an unbiased distribution, predicted by 2LPT, see e.g. Fig. 48) is clearly seen in the data. This is also the case for all triangles, not just those shown in Fig. 58, see Fig. 1 in [211].

Implications of these results for galaxy biasing and primordial non-Gaussianity are discussed in Section 8.3.5.

8.3.4. Skewness, kurtosis and higher-order cumulants

Table 19 shows different estimates for $S_3 = \bar{\xi}_3/\bar{\xi}_2^2$ and $S_4 = \bar{\xi}_4/\bar{\xi}_2^3$, the ratios of the cumulants $\bar{\xi}_N$ obtained by counts-in-cells. The shape of the cells correspond to top-hat spheres, unless stated otherwise.

The QDOT results by Saunders et al. [549] were obtained from counts-in-cells with a Gaussian window. The errors, from a minimum variance scheme, are quite large but they suggest a hierarchical scaling $\bar{\xi}_3 \simeq \bar{\xi}_2^2$, with a value of S_3 consistent with gravity from Gaussian initial conditions, as argued in [137].

Fig. 59 displays the 1.2Jy IRAS results ([88,92], left panel) and CfA-SSRS results ([246], right panel). There is a convincing evidence for the hierarchical scaling in $\bar{\xi}_3$ and $\bar{\xi}_4$ (denoted by straight lines) but the resulting S_3 and S_4 amplitudes are probably affected by sampling biases (see discussion below). Note that the scaling is preserved well into the non-linear regime, this is in agreement with expectations from N -body simulations which show that in redshift space the growth of S_p parameters toward the non-linear regime is suppressed by velocity dispersion from virialized regions ([391,437], see e.g. Fig. 49).

In their analysis of higher-order moments in the CfA, SSRS and IRAS 1.9 Jy catalogs, Frieman and Gaztanaga [224] studied the sensitivity of S_p to redshift distortions, by calculating moments in spherical cells and conical cells. The latter were argued to be less sensitive to the redshift space

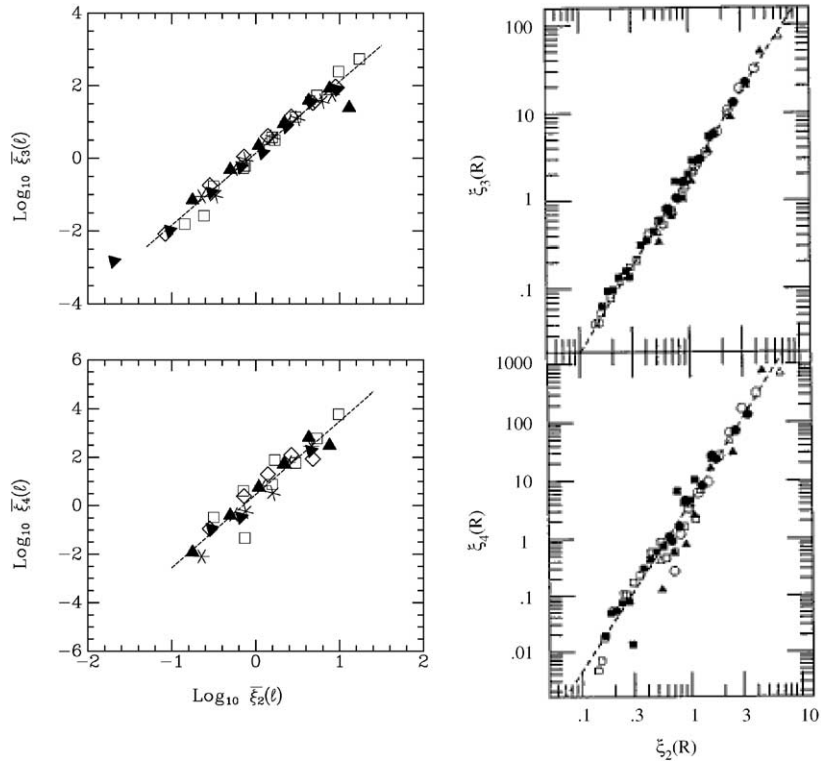


Fig. 59. Values of $\bar{\xi}_3(R)$ and $\bar{\xi}_4(R)$, as a function of $\bar{\xi}_2(R)$ in the IRAS (left, from [92]) and in the CfA and SSRS (right, from [246]) redshift catalogs. The lines show the best-fit amplitude for the hierarchical scaling $\bar{\xi}_N = S_N \bar{\xi}_2^{N-1}$.

mapping that acts along the line of sight.¹⁴⁰ They find that although cumulants $\bar{\xi}_p$ are sensitive to the change in cell geometry, the S_p parameters were not.

On the other hand, Ghinaa et al. [267] estimated the third- and fourth-order cumulants using moments of counts centered in galaxies [84] in the PPS. After a somewhat ad-hoc correction for virial fingers to recover “real space” quantities, they find a variation of S_3 and S_4 with scale, compatible with a non-negligible cubic term, e.g. $\bar{\xi}_3 \sim S_3 \bar{\xi}_2^2 + C_3 \bar{\xi}_2^3$. Since the scale where the cubic term becomes important is found to be about 5 Mpc/h, this is perfectly consistent with gravitational clustering: at these scales loop corrections are expected to increase (the real space) S_3 and S_4 , see e.g. Figs. 28 and 49.

An alternative method to moments of count-in-cells was proposed by Kim and Strauss [377], who parameterized the count PDF by an Edgeworth expansion (see Section 3.5) convolved with a Poisson distribution to take into account discreteness effects. This method is only applicable at large enough scales (and small enough δ/σ) so that the Edgeworth expansion holds; however convolution

¹⁴⁰ This is certainly true in the limit of large radial distances. At finite size, structures will still look less concentrated in conical cells than in real space due to velocity dispersion. Note that the conical geometry may introduce a change in $\bar{\xi}_N$ since not all N -point configurations are weighted equally.

with a Poisson distribution helps to regularize the resulting PDF (i.e. it is positive definite).¹⁴¹ The advantage of this method is that one can obtain the S_p from a likelihood analysis of the shape of the PDF near its maximum, rather than relying on the tails of the distribution which are sensitive to rare clusters, as in the moments method.¹⁴² One disadvantage is that error estimation in this framework is more complicated, although in principle not insurmountable. Results from N -body simulations show this method to be more reliable at large scales [377] than the standard approach. Application to the 1.2Jy survey [377] resulted in values for S_3 and S_4 significantly higher than in previous work using moments [92], see Table 19.

Measurements of the higher-order moments in the SSRS2 survey were obtained in [35]. Results for S_3 and S_4 were shown to be consistent with hierarchical at all scales probed (the error bars quoted in Table 19 were found by averaging over all scales assuming uncorrelated measurements). A study of the errors in numerical simulations showed that bootstrap resampling errors were underestimated by a factor of order two. A re-analysis of the data using the Edgeworth method of [377] showed that S_3 changed upward by a factor of about two to $S_3 \sim 3$, similar to the change seen in the IRAS 1.2Jy survey.

A recent analysis of the PSCz survey [632], which should be affected much less than previous IRAS surveys by finite-volume effects, was carried out by using minimum variance estimates of moments of counts-in-cells in volume-limited subsamples (see Section 6.9). The values of S_3 and S_4 found, shown in Fig. 60, are consistent within the errors¹⁴³ with that of previous IRAS results, including those found by deprojection from angular counts [451,92,239] and also (for S_3) in agreement with the amplitude obtained from measurements of the bispectrum [211] (see Fig. 58). They also found that the measurements of S_3 and S_4 agreed very well with the predictions of the semi-analytic galaxy formation model in [36], based on models of spiral galaxies in the framework of Λ CDM models.

A similar analysis technique was used in the Stromlo-APM and Durham/UKST surveys [319]. In this case measurements of the skewness are in agreement with those found in shallower redshift surveys (CfA, IRAS 1.2Jy, SSRS2) but with larger (but more realistic) errors. Comparison with deprojected values for S_3 and S_4 obtained from the parents catalogs APM [249] and EDSGC [622] shows a systematic trend where redshift surveys give systematically smaller values than angular surveys. The most significant contribution to this apparent discrepancy is likely to be redshift distortions: as shown in Fig. 49 for scales $R \lesssim 20$ Mpc/ h the S_p parameters are suppressed in redshift space.¹⁴⁴ At scales larger than 20 Mpc/ h results from the redshift and parent angular surveys should agree since redshift distortions do not affect the S_p significantly [313]. In this regime, the results from APM/EDSGC surveys seem systematically higher, although no more than 1σ given the large error bars. In this case other systematic effects might be taking place. Deprojection from angular

¹⁴¹ However, for future applications to surveys not as sparse as the IRAS galaxy distribution, such as 2dFGRS and SDSS, this will not be the case.

¹⁴² The peak of the PDF is however sensitive to the largest voids in the sample (see e.g. Fig. 20), which can influence the most likely value of δ and thus the S_p derived from such method.

¹⁴³ One should take into account that errors in previous analyses have been underestimated. The more realistic errors in [632] were obtained using the FORCE code [621,152,630], which is based on the full theory of cosmic errors as described in Section 6.

¹⁴⁴ This is for dark matter, however at these scales bias should not make a qualitative difference. Furthermore, deviations in galaxy surveys are seen at similar scales [319].

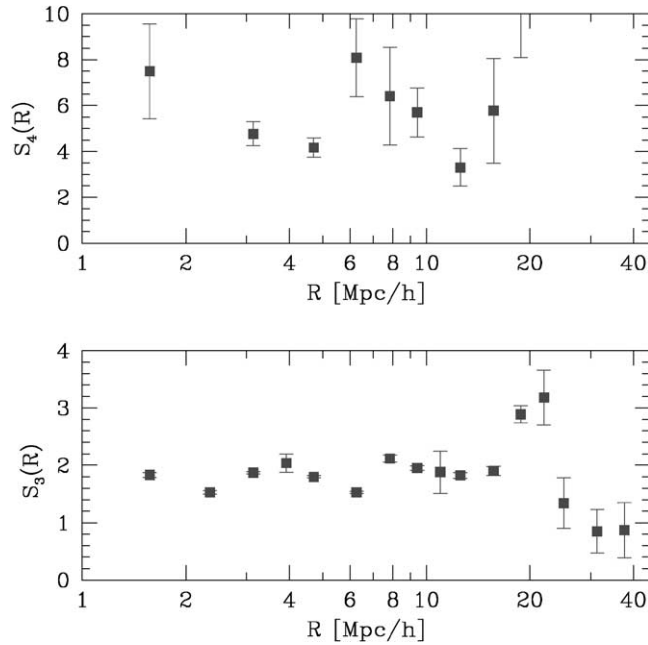


Fig. 60. The redshift-space skewness S_3 and kurtosis S_4 as a function of smoothing scale in the PSCz survey [211].

surveys using the hierarchical model rather than the configuration dependence predicted by PT can cause an overestimation of the 3D S_p that can be as much as 20% for S_3 (see e.g. Fig. 47). In addition, finite-volume effects [621,328,630] as discussed in Section 6 can lead to underestimation of S_p from redshift surveys that are typically sampling a smaller volume.¹⁴⁵

8.3.5. Constraints on biasing and primordial non-Gaussianity

We now review implications of the above results for biasing and primordial non-Gaussianity, concentrating on higher-order statistics. Effects of primordial non-Gaussianity on the power spectrum have been considered in [212,605,612]. The results presented here are complementary to recent studies of the impact of primordial non-Gaussian models in other aspects of large-scale structure such as the abundance of massive clusters [538,385,691,522].

Results on the redshift-space bispectrum in the CfA/PPS sample [31] (see Fig. 57) and the skewness of CfA/SSRS surveys [246] were used in [224] to put constraints on the non-local (scale-dependent) bias in the cooperative galaxy formation (CGF) model [96] proposed to generate enough large-scale power in the context of otherwise-standard CDM. This model corresponds to a (density-dependent) threshold bias model where galaxies form in regions satisfying $\delta > \nu\sigma - \kappa\delta(R_s)$, where κ is the strength of cooperative effects and R_s describes the “scale of influence” of non-locality. Fig. 57 shows the predictions of CGF models for $(\kappa, R_s) = (0.84, 10h^{-1}\text{Mpc})$ (dot-long-dashed), $(2.29, 20h^{-1}\text{Mpc})$ (solid) and $(4.48, 30h^{-1}\text{Mpc})$ (dot-short-dashed), all of which have similar large-scale power to a $\Gamma = 0.2$ CDM model. Because of the scale dependence induced by

¹⁴⁵ These effects are thought to be dominant for smaller surveys such as CfA/SSRS, see [328] for a detailed discussion.

the CGF models, additional linear bias is required to suppress these features, which in turn implies non-zero non-linear bias to maintain agreement with $Q_3 \sim 0.5$ and also would be in disagreement with the normalization implied by the CMB [596]. In addition, this would make the agreement with the simple prediction of PT from Gaussian initial conditions purely accidental. Similar results follow from the analysis of the skewness S_3 , see [224].

As discussed in Section 8.3.3, the detection of the configuration dependence of the bispectrum in IRAS surveys (see e.g. Fig. 58) gives a tool to constrain galaxy bias, primordial non-Gaussianity and break degeneracies present in two-point statistics. Using a maximum likelihood method that takes into account the non-Gaussianity of the cosmic distribution function and the covariance matrix of the bispectrum [566], the constraints on local bias parameters from IRAS surveys assuming Gaussian initial conditions¹⁴⁶ read [567,211]

$$\frac{1}{b_1} = 1.32^{+0.36}_{-0.58}, \quad \frac{b_2}{b_1^2} = -0.57^{+0.45}_{-0.30} \quad (2\text{Jy.}), \quad (630)$$

$$\frac{1}{b_1} = 1.15^{+0.39}_{-0.39}, \quad \frac{b_2}{b_1^2} = -0.50^{+0.31}_{-0.51} \quad (1.2\text{Jy.}), \quad (631)$$

$$\frac{1}{b_1} = 1.20^{+0.18}_{-0.19}, \quad \frac{b_2}{b_1^2} = -0.42^{+0.19}_{-0.19} \quad (\text{PSCz}), \quad (632)$$

with the best-fit model shown as a dashed line in Fig. 58 for the PSCz case. These results for the linear bias of IRAS galaxies, when coupled with measurements of the power spectrum redshift distortions, which determine $\beta = \Omega_m^{0.6}/b_1 \simeq 0.4 \pm 0.12$ for the PSCz survey [299,643], allow the break of the degeneracy between linear bias and Ω_m , giving $\Omega_m = 0.16 \pm 0.1$.

If bias is local in Lagrangian, rather than Eulerian space, the bispectrum shape depends differently on bias parameters [120], see Eq. (537). Physically this corresponds to galaxies that form depending exclusively on the initial density field, and then evolved by gravity. Eulerian bias, on the other hand, corresponds to the other extreme limit where galaxies form depending exclusively on the present (non-linear) density field. Both limits are undoubtedly simplistic, but analysis of the bispectrum in the PSCz survey suggests that the Eulerian bias model is more likely than the Lagrangian one [211].

The bispectrum results can also be used to constraint non-Gaussian initial conditions. In this case one must also take into account the possibility of galaxy biasing, which is more complicated since the usual formula for Gaussian initial conditions, Eq. (528), is not valid anymore, but it is calculable in terms of the primordial statistics [565]. Using a χ^2 model as an example of dimensional scaling models (where $\xi_N \sim \xi_2^{N/2}$, see Section 4.4.2), it was shown that the IRAS 1.2Jy bispectrum is inconsistent with the amplitude and scaling of this type of initial conditions at the 95% level [567].

The PSCz bispectrum provides stronger constraints upon non-Gaussian initial conditions. In [211] χ_N^2 statistics were considered as a general example of dimensional scaling models. For $N=1$, this corresponds to the predictions of some inflationary models with isocurvature perturbations [531,7,404]; as $N \rightarrow \infty$ the model becomes effectively Gaussian, and for a fixed power spectrum (taken to fit that of PSCz) the primordial bispectrum obeys $Q^I \propto N^{-1/2}$ [565]. From the PSCz data, it follows

¹⁴⁶In addition, these constraints assume a fixed linear power spectrum shape given by $\Gamma = 0.21$, in agreement with power spectrum measurements. See [566,567] for sensitivity of bias parameters on the assumed power spectrum shape. The dependence of the bispectrum on the assumed Ω_m is negligible, as first pointed out in [313].

that $N > 49(22)$ at 68% (95%) CL. Since the primordial dimensionless skewness is $B_3 = 2.46$ for a χ_1^2 field [514], the PSCz bispectrum constrains $B_3 < 0.35(0.52)$. These results are independent of (local) biasing, and they are obtained by marginalizing over bias parameters [211].

8.4. Recent results from 2dFGRS and SDSS

Looking at the overall picture, clustering statistics have been measured in a wide range of observational data. The catalogs listed in Tables 14 and 17 cover angular surface densities that are almost six orders of magnitude apart, solid angles ranging over more than three orders of magnitude, depths that go from $50\text{--}2000h^{-1}$ Mpc, and volumes ranging over three orders of magnitude. They also involve quite different systematics, from photographic plates to satellite missions and different observational filters. Despite these large differences, and after carefully correcting for systematic effects, all data on higher-order statistics in the weakly non-linear regime seem to be in good agreement with gravitational instability from Gaussian initial conditions. This provides a remarkable step forward in our understanding of structure formation and points to gravity as the basic mechanism to build cosmic structure from small primordial fluctuations generated in the early universe.

Needless to say, the observational results reviewed here, although providing a consistent picture, have significant limitations. The magnitude of statistical and systematic errors is still rather large and the range of scales available in the weakly non-linear regime is quite restricted. In the next few years this is expected to change significantly, with the completion of the new generation of wide-field surveys such as 2dFGRS and SDSS. Here we provide a brief summary of the results that have been recently reported in the literature from these preliminary samples.

The 2dFGRS has recently publically released their first versions of galaxy and quasars catalogs, containing 100 000 [142] and 10 000 redshifts [167], respectively. The completed survey is expected to reach 250 000 galaxies and 25 000 quasars. The parent source catalog is an extended and revised version of the APM survey [425], with galaxies with magnitudes $b_J < 19.45$. For a review of the recent results see [496].

A measurement of the redshift-space two-point correlation function was presented in [497] from analysis of 141 402 galaxies. Using a phenomenological model similar to that in Eq. (617) with input real-space power spectrum obtained by deprojection from the APM survey [27], they obtain a velocity dispersion parameter $\sigma_v = 385$ km/s and, after marginalizing over σ_v , a best-fit estimate of $\beta = 0.43 \pm 0.07$. These results are obtained by considering only the two-point function data for $8h^{-1}$ Mpc $< r < 25h^{-1}$ Mpc.

A preliminary analysis of the redshift-space power spectrum is presented in [518], based on a sample of 147 024 galaxies. After taking into account the window of the survey, and assuming linear perturbation theory at scales $0.02 \lesssim k \lesssim 0.15$ h/Mpc, they obtain that models containing baryons oscillations are marginally ($\sim 2\sigma$) preferred over featureless spectra. Assuming scale invariance for the primordial power spectrum, their analysis gives $\Omega_m h = 0.20 \pm 0.03$ and a baryon fraction $\Omega_b/\Omega_m = 0.15 \pm 0.07$, in good agreement with recent determinations from measurements of the CMB power spectrum [476,282]. The most recent analysis [652] of the publically released 100 000 galaxy sample using KL eigenmodes finds however no significant detection of baryonic wiggles, although their results are consistent with the previous analyses using a larger sample, but less sophisticated techniques.

Using a series of volume-limited samples, Norberg et al. [480] present a measurement of the projected correlation function by integrating the redshift-space two-point function along the line of sight. The result is well described by a power law in pair separation over the range $0.1h^{-1} \text{ Mpc} < r < 10h^{-1} \text{ Mpc}$, with $r_0 = 4.9 \pm 0.3h^{-1} \text{ Mpc}$ and $\gamma = 1.71 \pm 0.06$, see Eq. (626). Measurements for different samples spanning a factor of 40 in luminosity show a remarkable little variation in the power-law slope, with all correlation functions being almost parallel with amplitudes spanning a factor of about three.

These results have been confirmed by recent measurements in a preliminary sample of the SDSS survey [704] containing 29 300 galaxy redshifts. They find a scale-independent luminosity bias for scales $r < 10h^{-1} \text{ Mpc}$, with different subsamples having nearly parallel projected correlation functions with power-law slope $\gamma \sim 1.8$. For the whole sample, the correlation length is $r_0 = 6.1 \pm 0.2h^{-1} \text{ Mpc}$ and the power-law slope $\gamma = 1.75 \pm 0.03$, for scales $0.1h^{-1} \text{ Mpc} < r < 16h^{-1} \text{ Mpc}$. The inferred velocity dispersion is $\sigma_v \simeq 600 \pm 100 \text{ km/s}$, nearly independent of scale for projected separations $0.15h^{-1} \text{ Mpc} < r_p < 5h^{-1} \text{ Mpc}$.

A series of papers have recently analyzed angular clustering of over a million galaxies in a rectangular stripe of $2.5^\circ \times 90^\circ$ from early SDSS data. The analysis of systematic effects and statistical uncertainties is presented in [572], where the angular correlation function is calculated and the impacts of several potential systematic errors is evaluated, from star/galaxy separation to the effects of seeing variations and CCD systematics, finding all of them to be under control. The Limber scaling test is performed and is shown to make angular correlation functions corresponding to all four magnitude bins agree when scaled to the same depth.¹⁴⁷ Analysis of statistical errors includes calculation of covariance matrices for $w_2(\theta)$ in the four slices using 200 realizations of mock catalogs constructed using the PTHaLos code [571] and also using the subsampling and jackknife methods.

Analysis of the angular correlation function is presented in [157], which is found to be consistent with results from previous surveys (see also [261]). On scales between 1° and 1 arcmin, the correlation functions are well described by a power law with an exponent of about -0.7 , in agreement with Eq. (625). The amplitude of the correlation function within this angular interval decreases with fainter magnitudes in accordance with previous galaxy surveys. There is a characteristic break in the correlation functions on scales close to $1\text{--}2^\circ$ degrees, showing a somewhat smaller amplitude at large scales (for the corresponding magnitude slice) than the APM correlation function. On small scales, less than an arcminute, the SDSS correlation function does not appear to be consistent with the same power law fitted to the larger angular scales. This result should however be regarded as preliminary due to the still limited amount of data (only 1.6% of the final size of the SDSS photometric sample) and the uncertainties in modeling the covariance matrix of $w_2(\theta)$ obtained from the mock catalogs described above.

The angular power spectrum $P_{2D}(l)$ is obtained in [651] for large angular scales corresponding to multiple moments $\ell \lesssim 600$. The data in all four magnitude bins are shown to be consistent with a simple Λ CDM “concordance” model with non-linear evolution (particularly evident for the brightest galaxies) and linear bias factors of order unity. The results were obtained using KL compression, quadratic estimators and presented in terms of uncorrelated band powers (Section 6.11). These results, together with those of the angular correlation function [572,157], are used in [189] to perform an inversion to obtain the 3D power spectrum, using a variant of the SVD decomposition method of

¹⁴⁷ These correspond to $r^* = 18\text{--}19, 19\text{--}20, 20\text{--}21, 21\text{--}22$, with median redshifts $\bar{z} \sim 0.17, 0.25, 0.35, 0.46$ [189].

[204]¹⁴⁸ with the corresponding covariance matrix computed from the mock catalogs. The resulting 3D power spectrum estimates from both inversions agree with each other and with previous estimates from the APM survey for $0.03 h/\text{Mpc} < k < 1 h/\text{Mpc}$. These results are shown to agree with an alternative method presented in [617], where the projected galaxy distribution is expanded in KL eigenmodes and the 3D power spectrum parameters recovered are $\Gamma = 0.188 \pm 0.04$ and $b_1\sigma_8 = 0.915 \pm 0.06$.

Preliminary results for the higher-order correlations in the SDSS have been presented in [261,262, 637], including s_3, s_4, q_3 and c_{12} statistics. In all cases a very good agreement with previous surveys was found. In particular, at the bright end, the agreement with the APM results is quite remarkable despite the important differences in survey design and calibration. These results confirm the need for non-trivial biasing at small scales, as discussed in Sections 8.2.4 and 8.2.5 (see also Fig. 54).

9. Summary and conclusions

As illustrated throughout this work, PT provides a valuable tool to understand and calculate predictions for the evolution of large-scale structure in the universe. The last decade has witnessed a substantial activity in this area, with strong interplay with numerical simulations of structure formation and observations of clustering of galaxies and, more recently, weak gravitational lensing. As galaxy surveys become larger probing more volume in the weakly non-linear regime, new applications of PT are likely to flourish to provide new ways of learning about cosmology, the origin of primordial fluctuations, and the relation between galaxies and dark matter.

The general framework of these calculations is well established and calculations have been pursued for a number of observational situations, whether it is for the statistical properties of the local density contrast, the velocity divergence, for the projected density contrast, redshift measurements or for more elaborate statistics such as joint density cumulants. All these results provide robust frameworks for understanding the observations or for reliable error computations. There are, however, a number of outstanding issues that remain to be addressed in order to improve our understanding of gravitational instability at large scales:

- Most of the calculations have been done assuming Gaussian initial conditions, except for some specific cases such as χ^2 models. Although present observations are consistent with Gaussian initial conditions, deriving quantitative constraints on primordial non-Gaussianity requires some knowledge or useful parametrization of non-Gaussian initial conditions and how they evolve by gravity.
- Predictions of PT for velocity field statistics are still in a rudimentary state compared to the case of the density field. Upcoming velocity surveys will start probing scales where PT predictions can be used. In addition, robust methods for calculating redshift distortions including the non-linear effects due to the redshift-space mapping are needed to fully extract information from the next generation of galaxy redshift surveys.
- Another observational context in which a PT approach can be very valuable is the Lyman- α forest observed in quasar spectra. The statistical properties of these systems should be accessible to

¹⁴⁸ See Section 8.2.3 for a brief discussion of inversion procedures and results.

perturbative methods since most of the absorption lines correspond to modest density contrasts (from 1 to 10). This is a very promising field for observational cosmology.

- Accurate constraints on cosmological parameters from galaxy surveys require precise models of the joint likelihood of low- and higher-order statistics including their covariance matrices. To date this has only been investigated in detail numerically, or analytically in some restricted cases.

In addition, as we probe the transition to the non-linear regime, there are a few technical issues that need more investigation,

- Most results have been obtained in the tree-level approximation, for which systematic calculations can be done and the emergence of non-Gaussianity can be characterized in an elegant way. There is no such systematic framework for loop corrections, and only a few general results are known in this case. Furthermore, loop corrections are found to be divergent for power-law spectra with index $n \geq -1$, the interpretation of which is still not clear. Although this issue is irrelevant for realistic spectra such as CDM, its resolution may shed some light into the physics of the non-linear regime.
- The SC collapse prescription (Section 5.5.2) leads to a good description of S_p parameters in the transition to the non-linear regime when compared to N -body simulations and exact one-loop corrections when known. Is it possible to improve on this approximation, or make it more rigorous in any well-controlled way while maintaining its simplicity?
- The development of HEPT (Section 4.5.6) and EPT (Section 5.13) suggests that there is a deep connection between gravitational clustering at large and small scales. Is this really so, or is it just an accident? Why do strongly non-linear clustering amplitudes seem to be so directly related to initial conditions?

From the observational point of view, the next few years promise to be extremely exciting, with the completion of 2dFGRS and SDSS and deep surveys that will trace the evolution of large-scale structure toward high redshift.¹⁴⁹ Observations of the so-called Lyman break galaxies [603] should soon provide a precious probe of the high-redshift universe, in particular regarding the evolution of galaxy bias [2,528,110]. Furthermore, weak lensing observations will provide measurements of the projected mass density that can be directly compared with theoretical predictions. In addition, CMB satellites and high-resolution experiments will probe scales that overlap with galaxy surveys and thus provide a consistency check on the framework of the growth of structure.

Outstanding observational issues abound, most of them perhaps related to the way galaxies form and evolve. One of the most pressing ones, as discussed many times in Section 8, is probably to have a convincing explanation of why correlation functions scale as power laws at non-linear scales. The scaling in Figs. 50 and 56 is certainly remarkable and preliminary results from 2dFGRS [280] and SDSS [704,157] seem already to confirm and extend these results. In the CDM framework, however, this simple behavior is thought to be the result of accidental cancellation of the dark matter non-power-law form by scale-dependent bias due to the way dark matter halos are populated by galaxies (see discussion in Section 7.1.3). Although this may seem rather adhoc, this model has, on the other hand, many observable consequences. The same weighting that makes the two-point function depend as a power law of separation [579,495,570] suppresses the velocity dispersion and

¹⁴⁹ See e.g. [130] for a recent assessment of how well upcoming deep surveys will determine correlation functions.

mean streaming of galaxies [591,592] as observed, see e.g. [349]. In addition, this weighting affects higher-order statistics in the non-linear regime, suppressing them in comparison with their dark matter counterparts [570] (see Fig. 45) as observed, see e.g. Fig. 54 for a comparison between dark matter and S_p in the APM survey. There are also complementary indications that galaxies do not trace the underlying dark matter distribution at small scales from measurements of higher-order statistics. As discussed in Section 8.2.3, reconstruction of the linear power spectrum from galaxy surveys leads to significant disagreement of higher-order moments if no biasing is imposed at small scales, as shown in Fig. 54 for APM galaxies. A promising way to confirm that the underlying higher-order statistics of the dark matter are much higher than those of galaxies at small scales is by measuring higher-order moments in weak gravitational lensing. This will likely be done in the near future, as weak lensing surveys are already beginning to probe the relation between dark matter and galaxies [315].

In any case, statistical analysis of future observations are going to decide whether the small-scale behavior of correlations is dictated by biasing or if a new framework is needed to understand galaxy clustering at non-linear scales. What seems clear, whatever the outcome, is that the techniques described here will be a valuable tool to achieve that goal.

Acknowledgements

This project was possible thanks to the hospitality of several institutions that supported frequent visits. We thank CSIC, IAP, IAS, IEEC, SPhT, and also CITA during the initial stages of this work. We also benefited greatly from discussions with F. Bouchet, J. Frieman, J. Fry, R. Juszkiewicz and I. Szapudi that help set the structure of this review. We thank Marc Kamionkowski for many helpful comments about a previous version of this work. This project has made extensive use of NASA's Astrophysics Data System Abstract Service.

Appendix A. The spherical collapse dynamics

The spherical collapse dynamics can be obtained from the Friedmann equations of the expansion factor in different cosmologies. It amounts to solve the motion equation for the radius R of a shell collapsing under its own gravity,

$$\frac{d^2R}{dt^2} = -G \frac{M(< R)}{R^2}, \quad (\text{A.1})$$

where $M(< R)$ is the mass encompassed in a radius R . The corresponding density contrast can be defined as

$$\delta_{\text{sc}}(t) = \frac{M(< R)}{\bar{\rho}4\pi R^3/3} - 1. \quad (\text{A.2})$$

Explicit solutions are known for open or closed universes without cosmological constant. The complete derivation of them can be found in [508] where the density contrast is expressed as a function of time t . We present the results here in a slightly different way by expressing the non-linear density contrast as a function of the linear density contrast, $\varepsilon (\equiv D_+(t)\delta_{\text{init}})$ [43].

For an open universe the background evolution is described by parameter ψ_0 so that the current value of the density parameter is given by

$$\Omega_0 = \frac{2}{1 + \cosh \psi_0} . \tag{A.3}$$

Similarly the density fluctuation is characterized by a parameter θ . There is a minimal value of the linear density contrast below which the density fluctuation is still below critical and does not collapse. This is given by

$$\epsilon_{\min} = \frac{9}{2} \frac{\sinh \psi_0 (\sinh \psi_0 - \psi_0)}{(\cosh \psi_0 - 1)^2} . \tag{A.4}$$

As a result, if the linear density contrast $\epsilon > \epsilon_{\min}$, the evolution of the perturbation density is given by

$$\delta_{\text{sc}}(\epsilon) = \left(\frac{\cosh \psi_0 - 1}{-\cos \theta + 1} \right)^3 \left(\frac{-\sin \theta + \theta}{\sinh \psi_0 - \psi_0} \right)^2 - 1 \tag{A.5}$$

with

$$\epsilon = \epsilon_{\min} \left[\left(\frac{-\sin \theta + \theta}{\sinh \psi_0 - \psi_0} \right)^{2/3} + 1 \right] . \tag{A.6}$$

If $\epsilon < \epsilon_{\min}$, we have

$$\delta_{\text{sc}}(\epsilon) = \left(\frac{\cosh \psi_0 - 1}{\cosh \theta - 1} \right)^3 \left(\frac{\sinh \theta - \theta}{\sinh \psi_0 - \psi_0} \right)^2 - 1 \tag{A.7}$$

with

$$\epsilon = -\epsilon_{\min} \left[\left(\frac{\sinh \theta - \theta}{\sinh \psi_0 - \psi_0} \right)^{2/3} - 1 \right] . \tag{A.8}$$

The Einstein–de Sitter case is recovered when $\psi_0 \rightarrow 0$. It implies that $\epsilon_{\min} \rightarrow 0$. In this case the solution reads, for $\epsilon < 0$,

$$\delta_{\text{sc}}(\epsilon) = \frac{9}{2} \frac{(\sinh \theta - \theta)^2}{(\cosh \theta - 1)^3} - 1 , \tag{A.9}$$

$$\epsilon = -\frac{3}{5} \left[\frac{3}{4} (\sinh \theta - \theta) \right]^{2/3} , \tag{A.10}$$

and for $\epsilon > 0$,

$$\delta_{\text{sc}}(\epsilon) = \frac{9}{2} \frac{(\theta - \sin \theta)^2}{(1 - \cos \theta)^3} - 1 , \tag{A.11}$$

$$\epsilon = \frac{3}{5} \left[\frac{3}{4} (\theta - \sin \theta) \right]^{2/3} . \tag{A.12}$$

In the limit $\Omega_0 \rightarrow 0$ we have $\psi_0 \rightarrow \infty$. It implies that $\epsilon_{\min} \rightarrow \frac{3}{2}$. Moreover, ϵ is finite when θ is close to ψ_0 so that

$$\epsilon = \frac{3}{2} \left(\frac{\exp \theta}{\exp \psi_0} \right) - 1, \quad \delta_{\text{sc}} = \frac{\exp \psi_0}{\exp \theta}, \quad (\text{A.13})$$

which gives

$$\delta_{\text{sc}}(\epsilon) = \frac{1}{(1 - 2\epsilon/3)^{3/2}}. \quad (\text{A.14})$$

The case of a closed universe is obtained by the change of variable $\psi_0 \rightarrow i\psi_0$.

Appendix B. Tree summations

In this appendix we present methods for performing tree summations. These calculations have been developed initially in different contexts (such as polymer physics, see e.g. [185]). In cosmology, these computation techniques have been introduced in [551] and presented in detail in a more complex situation in [41].

B.1. For one field

The issue we address is the computation of the sum of all tree diagrams (in a specific sense given in the following) connecting an arbitrary number of points. More specifically we define $\varphi(y)$ as (minus) the sum of all diagrams with the weight $(-y)^n$ for diagrams of n points.

For computing the contribution of each order the rule is to build all possible minimal connection (that means $n - 1$ connections for n points) and to affect the value v_p to points connected to p neighbors. The value of each diagram is then given by the product of the vertices v_p it is composed of.

The function $\varphi(y)$ then corresponds to the cumulant generating function,

$$\varphi(y, v_1, v_2, \dots) = - \sum_{n=2}^{\infty} (-y)^n \sum_{\text{trees connecting } n \text{ points}} \left(\prod_{\text{vertices}} v_p \right). \quad (\text{B.1})$$

At the end of the calculation the value of v_1 will be unity, but for the time being we assume it is a free parameter. Then φ is a function of y and of the vertices v_p . We can then define τ as

$$-\tau = \frac{1}{-y} \frac{\partial(-\varphi)}{\partial v_1}. \quad (\text{B.2})$$

Like φ , $(-\tau)$ is a function of y and of the vertices v_p . This corresponds to all the diagrams for which *one* external line (connected to a v_1 vertex) has been marked away. This is the sum of the so-called diagrams with one free external line. It is possible to write down an implicit equation for τ ,

$$-\tau = -y \left(v_1 - v_2 \tau + v_3 \frac{\tau^2}{2} + \dots + v_p \frac{(-\tau)^{p-1}}{(p-1)!} + \dots \right). \quad (\text{B.3})$$

This equation expresses the fact that τ can be reconstructed in a recursive way (see Fig. 25). Note the factor $(p - 1)!$ which corresponds to the symmetry factor. If one defines the generating function of the vertices,

$$\zeta(\tau) = \sum_{p=1}^{\infty} v_p \frac{(-\tau)^p}{p!}, \tag{B.4}$$

then we have

$$\tau = -y \frac{\partial \zeta}{\partial \tau}. \tag{B.5}$$

To complete the calculation we need to introduce the Legendre transform $\mathcal{L}(\tau, v_2, \dots)$ defined as

$$\mathcal{L}(y, \tau, v_2, \dots) = \varphi + y v_1 \tau. \tag{B.6}$$

It is important to note that \mathcal{L} is viewed as a function of τ and not of v_1 . We then have the remarkable property due to the Legendre transform,

$$\frac{\partial \mathcal{L}}{\partial \tau} = \frac{\partial \varphi}{\partial v_1} \frac{\partial v_1}{\partial \tau} + y \tau \frac{\partial v_1}{\partial \tau} + y v_1 = y v_1. \tag{B.7}$$

From Eq. (B.3) we have

$$y v_1 = \tau - y \sum_{p=2}^{\infty} v_p \frac{(-\tau)^{p-1}}{(p-1)!}, \tag{B.8}$$

which after integrating relation (B.7) implies that

$$\mathcal{L} = c + \frac{\tau^2}{2} + y \sum_{p=2}^{\infty} v_p \frac{(-\tau)^p}{p!} = c + \frac{\tau^2}{2} + y \zeta(\tau) + y v_1 \tau, \tag{B.9}$$

which leads to (the integration constant $c = 0$ is such that $\varphi(y) \sim -y^2$ at leading order in y)

$$\varphi(y) = y \zeta(\tau) - \frac{1}{2} y \tau \zeta'(\tau). \tag{B.10}$$

This equation, with Eq. (B.3), gives the tree generating function expressed as a function of the vertex generating function ζ .

B.2. For two fields

We can extend the previous results to joint tree summations. It corresponds to either two different fields taken at the same position (as the density and the velocity divergence for instance), or to two fields taken at different locations. We want to construct the joint generating function, $\varphi(y_1, y_2)$, of the joint cumulants,

$$\varphi(y_1, y_2) = - \sum_{n, m, n+m \geq 2} C_{nm} \frac{(-y_1)^n}{n!} \frac{(-y_2)^m}{m!}, \tag{B.11}$$

where C_{nm} is the value of each cumulant. In this case for each diagram there are n vertices of type 1, and m of type 2. They take respectively the value v_p and μ_q if they are connected respectively to p or q neighbors. Obviously, if the two fields are identical the two series identify. Moreover, in order to account for cell separation, a weight ξ is put for each line connecting points of different nature.

The generating function φ is then a function of $y_1, y_2, \xi, v_1, \dots, \mu_1, \dots$. One can define the two functions τ_1 and τ_2 by

$$\tau_1 = \frac{1}{-y_1} \frac{\partial(-\varphi)}{\partial v_1}, \quad \tau_2 = \frac{1}{-y_2} \frac{\partial(-\varphi)}{\partial \mu_1}. \tag{B.12}$$

It is easy to see that the functions τ_1 and τ_2 are given respectively by

$$\tau_1 = y_1 \sum_{p=1}^{\infty} v_p \frac{(-\tau_1)^{p-1}}{(p-1)!} + \xi y_2 \sum_{p=1}^{\infty} \mu_p \frac{(-\tau_2)^{p-1}}{(p-1)!}, \tag{B.13}$$

$$\tau_2 = \xi y_1 \sum_{p=1}^{\infty} v_p \frac{(-\tau_1)^{p-1}}{(p-1)!} + y_2 \sum_{p=1}^{\infty} \mu_p \frac{(-\tau_2)^{p-1}}{(p-1)!}. \tag{B.14}$$

This expresses the fact that there is a joint recursion between the two functions. A factor ξ is introduced whenever a vertex of a given type is connected to vertex of the other type.

Defining the Legendre transform as $\mathcal{L} = \varphi + y_1 \tau_1 v_1 + y_2 \tau_2 \mu_1$, one obtains

$$\frac{\partial \mathcal{L}}{\partial \tau_1} = y_1 v_1, \quad \frac{\partial \mathcal{L}}{\partial \tau_2} = y_2 \mu_1. \tag{B.15}$$

One should then solve the linear system for v_1 and μ_1 given by Eqs. (B.13) and (B.14). One eventually gets for φ ,

$$\varphi = y_1 \zeta_1(\tau_1) + y_2 \zeta_2(\tau_2) + \frac{1}{2(1-\xi^2)} (\tau_1^2 - 2\xi \tau_1 \tau_2 + \tau_2^2), \tag{B.16}$$

where ζ_1 and ζ_2 are, respectively, the generating functions of v_p and μ_p . This result can be rewritten in a more elegant form,

$$\varphi(y_1, y_2) = y_1 \zeta_1(\tau_1) + y_2 \zeta_2(\tau_2) - \frac{1}{2} y_1 \tau_1 \zeta_1'(\tau_1) - \frac{1}{2} y_2 \tau_2 \zeta_2'(\tau_2). \tag{B.17}$$

If ξ is unity, for instance for the computation of the joint density distribution of (δ, θ) , we have

$$\tau = \tau_1 = \tau_2 = -y_1 \zeta_1'(\tau) - y_2 \zeta_2'(\tau). \tag{B.18}$$

B.3. The large separation limit

The other case of interest is when ξ is small (which means that the correlation function at the cell separation is much smaller than the average correlation function at the cell size).

It is then possible to expand $\varphi(y_1, y_2)$ at leading order in ξ . The results read

$$\varphi(y_1, y_2) = \varphi_1(y_1) + \varphi_2(y_2) - \tau_1^{(0)}(y_1)\xi\tau_2^{(0)}(y_2), \quad (\text{B.19})$$

where $\tau_1^{(0)}$ and $\tau_2^{(0)}$ are, respectively, the functions τ_1 and τ_2 computed when $\xi = 0$.

Appendix C. Geometrical properties of top-hat window functions

In this section we recall the properties of top-hat window function. The derivations are presented in a systematic way for any dimension of space D . The window function W_D in Fourier space is given by

$$W_D(k) = 2^{D/2} \Gamma(D/2 + 1) \frac{J_{D/2}(k)}{k^{D/2}}. \quad (\text{C.1})$$

We are interested in computing the angle integrals of $W_D(|\bar{l}_1 - \bar{l}_2|)$ times a geometrical function that can be expressed in terms of Legendre polynomials. In particular we want to compute $\int d^D\Omega W_D(|\bar{l}_1 - \bar{l}_2|)[1 - (\bar{l}_1\bar{l}_2)/(l_1^2 l_2^2)]$ and $\int d^D\Omega W_D(|\bar{l}_1 - \bar{l}_2|)[1 + \bar{l}_1\bar{l}_2/l_1^2]$. In general, the only angle that intervenes in the angular integral, $d^D\Omega$, is the relative angle φ so that $d^D\Omega/\Omega_{\text{tot}}$ reduces to $\Gamma(D/2)/(\sqrt{\pi}\Gamma[(D-1)/2])\sin(\varphi)^{D-2}d\varphi$, $0 \leq \varphi \leq \pi$.

In order to complete these calculations, we need the summation theorem (GR, 8.532) for Bessel function,

$$\frac{J_\nu(|\bar{l}_1 - \bar{l}_2|)}{|\bar{l}_1 + \bar{l}_2|^\nu} = 2^\nu \Gamma(\nu) \sum_{k=0}^{\infty} (\nu + k) \frac{J_{\nu+k}(l_1)}{l_1^\nu} \frac{J_{\nu+k}(l_2)}{l_2^\nu} C_k^\nu(\cos \varphi), \quad (\text{C.2})$$

where C_k^ν are Gegenbauer polynomials. Note that in the case of $\nu = 0$ the previous equation reads

$$J_0(|\bar{l}_1 - \bar{l}_2|) = J_0(l_1)J_0(l_2) + 2 \sum_{k=1}^{\infty} J_k(l_1)J_k(l_2) \cos(k\varphi). \quad (\text{C.3})$$

In the following, the only property of interest for the Gegenbauer polynomials is (GR, 7.323)

$$\int_0^\pi C_k^\nu(\cos \varphi) \sin^{2\nu}(\varphi) d\varphi = 0 \quad \text{for } k \geq 1, \quad (\text{C.4})$$

$$\int_0^\pi C_0^\nu(\cos \varphi) \sin^{2\nu}(\varphi) d\varphi = \frac{\pi \Gamma(2\nu + 1)}{2^{2\nu} \Gamma^2(1 + \nu)}. \quad (\text{C.5})$$

As a result we have

$$\begin{aligned} & \frac{\Gamma(D/2)}{\sqrt{\pi}\Gamma[(D-1)/2]} \int_0^\pi \sin(\varphi)^{D-2} d\varphi W_D(|\bar{l}_1 - \bar{l}_2|) \left[1 - \frac{(\bar{l}_1\bar{l}_2)^2}{l_1^2 l_2^2} \right] \\ &= \frac{2^D \Gamma^2(D/2) \Gamma(D/2 + 1)}{\sqrt{\pi}\Gamma[(D-1)/2]} \sum_{k=0}^{\infty} \left(\frac{D}{2} + k \right) \frac{J_{\nu+k}(l_1)}{l_1^\nu} \frac{J_{\nu+k}(l_2)}{l_2^\nu} \\ & \quad \times \int \sin(\varphi)^D d\varphi C_k^{D/2}(\cos \varphi). \end{aligned} \quad (\text{C.6})$$

The only non-vanishing term of this summation is the one corresponding to $k = 0$. We finally have

$$\int \frac{d^D \Omega}{\Omega_{\text{tot}}} W_D(|\bar{l}_1 - \bar{l}_2|) \left[1 - \frac{(\bar{l}_1 \bar{l}_2)^2}{l_1^2 l_2^2} \right] = \left(1 - \frac{1}{D} \right) W_D(l_1) W_D(l_2). \quad (\text{C.7})$$

This result writes as a kind of commutation rule: the filtering can be applied to the wave vectors separately provided the angular kernel is properly averaged.

The second relation can be obtained from the observation that

$$l \frac{d}{dl} \left[\frac{J_{D/2-1}(l)}{l^{D/2-1}} \right] = -\frac{J_{D/2}(l)}{l^{D/2}}, \quad (\text{C.8})$$

$$\frac{J_{D/2-1}(l)}{l^{D/2-1}} = l \frac{d}{dl} \left[\frac{J_{D/2}(l)}{l^{D/2}} \right] + d \frac{J_{D/2}(l)}{l^{D/2}}. \quad (\text{C.9})$$

The summation theorem applied to $J_{D/2-1}(|\bar{l}_1 + \bar{l}_2|)/|\bar{l}_1 + \bar{l}_2|^{D/2-1}$ leads to

$$\int \frac{d^D \Omega}{\Omega_{\text{tot}}} \frac{J_{D/2-1}(|\bar{l}_1 + \bar{l}_2|)}{|\bar{l}_1 + \bar{l}_2|^{D/2-1}} = \frac{2(D-2)\sqrt{\pi}\Gamma(-2+D)}{2^{D/2}\Gamma((-1+D)/2)} \frac{J_{D/2-1}(l_1)}{l_1^{D/2-1}} \frac{J_{D/2-1}(l_2)}{l_2^{D/2-1}}. \quad (\text{C.10})$$

Taking the derivative of this equality with respect to l_1 leads to

$$\int \frac{d^D \Omega}{\Omega_{\text{tot}}} W_D(|\bar{l}_1 - \bar{l}_2|) \left[1 - \frac{\bar{l}_1 \bar{l}_2}{l_1^2} \right] = W_D(l_1) \left[W_D(l_2) + \frac{l_2}{D} \frac{d}{dl_2} W_D(l_2) \right]. \quad (\text{C.11})$$

Appendix D. One-loop calculations: dimensional regularization

To obtain the behavior of the one-loop p -point spectra for $n < -1$, one can use dimensional regularization (see e.g. [143]) to simplify considerably the calculations. Since we are interested in the limit where the ultraviolet cutoff $k_c \rightarrow \infty$, all the integrals run from 0 to ∞ , and divergences are regulated by changing the dimensionality d of space: we set $d = 3 + \epsilon$ and expand in $\epsilon \ll 1$. For example, for one-loop bispectrum calculations, we need the following one-loop three-point integral:

$$J(v_1, v_2, v_3) \equiv \int \frac{d^d \mathbf{q}}{(q^2)^{v_1} [(\mathbf{k}_1 - \mathbf{q})^2]^{v_2} [(\mathbf{k}_2 - \mathbf{q})^2]^{v_3}}. \quad (\text{D.1})$$

When one of the indices vanishes, e.g. $v_3 = 0$, this reduces to the standard formula for dimensional-regularized two-point integrals [595]

$$J(v_1, v_2, 0) = \frac{\Gamma(d/2 - v_1)\Gamma(d/2 - v_2)\Gamma(v_1 + v_2 - d/2)}{\Gamma(v_1)\Gamma(v_2)\Gamma(d - v_1 - v_2)} \pi^{d/2} k_1^{d-2v_1-2v_2}, \quad (\text{D.2})$$

which is useful for one-loop power spectrum calculations. The integral $J(v_1, v_2, v_3)$ appears in triangle diagrams for massless particles in quantum field theory, and can be evaluated for arbitrary values of

its parameters in terms of hypergeometric functions of two variables. The result is [178]

$$\begin{aligned}
 J(v_1, v_2, v_3) = & \frac{\pi^{d/2} k_1^{d-2v_{123}}}{\Gamma(v_1)\Gamma(v_2)\Gamma(v_3)\Gamma(d-v_{123})} \times (\Gamma(v_3)\Gamma(v_{123}-d/2) \\
 & \times F_4(v_3, v_{123}-d/2; 1+v_{23}-d/2, 1+v_{13}-d/2; x, y) \\
 & \times \Gamma(d/2-v_{13})\Gamma(d/2-v_{23}) + y^{d/2-v_{13}} \Gamma(v_2)\Gamma(d/2-v_1) \\
 & \times F_4(v_2, d/2-v_1; 1+v_{23}-d/2, 1-v_{13}+d/2; x, y) \\
 & \times \Gamma(v_{13}-d/2)\Gamma(d/2-v_{23}) + x^{d/2-v_{23}} \Gamma(v_1)\Gamma(d/2-v_2) \\
 & \times F_4(v_1, d/2-v_2; 1-v_{23}+d/2, 1+v_{13}-d/2; x, y) \\
 & \times \Gamma(d/2-v_{13})\Gamma(v_{23}-d/2) + x^{d/2-v_{23}} y^{d/2-v_{13}} \Gamma(d/2-v_3) \\
 & \times F_4(d-v_{123}, d/2-v_3; 1-v_{23}+d/2, 1-v_{13}+d/2; x, y) \\
 & \times \Gamma(d-v_{123})\Gamma(v_{23}-d/2)\Gamma(v_{13}-d/2) \ , \tag{D.3}
 \end{aligned}$$

where $v_{123} \equiv v_1 + v_2 + v_3$, $v_{ij} \equiv v_i + v_j$, $x \equiv (\mathbf{k}_2 - \mathbf{k}_1)^2/k_1^2$, $y \equiv k_2^2/k_1^2$, and F_4 is Apell’s hypergeometric function of two variables, with the series expansion:

$$F_4(a, b; c, d; x, y) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \frac{x^i y^j}{i! j!} \frac{(a)_{i+j} (b)_{i+j}}{(c)_i (d)_j} \ , \tag{D.4}$$

where $(a)_i \equiv \Gamma(a+i)/\Gamma(a)$ denotes the Pochhammer symbol. When the spectral index is $n = -2$, the hypergeometric functions reduce to polynomials in their variables due to the following useful property for $-a$ a positive integer:

$$F_4(a, b; c, d; x, y) = \sum_{i=0}^{-a} \sum_{j=0}^{-a-i} \frac{x^j y^i}{j! i!} \frac{(b)_{i+j}}{(c)_i (d)_j} \frac{(-1)^{i+j} (-a)!}{(-a-i-j)!} \ . \tag{D.5}$$

When using expressions such as Eq. (D.3), divergences appear as poles in the gamma functions; these can be handled by the following expansion ($n = 0, 1, 2, \dots$ and $\epsilon \rightarrow 0$),

$$\Gamma(-n + \epsilon) = \frac{(-1)^n}{n!} \left[\frac{1}{\epsilon} + \psi(n+1) + \frac{\epsilon}{2} \left(\frac{\pi^2}{3} + \psi^2(n+1) - \psi'(n+1) \right) \right] \ , \tag{D.6}$$

plus terms of order ϵ^2 and higher. Here $\psi(x) \equiv d \ln \Gamma(x)/dx$ and

$$\psi(n+1) = 1 + \frac{1}{2} + \dots + \frac{1}{n} - \gamma_e \ , \tag{D.7}$$

$$\psi'(n+1) = \frac{\pi^2}{6} - \sum_{k=1}^n \frac{1}{k^2} \ , \tag{D.8}$$

with $\psi(1) = -\gamma_e = -0.577216\dots$ and $\psi'(1) = \pi^2/6$.

Appendix E. PDF construction from cumulant generating function

In this section we present the mathematical relation between the cumulant generating function defined in Section 3 and the one-point probability distribution function of the local density, and more generally the counts-in-cells probabilities.

In this presentation we follow the calculations (and most of notations) developed in [16].

E.1. Counts-in-cells and generating functions

Let us consider a cell of volume V placed at random in the field. We denote $P(N)$ as the probability that this cell contains N particles. One can define the probability distribution function $\mathcal{P}(\lambda)$ as

$$\mathcal{P}(\lambda) = \sum_{N=0}^{\infty} \lambda^N P(N). \quad (\text{E.1})$$

By construction, the counts-in-cells probabilities $P(N)$ are obtained by a Taylor expansion of $\mathcal{P}(\lambda)$ around $\lambda = 0$:

$$P(N) = \frac{1}{N!} \frac{d^N}{d\lambda^N} \mathcal{P}(\lambda = 0). \quad (\text{E.2})$$

Remarkably, the (factorial) moments of this distribution are obtained by a Taylor expansion of $\mathcal{P}(\lambda)$ around $\lambda = 1$:

$$\begin{aligned} \mathcal{P}(1) &= 1, \\ \frac{d}{d\lambda} \mathcal{P}(1) &= \bar{N}, \\ \frac{d^2}{d\lambda^2} \mathcal{P}(1) &= \langle N(N-1) \rangle, \\ &\vdots \\ \frac{d^p}{d\lambda^p} \mathcal{P}(1) &= \langle N(N-1) \cdots (N-p+1) \rangle. \end{aligned} \quad (\text{E.3})$$

If the field is an underlying Poisson distribution of a continuous field, then the factorial moments, $\langle N(N-1) \cdots (N-p+1) \rangle$, are equal to $\bar{N}^p M_p$ where M_p is the p th moment of the local density distribution. $\mathcal{P}(\lambda)$ can therefore be written in terms of the moment-generating function (see Section 3.3.3), $\mathcal{P}(\lambda) = \mathcal{M}[(\lambda-1)\bar{N}]$, which in turn can be written in terms of the cumulant generating function, $\mathcal{C}(\lambda-1)$:

$$\mathcal{P}(\lambda) = \exp(\mathcal{C}[(\lambda-1)\bar{N}]). \quad (\text{E.4})$$

When the cumulant generating function is written in terms of the S_p generating function, the counts-in-cells read

$$P(N) = \oint \frac{1}{2\pi i} \frac{d\lambda}{\lambda^{N+1}} \exp \left[-\frac{\bar{N}\bar{\xi}(1-\lambda) + \varphi(\bar{N}\bar{\xi}(1-\lambda))}{\bar{\xi}} \right], \quad (\text{E.5})$$

where the integral is made in the complex plane around the singularity $\lambda = 0$. One can change the variable to use $y = \bar{\xi}(1 - \lambda)$ so that

$$P(N) = \frac{-1}{\bar{N}\bar{\xi}} \oint \frac{dy}{2\pi i} \left(1 - \frac{y}{\bar{N}\bar{\xi}}\right)^{-(N+1)} \exp\left[-\frac{y + \varphi(y)}{\bar{\xi}}\right]. \tag{E.6}$$

E.2. The continuous limit

The contributing values for y are finite so that in the continuous limit λ should be close to unity. As a result one can write

$$\left(1 - \frac{y}{\bar{N}\bar{\xi}}\right)^{N+1} = \exp\left[-(N+1)\log\left(1 - \frac{y}{\bar{N}\bar{\xi}}\right)\right] = \exp\left[-\frac{N}{\bar{N}\bar{\xi}}y\right]. \tag{E.7}$$

It implies that

$$P(\rho) d\rho = -\frac{d\rho}{\bar{\xi}} \int_{-i\infty}^{+i\infty} \frac{dy}{2\pi i} \exp\left[-\frac{y + \varphi(y)}{\bar{\xi}} + \frac{y\rho}{\bar{\xi}}\right]. \tag{E.8}$$

This is the inverse Laplace transform.

It is important to note that the counts-in-cells $P(N)$ can be recovered by a Poisson convolution of the continuous distribution. A Poisson distribution is given by

$$P_{\text{Poisson}}(N, \bar{N}) = \frac{\bar{N}^N}{N!} e^{-\bar{N}} = \oint \frac{1}{2\pi i} \frac{d\lambda}{\lambda^{N+1}} \exp(-\bar{N}(1 - \lambda)). \tag{E.9}$$

Then

$$\begin{aligned} &\int d\rho P(\rho) P_{\text{Poisson}}(N, \bar{N}\rho) \\ &= - \int \frac{d\rho}{\bar{\xi}} \int_{-i\infty}^{+i\infty} \frac{dy}{2\pi i} \oint \frac{d\lambda}{2\pi i} \frac{1}{\lambda^{N+1}} \exp\left[-\frac{y + \varphi(y)}{\bar{\xi}} + \frac{y\rho}{\bar{\xi}} - \bar{N}\rho(1 - \lambda)\right]. \end{aligned} \tag{E.10}$$

The integration over ρ leads to $\delta_D(y - \bar{N}\bar{\xi}(1 - \lambda))$, which finally implies

$$\int d\rho P(\rho) P_{\text{Poisson}}(N, \bar{N}\rho) = P(N). \tag{E.11}$$

This is not surprising since we assumed from the very beginning that any discrete field would be the Poisson realization of a continuous field.

E.3. Approximate forms for $P(\rho)$ when $\bar{\xi} \ll 1$

In this section, we review the various approximations that have been used for $P(\rho)$. It obviously depends on the regime we are interested in, that is on the amplitude of the density fluctuations $\bar{\xi}$.

When $\bar{\xi}$ is small, it is possible to apply a saddle-point approximation. This point is defined by

$$\rho_s = \frac{d\varphi(y_s)}{dy}. \tag{E.12}$$

It leads to

$$P(\rho) = \frac{1}{\sqrt{-2\pi\bar{\xi}\varphi''(y_s)}} \exp \left[-\frac{1}{\bar{\xi}}(y_s + \varphi(y_s) - y_s\varphi'(y_s)) \right] . \tag{E.13}$$

In case $\varphi(y)$ is obtained through a tree summation, as for the weakly non-linear regime, one finally gets formula (312).

Obviously, such a result makes sense only if $\varphi''(y)$ is negative. Because of the presence of a singular point on the real axis this will not be always the case. In practice, it will be true only for values of the density smaller than a critical value, ρ_c . These values are given in Table 9 for the results obtained in the quasi-linear regime.

For $\rho > \rho_c$, the shape the saddle-point position is pushed toward the singularity. The behavior of the PDF will then be dominated by the behavior of $\varphi(y)$ around this point. Let us write generally $\varphi(y)$ as

$$\varphi(y) = \varphi_s + r_s(y - y_s) + \dots - a_s(y - y_s)^{\omega_s} , \tag{E.14}$$

where the expansion around the singular point has been decomposed into its regular part $\varphi_s + r_s(y - y_s) + \dots$ and singular part $a_s(y - y_s)^{\omega_s}$, where ω_s is a non-integer value ($\omega_s = \frac{3}{2}$ in the quasi-linear theory). In (E.8) the integration path for y will be pushed toward the negative part of the real axis ($y < y_s$). It can thus be described by the real variable u varying from 0 to ∞ with

$$y = y_s + ue^{\pm i\pi} , \tag{E.15}$$

where the sign changes according to whether y is above or under the real axis. Expanding the singular part in the exponential, one gets

$$P(\rho) = \frac{-a_s}{\bar{\xi}^2} \int_0^\infty du u^{\omega_s} \frac{e^{\pm i\pi(\omega_s-1)}}{2\pi i} \exp \left(-\frac{\rho - r_s}{\bar{\xi}} u \right) , \tag{E.16}$$

which gives

$$P(\rho) = \frac{a_s}{\Gamma(-\omega_s)\bar{\xi}^2} \left(\frac{\rho - r_s}{\bar{\xi}} \right)^{-\omega_s-1} \exp \left(-\frac{\varphi_s}{\bar{\xi}} - |y_s| \frac{\rho}{\bar{\xi}} \right) \tag{E.17}$$

taking advantage of the relation $\Gamma(\omega_s + 1)\Gamma(-\omega_s) = -\sin(\pi\omega_s)/\pi$. For the parameters describing the quasi-linear theory one gets relation (314).

E.4. Approximate forms for $P(\rho)$ when $\bar{\xi} \gg 1$

Two scaling domains have been found (see [16] for a comprehensive presentation of the scaling laws). One corresponds to the rather dense regions. It corresponds to cases where $\varphi(y)$ is always finite in (E.8) For large values of $\bar{\xi}$, it is therefore possible to write

$$P(\rho) = \frac{-1}{\bar{\xi}^2} \int_{-i\infty}^{+i\infty} \frac{dy}{2\pi i} \varphi(y) \exp(xy) \quad \text{with } x = \rho/\bar{\xi} . \tag{E.18}$$

One can see that the PDF is a function of x only. Roughly speaking, in this integral, $y \sim 1/x$ so that the validity domain of this expression is limited to cases where $\varphi(1/x) \ll \bar{\xi}$. It will be limited to a regime where

$$x \gg (\bar{\xi}/a)^{1/(1-\omega)}, \tag{E.19}$$

if $\varphi(y)$ behaves like $ay^{1-\omega}$ at large y .

If x is small, in a regime where $\varphi(y)$ can be approximated by its power-law asymptotic shape, the PDF eventually reads

$$P(\rho) = \frac{1}{\bar{\xi}^2} \frac{a(1-\omega)}{\Gamma(\omega)} x^{2-\omega}. \tag{E.20}$$

For large values of x , one recovers the exponential cut-off found in the previous regime, (E.17), with further simplifications since $\bar{\xi} \gg 1$:

$$P(\rho) = \frac{a_s}{\Gamma(-\omega_s)\bar{\xi}^2} \left(\frac{\rho}{\bar{\xi}}\right)^{-\omega_s-1} \exp(-|y_s|\rho/\bar{\xi}). \tag{E.21}$$

The second scaling regime corresponds to the underdense regions. They are described by the asymptotic form of $\varphi(y)$, which implies

$$P(\rho) = \frac{-1}{\bar{\xi}} \int_{-i\infty}^{+i\infty} \frac{dy}{2\pi i} \exp\left(-a\frac{y^{1-\omega}}{\bar{\xi}} + \frac{\rho y}{\bar{\xi}}\right). \tag{E.22}$$

A simple change of variable, $t^{1-\omega} = y^{1-\omega}\bar{\xi}/a$, shows that it can be written as

$$P(\rho) = \frac{-1}{\bar{\xi}} \left(\frac{a}{\bar{\xi}}\right)^{-1/(1-\omega)} \int_{-i\infty}^{+i\infty} \frac{dt}{2\pi i} \exp(-t^{1-\omega} + zt) \tag{E.23}$$

with

$$z = \frac{\rho}{\rho_v}, \quad \rho_v = \bar{\xi} \left(\frac{a}{\bar{\xi}}\right)^{1/(1-\omega)}, \tag{E.24}$$

which can be written

$$P(\rho) = \frac{1}{\pi\rho_v} \int_0^\infty du \sin[u^{1-\omega} \sin \pi u] e^{-zu + u^{1-\omega} \cos \pi u}. \tag{E.25}$$

For large values of z , the power-law behavior of (E.19) is recovered, and the two regimes overlap.

Small values of z however describe the small-density cut-off. The expression of the PDF can be obtained by a saddle-point approximation, and it appears to be a particular case of the results obtained in Eq. (E.13). Note that the shape of this cut-off depends only on ω ,

$$P(\rho) = \frac{1}{\rho_v} \frac{(1-\omega)^{1/2\omega}}{\sqrt{2\pi\omega}} z^{-1/2-1/2\omega} \exp[-\omega(1-\omega)^{(1-\omega)/\omega} z^{(1-\omega)\omega}]. \tag{E.26}$$

E.5. Numerical computation of the Laplace inverse transform

The starting point of the numerical computation of the local density PDF from the cumulant generating function is Eq. (E.8). In case the cumulant generating function can be obtained from

a vertex generating function \mathcal{G} , the latter is the natural variable to use. The technical difficulty is actually to choose the path to follow in the y or \mathcal{G} complex plane. The original path for y runs from $-i\infty$ to $+i\infty$ along the imaginary axis. But as the functions $\tau(y)$ or $\varphi(y)$ are not analytic over the complex plane (there is at least one singularity on the local axis for $y = y_s < 0$) the crossing point of the path with the real axis cannot be moved to the left side of y_s (otherwise, the PDF would simply vanish!). Actually the crossing point of the path for the numerical integration is the position of the saddle point, y_{saddle} defined by

$$0 = \rho - 1 - \left. \frac{d\varphi(y)}{dy} \right|_{y=y_{\text{saddle}}} . \quad (\text{E.27})$$

This equation has a solution as long as $\rho < 1 + \delta_c$ and it is then at a point $y_{\text{saddle}} > y_s$ (see Section 5.8). In the case of $\rho > 1 + \delta_c$, the crossing point of the integration path is then simply chosen to be $y = y_s$. The integration path is subsequently built in such a way that $\rho y - 1 - \varphi(y)$ is kept real and negative to avoid unnecessary oscillation of the function to integrate. In practice, the path is built step by step with an adaptive integration scheme [44,151].

Appendix F. Cosmic errors: expressions for the factorial moments

In this appendix, we first explain how the cosmic error on the factorial moments of count-in-cells is calculated. We then list the corresponding analytic expression for the cosmic covariance matrix up to third order in the 3D case.

F.1. Method

From now, to simplify we assume that the cells are spherical (or circular, in two dimensions), but the results are valid in practice with the obvious appropriate corrections for any compact cell.

The local Poisson assumption allows us to neglect correlations inside the union C_{\cup} of volume v_{\cup} of two overlapping cells and the non-spherical contribution of C_{\cup} . As a result, the generating function for bicounts in overlapping cells reads [621]

$$\mathcal{P}_{\text{over}}(x, y) = \mathcal{P}_{\cup}[q(x + y) + pxy] . \quad (\text{F.1})$$

The generating function $\mathcal{P}_{\cup}(x)$ is the same as $\mathcal{P}(x)$ but for a cell of volume v_{\cup} , and

$$p = [1 - f_{\mathcal{Q}}(r/R)]/[1 + f_{\mathcal{Q}}(r/R)], \quad q = f_{\mathcal{Q}}(r/R)/[1 + f_{\mathcal{Q}}(r/R)] , \quad (\text{F.2})$$

where $f_{\mathcal{Q}}(r/R)$ represents the excess of volume (or area) of v_{\cup} compared to v_R ,

$$v_{\cup} = v_R[1 + f_{\mathcal{Q}}(r/R)] , \quad (\text{F.3})$$

and r is the separation between the two cells. We have $f_3(\psi) = (\frac{3}{4})\psi - (\frac{1}{16})\psi^3$, and $f_2(\psi) = 1 - (1/\pi)[2\arccos(\psi/2) - \sqrt{1 - \psi^2/2}]$ in three and two dimensions, respectively.

The generating function for disjoint cells is Taylor expanded:

$$\mathcal{P}_{\text{disjoint}}(x, y) \simeq \mathcal{P}(x)\mathcal{P}(y)[1 + \mathcal{R}(x, y)] + O(\xi/\bar{\xi}^2) \quad (\text{F.4})$$

with

$$\mathcal{R}(x, y) = \xi \sum_{M=1, N=1}^{\infty} (x-1)^M (y-1)^N \frac{S_{NM}}{N!M!} \bar{N}^{N+M} \bar{\xi}^{N+M-2}. \quad (\text{F.5})$$

It is then easy to calculate cross-correlations on factorial moments, $\Delta_{k,l}$, by computing the double integral in Eq. (449) after applying partial derivatives in Eq. (447), with the further assumption that the two-point correlation function is well approximated by a power law of index $-\gamma \simeq -1.8$ for $r \leq 2R$.¹⁵⁰

F.2. Analytic results

The cosmic errors for the factorial moments as discussed in Section 6.7.4, Eq. (451), are now detailed here, up to third order (in the 3D case):

$$\Delta_{11}^F = \bar{N}^2 \bar{\xi}(\hat{L}), \quad (\text{F.6})$$

$$\Delta_{11}^E = 5.508 \bar{N}^2 \bar{\xi} \frac{v}{V}, \quad (\text{F.7})$$

$$\Delta_{11}^D = \bar{N} \frac{v}{V}, \quad (\text{F.8})$$

$$\Delta_{22}^F = 4\bar{N}^4 \bar{\xi}(\hat{L}) (1 + 2\bar{\xi} Q_{12} + \bar{\xi}^2 Q_{22}), \quad (\text{F.9})$$

$$\Delta_{22}^E = \bar{N}^4 \bar{\xi} \frac{v}{V} (17.05 + 3.417\bar{\xi} + 45.67\bar{\xi} Q_3 + 42.24\bar{\xi}^2 Q_4), \quad (\text{F.10})$$

$$\Delta_{22}^D = \bar{N}^2 \frac{v}{V} (0.648 + 4\bar{N} + 0.502\bar{\xi} + 8.871\bar{N}\bar{\xi} + 6.598\bar{N}^2 \bar{\xi}^2 Q_3), \quad (\text{F.11})$$

$$\begin{aligned} \Delta_{33}^F = 9\bar{N}^6 \bar{\xi}(\hat{L}) (1 + 2\bar{\xi} + \bar{\xi}^2 + 4\bar{\xi} Q_{12} + 4\bar{\xi}^2 Q_{12} + 6\bar{\xi}^2 Q_{13} \\ + 6\bar{\xi}^3 Q_{13} + 4\bar{\xi}^2 Q_{22} + 12\bar{\xi}^3 Q_{23} + 9\bar{\xi}^4 Q_{33}), \end{aligned} \quad (\text{F.12})$$

$$\begin{aligned} \Delta_{33}^E = \bar{N}^6 \bar{\xi} \frac{v}{V} (34.62 + 99.26\bar{\xi} + 39.60\bar{\xi}^2 + 180.3\bar{\xi} Q_3 + 331.1\bar{\xi}^2 Q_3 \\ + 93.50\bar{\xi}^3 Q_3^2 + 633.5\bar{\xi}^2 Q_4 + 441.3\bar{\xi}^3 Q_4 \\ + 1379\bar{\xi}^3 Q_5 + 1668\bar{\xi}^4 Q_6), \end{aligned} \quad (\text{F.13})$$

$$\begin{aligned} \Delta_{33}^D = \bar{N}^3 \frac{v}{V} (0.879 + 5.829\bar{N} + 9\bar{N}^2 + 2.116\bar{\xi} + 27.13\bar{N}\bar{\xi} \\ + 66.53\bar{N}^2 \bar{\xi} + 10.59\bar{N}\bar{\xi}^2 + 74.23\bar{N}^2 \bar{\xi}^2 + 1.709\bar{\xi}^2 Q_3 \end{aligned}$$

¹⁵⁰ The results do not depend significantly on the value of γ [621].

$$\begin{aligned}
& + 42.37\bar{N}\bar{\xi}^2 Q_3 + 148.5\bar{N}^2\bar{\xi}^2 Q_3 + 111.2\bar{N}^2\bar{\xi}^3 Q_3 \\
& + 44.40\bar{N}\bar{\xi}^3 Q_4 + 296.4\bar{N}^2\bar{\xi}^3 Q_4 + 349.3\bar{N}^2\bar{\xi}^4 Q_5) .
\end{aligned} \tag{F.14}$$

The cosmic cross-correlations read

$$A_{12}^F = 2\bar{N}^3\bar{\xi}(\hat{L})(1 + \bar{\xi}Q_{12}) , \tag{F.15}$$

$$A_{12}^E = \bar{N}^3\bar{\xi}\frac{v}{V}(8.525 + 11.42\bar{\xi}Q_3) , \tag{F.16}$$

$$A_{12}^D = \bar{N}^2\frac{v}{V}(2.0 + 1.478\bar{\xi}) , \tag{F.17}$$

$$A_{13}^F = 3\bar{N}^4\bar{\xi}(\hat{L})(1 + \bar{\xi} + 2\bar{\xi}Q_{12} + 3\bar{\xi}^2Q_{13}) , \tag{F.18}$$

$$A_{13}^E = \bar{N}^4\bar{\xi}\frac{v}{V}(9.05 + 11.42\bar{\xi} + 21.67\bar{\xi}Q_3 + 42.24\bar{\xi}^2Q_4) , \tag{F.19}$$

$$A_{13}^D = \bar{N}^3\frac{v}{V}(3.0 + 6.653\bar{\xi} + 4.949\bar{\xi}^2Q_3) , \tag{F.20}$$

$$A_{23}^F = 6\bar{N}^5\bar{\xi}(\hat{L})(1 + \bar{\xi} + 3\bar{\xi}Q_{12} + 3\bar{\xi}^2Q_{13} + \bar{\xi}^2Q_{12} + 2\bar{\xi}^2Q_{22} + 3\bar{\xi}^3Q_{23}) , \tag{F.21}$$

$$A_{23}^E = \bar{N}^5\bar{\xi}\frac{v}{V}(23.08 + 33.09\bar{\xi} + 90.17\bar{\xi}Q_3 + 55.19\bar{\xi}^2Q_3 + 211.2\bar{\xi}^2Q_4 + 229.9\bar{\xi}^3Q_5) , \tag{F.22}$$

$$\begin{aligned}
A_{23}^D = \bar{N}^3\frac{v}{V}(1.943 + 6\bar{N} + 4.522\bar{\xi} + 26.61\bar{N}\bar{\xi} + 9.898\bar{N}\bar{\xi}^2 \\
+ 3.531\bar{\xi}^2Q_3 + 39.59\bar{N}\bar{\xi}^2Q_3 + 39.53\bar{N}\bar{\xi}^3Q_4) .
\end{aligned} \tag{F.23}$$

Note that the finite-volume effect terms A_{pq}^F would be the same in the 2D case. In the above equations, $\bar{\xi}(\hat{L})$ is the integral of the two-point correlation function over the survey volume [Eq. (452)] and

$$Q_N \equiv \frac{S_N}{N^{N-2}} , \quad Q_{NM} \equiv \frac{C_{NM}}{N^{N-1}M^{M-1}} . \tag{F.24}$$

Note that these Q_N and Q_{NM} are slightly different from what was defined in Eqs. (150) and (214). They are also often used in the literature instead of S_p or C_{pq} .

An accurate approximation for $\bar{\xi}(\hat{L})$ is [153,154]

$$\bar{\xi}(\hat{L}) \simeq \frac{1}{\bar{V}^2} \int_{\bar{V}} d^{\mathcal{D}}r_1 d^{\mathcal{D}}r_2 \zeta(r_{12}) - \frac{1}{\bar{V}} \int_{r \leq 2R} d^{\mathcal{D}}r \zeta(r) . \tag{F.25}$$

This actually means that, rigorously, the finite-volume error as we defined it here actually contains an edge effect term. For practical calculations, however, the following approximation generally works quite well:

$$\bar{\xi}(\hat{L}) \simeq \bar{\xi}(L) , \tag{F.26}$$

where $\bar{\xi}(L)$ was defined in Eq. (389).

References

- [1] S.J. Aarseth, E.L. Turner, J.R. Gott, *Astrophys. J.* 228 (1979) 664–683.
- [2] K.L. Adelberger, C.C. Steidel, M. Giavalisco, M. Dickinson, M. Pettini, M. Kellogg, *Astrophys. J.* 505 (1998) 18–24.
- [3] R.J. Adler, *The Geometry of Random Fields*, Wiley, New York, 1981.
- [4] C. Alcock, B. Paczyński, *Nature* 281 (1979) 358–359.
- [5] L. Amendola, *Mon. Not. R. Astron. Soc.* 283 (1996) 983–989.
- [6] P. Anninos, M.L. Norman, D.A. Clarke, *Astrophys. J.* 436 (1994) 11–22.
- [7] I. Antoniadis, P.O. Mazur, E. Mottola, *Phys. Rev. Lett.* 79 (1997) 14–17.
- [8] A.W. Appel, *SIAM, J. Sci. Stat. Comput.* 6 (1985) 85–103.
- [9] P.P. Avelino, E.P.S. Shellard, J.H.P. Wu, B. Allen, *Astrophys. J.* 507 (1998) L101–L104.
- [10] A. Babul, D.H. Weinberg, A. Dekel, J.P. Ostriker, *Astrophys. J.* 427 (1994) 1–24.
- [11] D. Bacon, A. Refregier, R. Ellis, *Mon. Not. R. Astron. Soc.* 318 (2000) 625–640.
- [12] J.S. Bagla, *astro-ph/9911025*, 1999.
- [13] J.S. Bagla, T. Padmanabhan, *Mon. Not. R. Astron. Soc.* 266 (1994) 227–237.
- [14] J.S. Bagla, T. Padmanabhan, *Mon. Not. R. Astron. Soc.* 286 (1997) 1023–1031.
- [15] N.A. Bahcall, R.M. Soneira, *Astrophys. J.* 270 (1983) 20–38.
- [16] R. Balian, R. Schaeffer, *Astron. Astrophys.* 220 (1989) 1–29.
- [17] R. Balian, R. Schaeffer, *Astron. Astrophys.* 226 (1989) 373–414.
- [18] W.E. Ballinger, J.A. Peacock, A.F. Heavens, *Mon. Not. R. Astron. Soc.* 282 (1996) 877–888.
- [19] J.M. Bardeen, *Phys. Rev. D* 22 (1980) 1882–1905.
- [20] J.M. Bardeen, P.J. Steinhardt, M.S. Turner, *Phys. Rev. D* 28 (1983) 679–693.
- [21] J.M. Bardeen, J.R. Bond, N. Kaiser, A.S. Szalay, *Astrophys. J.* 304 (1986) 15–61.
- [22] J. Barnes, P. Hut, *Nature* 324 (1986) 446–449.
- [23] J. Barriga, E. Gaztañaga, *Mon. Not. R. Astron. Soc.* 333 (2002) 443–453.
- [24] J.D. Barrow, S.P. Bhavsar, D.H. Sonoda, *Mon. Not. R. Astron. Soc.* 210 (1984) 19P–23P.
- [25] M. Bartelmann, P. Schneider, *Phys. Rep.* 340 (2001) 291–472.
- [26] C.M. Baugh, G. Efstathiou, *Mon. Not. R. Astron. Soc.* 265 (1993) 145–156.
- [27] C.M. Baugh, G. Efstathiou, *Mon. Not. R. Astron. Soc.* 270 (1994) 183–198.
- [28] C.M. Baugh, E. Gaztañaga, G. Efstathiou, *Mon. Not. R. Astron. Soc.* 274 (1995) 1049–1070.
- [29] C.M. Baugh, *Mon. Not. R. Astron. Soc.* 280 (1996) 267–275.
- [30] C.M. Baugh, E. Gaztañaga, *Mon. Not. R. Astron. Soc.* 280 (1996) L37–L41.
- [31] D.J. Baumgart, J.N. Fry, *Astrophys. J.* 375 (1991) 25–34.
- [32] A.J. Bean, R.S. Ellis, T. Shanks, G. Efstathiou, B.A. Peterson, *Mon. Not. R. Astron. Soc.* 205 (1983) 605–624.
- [33] R.H. Becker, R.L. White, D.J. Helfand, *Astrophys. J.* 450 (1995) 559–577.
- [34] K. Benabed, F. Bernardeau, *Phys. Rev. D* 64 (2001) 083501.
- [35] C. Benoist, A. Cappi, L.N. da Costa, S. Maurogordato, F.R. Bouchet, R. Schaeffer, *Astrophys. J.* 514 (1999) 563–578.
- [36] A.J. Benson, S. Cole, C.S. Frenk, C.M. Baugh, C.G. Lacey, *Mon. Not. R. Astron. Soc.* 311 (2000) 793–808.
- [37] A.J. Benson, *Mon. Not. R. Astron. Soc.* 325 (2001) 1039–1044.
- [38] A.A. Berlind, V.K. Narayanan, D.H. Weinberg, *Astrophys. J.* 549 (2001) 688–701.
- [39] A.A. Berlind, D.H. Weinberg, *astro-ph/0109001*, 2001.
- [40] F. Bernardeau, R. Schaeffer, *Astron. Astrophys.* 250 (1991) 23–42.
- [41] F. Bernardeau, R. Schaeffer, *Astron. Astrophys.* 255 (1992) 1–25.
- [42] F. Bernardeau, *Astrophys. J.* 390 (1992) L61–L64.
- [43] F. Bernardeau, *Astrophys. J.* 392 (1992) 1–14.
- [44] F. Bernardeau, *Astron. Astrophys.* 291 (1994) 697–712.
- [45] F. Bernardeau, *Astrophys. J.* 427 (1994) 51–71.
- [46] F. Bernardeau, *Astrophys. J.* 433 (1994) 1–18.
- [47] F. Bernardeau, T.P. Singh, B. Banerjee, S.M. Chitre, *Mon. Not. R. Astron. Soc.* 269 (1994) 947–952.

- [48] F. Bernardeau, *Astron. Astrophys.* 301 (1995) 309–317.
- [49] F. Bernardeau, L. Kofman, *Astrophys. J.* 443 (1995) 479–498.
- [50] F. Bernardeau, R. Juszkiewicz, A. Dekel, F.R. Bouchet, *Mon. Not. R. Astron. Soc.* 274 (1995) 20–26.
- [51] F. Bernardeau, *Astron. Astrophys.* 312 (1996) 11–23.
- [52] F. Bernardeau, R. Van De Weygaert, *Mon. Not. R. Astron. Soc.* 279 (1996) 693–711.
- [53] F. Bernardeau, L. Van Waerbeke, Y. Mellier, *Astron. Astrophys.* 322 (1997) 1–18.
- [54] F. Bernardeau, R. Van De Weygaert, E. Hivon, F.R. Bouchet, *Mon. Not. R. Astron. Soc.* 290 (1997) 566–576.
- [55] F. Bernardeau, *Astron. Astrophys.* 338 (1998) 375–382.
- [56] F. Bernardeau, M.J. Chodorowski, E.L. Łokas, R. Stompor, A. Kudlicki, *Mon. Not. R. Astron. Soc.* 309 (1999) 543–555.
- [57] F. Bernardeau, R. Schaeffer, *Astron. Astrophys.* 349 (1999) 697–728.
- [58] F. Bernardeau, P. Valageas, *Astron. Astrophys.* 364 (2000) 1–16.
- [59] G.M. Bernstein, *Astrophys. J.* 424 (1994) 569–577.
- [60] E. Bertschinger, *Astrophys. J.* 58 (Suppl.) (1985) 39–66.
- [61] E. Bertschinger, in: V.J. Martinez, M. Portilla, D. Saez (Eds.), *New Insights into the Universe*, Springer, Berlin, 1992, pp. 65–126.
- [62] E. Bertschinger, B. Jain, *Astrophys. J.* 431 (1994) 486–494.
- [63] E. Bertschinger, *Annu. Rev. Astron. Astrophys.* 36 (1998) 599–654.
- [64] E. Bertschinger, *Astrophys. J. Suppl.* 137 (2001) 1–20.
- [65] P. Binetruy, C. Deffayet, E. Dudas, P. Ramond, *Phys. Lett. B* 441 (1998) 163–172.
- [66] P. Binetruy, C. Deffayet, P. Peter, *Phys. Lett. B* 441 (1998) 52–59.
- [67] J.J. Binney, N.J. Dowrick, A.J. Fisher, M.E.J. Newman, *The Theory of Critical Phenomena*, Oxford University Press, Oxford, 1992.
- [68] A. Blanchard, J.-M. Alimi, *Astron. Astrophys.* 203 (1988) L1–L4.
- [69] R.D. Blandford, *Q. J. R. Astron. Soc.* 31 (1990) 305–331.
- [70] R.D. Blandford, A.B. Saust, T.G. Brainerd, J.V. Villumsen, *Mon. Not. R. Astron. Soc.* 251 (1991) 600–627.
- [71] M.R. Blanton, *Astrophys. J.* 544 (2000) 63–80.
- [72] M.R. Blanton, R. Cen, J. Ostriker, M. Strauss, *Astrophys. J.* 522 (1999) 590–603.
- [73] M.R. Blanton, R. Cen, J. Ostriker, M. Strauss, M. Tegmark, *Astrophys. J.* 531 (2000) 1–16.
- [74] M.R. Blanton, R.H. Lupton, F.M. Maley, N. Young, I. Zehavi, J. Loveday, *astro-ph/0105535*, 2001.
- [75] G.R. Blumenthal, S.M. Faber, J.R. Primack, M.J. Rees, *Nature* 311 (1984) 517–525.
- [76] J.R. Bond, G. Efstathiou, *Astrophys. J.* 285 (1984) L45–L48.
- [77] J.R. Bond, H.M.P. Couchman, in: *The Post-Recombination Universe*, Proceedings of the NATO Advanced Study Institute, Kluwer Academic Publishers, Dordrecht, 1988, pp. 263–265.
- [78] J.R. Bond, S. Cole, N. Kaiser, G. Efstathiou, *Astrophys. J.* 379 (1991) 440–460.
- [79] J.R. Bond, *Phys. Rev. Lett.* 74 (1995) 4369–4372.
- [80] J.R. Bond, A.H. Jaffe, L. Knox, *Phys. Rev. D* 57 (1998) 2117–2137.
- [81] J.R. Bond, A.H. Jaffe, L. Knox, *Astrophys. J.* 533 (2000) 19–37.
- [82] J.R. Bond, D. Pogosyan, T. Souradeep, *Phys. Rev. D* 62 (2000) 043005.
- [83] S.A. Bonometto, N.A. Sharp, *Astron. Astrophys.* 92 (1980) 222–224.
- [84] S.A. Bonometto, S. Borgani, S. Ghigna, A. Klypin, J.R. Primack, *Mon. Not. R. Astron. Soc.* 273 (1995) 101–121.
- [85] P. Boschán, I. Szapudi, A.S. Szalay, *Astrophys. J.* 93 (Suppl.) (1994) 65–71.
- [86] F.R. Bouchet, J.-C. Adam, R. Pellat, *Astron. Astrophys.* 144 (1995) 413–426.
- [87] F.R. Bouchet, R. Schaeffer, M. Davis, *Astrophys. J.* 383 (1991) 19–40.
- [88] F.R. Bouchet, M. Davis, M. Strauss, in: G.A. Mamon, D. Gerbal (Eds.), *Distribution of Matter in the Universe*, Obs. de Paris, Meudon, 1992, pp. 287–300.
- [89] F.R. Bouchet, L. Hernquist, *Astrophys. J.* 68 (Suppl.) (1988) 521–538.
- [90] F.R. Bouchet, D.P. Bennett, *Phys. Rev. D* 41 (1990) 2408–2433.
- [91] F.R. Bouchet, R. Juszkiewicz, S. Colombi, R. Pellat, *Astrophys. J.* 394 (1992) L5–L8.
- [92] F.R. Bouchet, M.A. Strauss, M. Davis, K.B. Fisher, A. Yahil, J.P. Huchra, *Astrophys. J.* 417 (1993) 36–53.
- [93] F.R. Bouchet, S. Colombi, E. Hivon, R. Juszkiewicz, *Astron. Astrophys.* 296 (1995) 575–608.

- [94] F.R. Bouchet, in: S. Bonometto, J.R. Primack, A. Provenzale (Eds.), *Proceedings of the International School Enrico Fermi, Course CXXXII (Dark Matter in the Universe)*, Varenna, Italy, IOS Press, Amsterdam.
- [95] R.J. Bower, *Mon. Not. R. Astron. Soc.* 248 (1991) 332–352.
- [96] R.G. Bower, P. Coles, C.S. Frenk, S.D.M. White, *Mon. Not. R. Astron. Soc.* 405 (1993) 403–412.
- [97] T.G. Brainerd, R.J. Scherrer, J.V. Villumsen, *Astrophys. J.* 418 (1993) 570–578.
- [98] C. Brans, R.H. Dicke, *Phys. Rep.* 124 (1961) 925–945.
- [99] T.J. Broadhurst, R.S. Ellis, D.C. Koo, A.S. Szalay, *Nature* 343 (1990) 726–728.
- [100] A. Buchalter, M. Kamionkowski, *Astrophys. J.* 521 (1999) 1–16.
- [101] A. Buchalter, M. Kamionkowski, A.H. Jaffe, *Astrophys. J.* 530 (2000) 36–52.
- [102] T. Buchert, *Astron. Astrophys.* 223 (1989) 9–24.
- [103] T. Buchert, *Mon. Not. R. Astron. Soc.* 254 (1992) 729–737.
- [104] T. Buchert, J. Ehlers, *Mon. Not. R. Astron. Soc.* 264 (1993) 375–387.
- [105] T. Buchert, *Mon. Not. R. Astron. Soc.* 267 (1994) 811–820.
- [106] T. Buchert, A.L. Melott, A.G. Weiß, *Astron. Astrophys.* 288 (1994) 349–364.
- [107] T. Buchert, in: S. Bonometto, J.R. Primack, A. Provenzale (Eds.), *Proceedings of the International School Enrico Fermi, Course CXXXII (Dark Matter in the Universe)*, Varenna, Italy, IOS Press, Amsterdam, 1996, pp. 543–564.
- [108] T. Buchert, G. Karakatsanis, R. Klaffl, P. Shiller, *Astron. Astrophys.* 318 (1997) 1–10.
- [109] T. Buchert, A. Dominguez, *Astron. Astrophys.* 335 (1999) 395–402.
- [110] J.S. Bullock, R.H. Wechsler, R.S. Somerville, *Mon. Not. R. Astron. Soc.* 329 (2002) 246–256.
- [111] A. Buonanno, K.A. Meissner, C. Ungarelli, G. Veneziano, *J. High Energy Phys.* 01 (1998) 004.
- [112] R.R. Caldwell, R. Juszkiewicz, P.J. Steinhardt, F.R. Bouchet, *Astrophys. J.* 547 (2001) L93–L96.
- [113] R.G. Carlberg, H.K.C. Yee, S.L. Morris, H. Lin, M. Sawicki, G. Wirth, D. Patton, C.W. Shepherd, E. Ellingson, D. Schade, C.J. Pritchett, F.D.A. Hartwick, *Phil. Trans. R. Soc. London A* 357 (1999) 167–183.
- [114] S. Carroll, W. Press, E.L. Turner, *Annu. Rev. Astron. Astrophys.* 30 (1992) 499–542.
- [115] E. Carretti, C. Bertoni, A. Messina, E. Zucca, L. Guzzo, *Mon. Not. R. Astron. Soc.* 324 (2001) 1029–1040.
- [116] P. Catelan, L. Moscardini, *Astrophys. J.* 426 (1994) 14–18.
- [117] P. Catelan, *Mon. Not. R. Astron. Soc.* 276 (1995) 115–124.
- [118] P. Catelan, F. Lucchin, S. Matarrese, L. Moscardini, *Mon. Not. R. Astron. Soc.* 276 (1995) 39–56.
- [119] P. Catelan, F. Lucchin, S. Matarrese, C. Porciani, *Mon. Not. R. Astron. Soc.* 297 (1998) 692–712.
- [120] P. Catelan, C. Porciani, M. Kamionkowski, *Mon. Not. R. Astron. Soc.* 318 (2000) L39–L44.
- [121] P. Catelan, M. Kamionkowski, R. Blandford, *Mon. Not. R. Astron. Soc.* 320 (2001) L7–L13.
- [122] R. Cen, J. Ostriker, *Astrophys. J.* 538 (2000) 83–91.
- [123] T. Chester, C. Beichmann, T. Conrow, *Revised IRAS Explanatory Supplement, 1987 (Chapter XII)*.
- [124] M. Chodorowski, F.R. Bouchet, *Mon. Not. R. Astron. Soc.* 279 (1996) 557–563.
- [125] M. Chodorowski, E.L. Łokas, *Mon. Not. R. Astron. Soc.* 287 (1997) 591–606.
- [126] M. Chodorowski, *Mon. Not. R. Astron. Soc.* 292 (1997) 695–702.
- [127] M. Chodorowski, E.L. Łokas, A. Pollo, A. Nusser, *Mon. Not. R. Astron. Soc.* 300 (1998) 1027–1034.
- [128] M. Chodorowski, *Mon. Not. R. Astron. Soc.* 308 (1999) 640–650.
- [129] M. Chodorowski, *Mon. Not. R. Astron. Soc.* 318 (2000) 375–380.
- [130] A.L. Coil, M. Davis, I. Szapudi, *Publ. Astron. Soc. Pac.* 113 (2001) 1312–1325.
- [131] S. Cole, K.B. Fisher, D.H. Weinberg, *Mon. Not. R. Astron. Soc.* 267 (1994) 785–799.
- [132] S. Cole, K.B. Fisher, D.H. Weinberg, *Mon. Not. R. Astron. Soc.* 275 (1995) 515–526.
- [133] S. Cole, S. Hatton, D.H. Weinberg, C.S. Frenk, *Mon. Not. R. Astron. Soc.* 300 (1998) 945–966.
- [134] S. Cole, C. Lacey, C.M. Baugh, C.S. Frenk, *Mon. Not. R. Astron. Soc.* 319 (2000) 168–204.
- [135] P. Coles, *Mon. Not. R. Astron. Soc.* 243 (1990) 171–176.
- [136] P. Coles, B. Jones, *Mon. Not. R. Astron. Soc.* 248 (1991) 1–13.
- [137] P. Coles, C.S. Frenk, *Mon. Not. R. Astron. Soc.* 253 (1991) 727–737.
- [138] P. Coles, A.L. Melott, S.F. Shandarin, *Mon. Not. R. Astron. Soc.* 260 (1993) 765–776.
- [139] P. Coles, L. Moscardini, F. Lucchin, S. Matarrese, A. Messina, *Mon. Not. R. Astron. Soc.* 264 (1993) 749–757.
- [140] P. Coles, *Mon. Not. R. Astron. Soc.* 262 (1993) 1065–1075.
- [141] P. Colín, A.A. Klypin, A.V. Kravtsov, A.M. Khokhlov, *Astrophys. J.* 523 (1999) 32–53.

- [142] M. Colless, G. Dalton, S. Maddox, W. Sutherland, P. Norberg, S. Cole, et al., *Mon. Not. R. Astron. Soc.* 328 (2001) 1039–1063.
- [143] J.C. Collins, *Renormalization*, Cambridge University Press, Cambridge, 1984.
- [144] C.A. Collins, R.C. Nichol, S.L. Lumsden, *Mon. Not. R. Astron. Soc.* 254 (1990) 295–300.
- [145] S. Colombi, F.R. Bouchet, R. Schaeffer, *Astron. Astrophys.* 263 (1992) 1–22.
- [146] S. Colombi, Ph.D. Thesis, 1993, Université Paris 7.
- [147] S. Colombi, F.R. Bouchet, R. Schaeffer, *Astron. Astrophys.* 281 (1994) 301–313.
- [148] S. Colombi, *Astrophys. J.* 435 (1994) 536–539.
- [149] S. Colombi, F.R. Bouchet, R. Schaeffer, *Astrophys. J.* 96 (Suppl.) (1995) 401–428.
- [150] S. Colombi, F.R. Bouchet, L. Hernquist, *Astrophys. J.* 465 (1996) 14–33.
- [151] S. Colombi, F. Bernardeau, F.R. Bouchet, L. Hernquist, *Mon. Not. R. Astron. Soc.* 287 (1997) 241–252.
- [152] S. Colombi, I. Szapudi, A.S. Szalay, *Mon. Not. R. Astron. Soc.* 296 (1998) 253–274.
- [153] S. Colombi, I. Szapudi, A. Jenkins, J. Colberg, *Mon. Not. R. Astron. Soc.* 313 (2000) 711–724.
- [154] S. Colombi, S. Charlot, J. Devriendt, M. Fioc, I. Szapudi, in: A. Mazure, O. Lefèvre, V. Lebrun (Eds.), *Clustering at High Redshift*, ASP Conference Series, Vol. 200, 2000, pp. 153–157.
- [155] S. Colombi, *New Astron. Rev.* 45 (2001) 373–377.
- [156] J.J. Condon, W.D. Cotton, E.W. Greisen, Q.F. Yin, R.A. Perley, G.B. Taylor, J.J. Broderick, *Astron. J.* 115 (1998) 1693–1716.
- [157] A. Connolly, R. Scranton, D. Johnston, et al., *Astrophys. J.* (2001), submitted for publication, astro-ph/0107417.
- [158] A. Cooray, W. Hu, *Astrophys. J.* 548 (2001) 7–18.
- [159] E.J. Copeland, R. Easther, D. Wands, *Phys. Rev. D* 56 (1997) 874–888.
- [160] S. Courteau, M.A. Strauss, J.A. Willick (Eds.), *Cosmic Flows: Towards an Understanding of Large-Scale Structure*, ASP Conference Series, 2000.
- [161] W.J. Couch, J.S. Jurcevic, B.J. Boyle, *Mon. Not. R. Astron. Soc.* 260 (1993) 241–252.
- [162] H.M.P. Couchman, *Astrophys. J.* 368 (1991) L23–L26.
- [163] H.M.P. Couchman, in: *Computational Astrophysics*, ASP Conference Series, Vol. 123, 1997, p. 340.
- [164] H.M.P. Couchman, P.J.E. Peebles, *Astrophys. J.* 497 (1998) 499–511.
- [165] C.M. Cress, M. Kamionkowski, *Mon. Not. R. Astron. Soc.* 297 (1998) 486–492.
- [166] R.A.C. Croft, C. Metzler, *Astrophys. J.* 545 (2000) 561–571.
- [167] S.M. Croom, R.J. Smith, B.J. Boyle, T. Shanks, N.S. Loaring, L. Miller, I.J. Lewis, astro-ph/0104095, 2001.
- [168] L.N. da Costa, P.S. Pellegrini, M. Davis, A. Meiksin, W.L.W. Sargent, J.L. Tonry, *Astrophys. J.* 75 (Suppl.) (1991) 935–964.
- [169] L.N. da Costa, M.S. Vogeley, M.J. Geller, J.P. Huchra, C. Park, *Astrophys. J.* 437 (1994) L1–L4.
- [170] R. Davé, D.N. Spergel, P.J. Steinhardt, B.D. Wandelt, *Astrophys. J.* 547 (2001) 574–589.
- [171] M. Davis, P.J.E. Peebles, *Astrophys. J.* 34 (Suppl.) (1977) 425–450.
- [172] M. Davis, P.J.E. Peebles, *Astrophys. J.* 267 (1983) 465–482.
- [173] M. Davis, G. Efstathiou, C.S. Frenk, S.D.M. White, *Astrophys. J.* 292 (1985) 371–394.
- [174] M. Davis, G. Efstathiou, C.S. Frenk, S.D.M. White, *Nature* 356 (1992) 489–494.
- [175] M. Davis, A. Meiksin, M.A. Strauss, L.N. da Costa, A. Yahil, *Astrophys. J.* 333 (1988) L9–L12.
- [176] M. Davis, A. Miller, S.D.M. White, *Astrophys. J.* 490 (1997) 63–71.
- [177] M. Davis, J.A. Newman, in: A.J. Banday, S. Zaroubi, M. Bartelmann (Eds.), *Mining the Sky*, ESO/MPA/MPE Astrophysics Symposia, Springer, Heidelberg, 2001.
- [178] A.I. Davydychev, *J. Phys. A* 25 (1992) 5587–5596.
- [179] A. Dekel, *Annu. Rev. Astron. Astrophys.* 32 (1994) 371–418.
- [180] A. Dekel, O. Lahav, *Astrophys. J.* 520 (1999) 24–34.
- [181] A.A. de Laix, G.D. Starkman, *Mon. Not. R. Astron. Soc.* 299 (1998) 977–984.
- [182] V. de Lapparent, M.J. Geller, J.P. Huchra, *Astrophys. J.* 302 (1986) L1–L5.
- [183] V. de Lapparent, M.J. Kurtz, M.J. Geller, *Astrophys. J.* 304 (1986) 585–598.
- [184] G. de Vaucouleurs, *Ann. Astrophys.* 11 (1948) 247.
- [185] J. Des Cloiseaux, G. Jannink, *Les polymères en solution*, Les éditions de physique, Les Ulis, France, 1987.
- [186] A. Diaferio, M.J. Geller, *Astrophys. J.* 467 (1996) 19–37.

- [187] S.G. Djorgovski, R.R. Gal, S.C. Odewahn, R.R. de Carvalho, R. Brunner, G. Longo, R. Scaramella, *Wide Field Surveys in Cosmology*, 14th IAP Meeting, 26–30 May 1998, Editions Frontieres, Paris, p. 89, astro-ph/9809187.
- [188] S. Dodelson, E. Gaztañaga, *Mon. Not. R. Astron. Soc.* 312 (2000) 774–780.
- [189] S. Dodelson, V.K. Narayanan, M. Tegmark, R. Scranton et al., *Astrophys. J.* (2001), submitted for publication, astro-ph/0107421.
- [190] A.G. Doroshkevich, *Astrofizika* 6 (1970) 581.
- [191] A.G. Doroshkevich, E.V. Kotok, I.D. Novikov, A.N. Polyudov, S.F. Shandarin, Yu.S. Sigov, *Mon. Not. R. Astron. Soc.* 192 (1980) 321–327.
- [192] A. Dressler, *Astrophys. J.* 236 (1980) 351–365.
- [193] J. Dubinski, *New Astron.* 1 (1996) 133–147.
- [194] R. Durrer, M. Kunz, A. Melchiorri, *Phys. Rev. D* 59 (1999) 123005.
- [195] R. Durrer, R. Juszkiewicz, M. Kunz, J-P. Uzan, *Phys. Rev. D* 62 (2000) 021301.
- [196] G. Efstathiou, in: R. Schaeffer, J. Silk, M. Spiro, J. Zinn-Justin (Eds.), *Cosmology and Large Scale Structure*, Les Houches Lectures, Session LX, Elsevier Science, Amsterdam, 1996, p. 133.
- [197] G. Efstathiou, in: J.A. Peacock, A.F. Heavens, A.T. Davies (Eds.), *Physics of the Early Universe*, Proceedings of the 36th Scottish Universities Summer School in Physics of the Early Universe, Adam Hilger, Bristol, 1990, pp. 361–463.
- [198] G. Efstathiou, R.I. Jedredjewski, *Adv. Space Res.* 3 (1984) 379–386.
- [199] G. Efstathiou, M. Davis, S.D.M. White, C. S. Frenk, *Astrophys. J.* 57 (Suppl.) (1985) 241–260.
- [200] G. Efstathiou, N. Kaiser, W. Saunders, A. Lawrence, M. Rowan-Robinson, R.S. Ellis, C.S. Frenk, *Mon. Not. R. Astron. Soc.* 247 (1990) 10P–14P.
- [201] G. Efstathiou, J.R. Bond, S.D.M. White, *Mon. Not. R. Astron. Soc.* 258 (1992) 1P–6P.
- [202] G. Efstathiou, *Mon. Not. R. Astron. Soc.* 276 (1995) 1425–1434.
- [203] G. Efstathiou, S.J. Moody, *Mon. Not. R. Astron. Soc.* 325 (2001) 1603–1615.
- [204] D.J. Eisenstein, M. Zaldarriaga, *Astrophys. J.* 546 (2001) 2–19.
- [205] E. Elizalde, E. Gaztañaga, *Mon. Not. R. Astron. Soc.* 254 (1992) 247–256.
- [206] A.E. Evrard, T.J. MacFarland, H.M.P. Couchman, J.M. Colberg, N. Yoshida, S.D.M. White, A.R. Jenkins, C.S. Frenk, F.R. Pearce, G. Efstathiou, J.A. Peacock, P.A. Thomas, astro-ph/0110246, 2001.
- [207] A.E. Evrard, M.M. Crone, *Astrophys. J.* 394 (1992) L1–L4.
- [208] E.E. Falco, M.J. Kurtz, M.J. Geller, J.P. Huchra, J. Peters, P. Berlind, D.J. Mink, S.P. Tokarz, B. Elwell, *Publ. Astron. Soc. Pac.* 111 (1999) 438–452.
- [209] S.M. Fall, S. Tremaine, *Astrophys. J.* 216 (1997) 682–689.
- [210] S.M. Fall, *Rev. Mod. Phys.* 51 (1979) 21–42.
- [211] H.A. Feldman, J.A. Frieman, J.N. Fry, R. Scoccimarro, *Phys. Rev. Lett.* 86 (2001) 1434–1438.
- [212] H.A. Feldman, N. Kaiser, J. Peacock, *Astrophys. J.* 426 (1994) 23–37.
- [213] P.G. Ferreira, R. Juszkiewicz, H.A. Feldman, M. Davis, A.H. Jaffe, *Astrophys. J.* 515 (1999) L1–L4.
- [214] J.A. Fillmore, P. Goldreich, *Astrophys. J.* 281 (1984) 1–8.
- [215] K.B. Fisher, M. Davis, M.A. Strauss, A. Yahil, J.P. Huchra, *Astrophys. J.* 402 (1993) 42–57.
- [216] K.B. Fisher, M. Davis, M.A. Strauss, A. Yahil, J.P. Huchra, *Mon. Not. R. Astron. Soc.* 266 (1994) 50–64.
- [217] K.B. Fisher, M. Davis, M.A. Strauss, A. Yahil, J.P. Huchra, *Mon. Not. R. Astron. Soc.* 267 (1994) 927–948.
- [218] K.B. Fisher, J.P. Huchra, M.A. Strauss, M. Davis, A. Yahil, D. Schlegel, *Astrophys. J.* 100 (Suppl.) (1995) 69–103.
- [219] K.B. Fisher, *Astrophys. J.* 448 (1995) 494–499.
- [220] K.B. Fisher, A. Nusser, *Astrophys. J.* 279 (1996) L1–L5.
- [221] R. Fong, D. Hale-Sutton, T. Shanks, *Mon. Not. R. Astron. Soc.* 257 (1992) 650–658.
- [222] P. Fosalba, E. Gaztañaga, *Mon. Not. R. Astron. Soc.* 301 (1998) 503–523.
- [223] P. Fosalba, E. Gaztañaga, *Mon. Not. R. Astron. Soc.* 301 (1998) 535–546.
- [224] J.A. Frieman, E. Gaztañaga, *Astrophys. J.* 425 (1994) 392–402.
- [225] J.A. Frieman, E. Gaztañaga, *Astrophys. J.* 521 (1999) L83–L86.
- [226] J.N. Fry, P.J.E. Peebles, *Astrophys. J.* 221 (1978) 19–33.
- [227] J.N. Fry, P.J.E. Peebles, *Astrophys. J.* 238 (1980) 785–792.
- [228] J.N. Fry, *Astrophys. J.* 262 (1982) 424–431.
- [229] J.N. Fry, M. Seldner, *Astrophys. J.* 259 (1982) 474–481.

- [230] J.N. Fry, *Astrophys. J.* 267 (1983) 483–487.
- [231] J.N. Fry, *Astrophys. J.* 277 (1984) L5–L8.
- [232] J.N. Fry, *Astrophys. J.* 279 (1984) 499–510.
- [233] J.N. Fry, *Astrophys. J.* 289 (1985) 10–17.
- [234] J.N. Fry, A.L. Melott, S.F. Shandarin, *Astrophys. J.* 412 (1993) 504–512.
- [235] J.N. Fry, E. Gaztañaga, *Astrophys. J.* 413 (1993) 447–452.
- [236] J.N. Fry, *Phys. Rev. Lett.* 73 (1994) 215–219.
- [237] J.N. Fry, *Astrophys. J.* 421 (1994) 21–26.
- [238] J.N. Fry, R.J. Scherrer, *Astrophys. J.* 429 (1994) 36–42.
- [239] J.N. Fry, E. Gaztañaga, *Astrophys. J.* 425 (1994) 1–13.
- [240] J.N. Fry, A.L. Melott, S.F. Shandarin, *Mon. Not. R. Astron. Soc.* 274 (1995) 745–754.
- [241] J.N. Fry, *Astrophys. J.* 461 (1996) L65–L67.
- [242] J.N. Fry, D. Thomas, *Astrophys. J.* 524 (1999) 591–602.
- [243] T. Fukushige, J. Makino, *Astrophys. J.* 477 (1997) L9–L12.
- [244] T. Fukushige, Y. Suto, *Astrophys. J.* 557 (2001) L11–L14.
- [245] A. Gangui, *Phys. Rev. D* 62 (2000) 103004.
- [246] E. Gaztañaga, *Astrophys. J.* 398 (1992) L17–L20.
- [247] E. Gaztañaga, J.I. Yokoyama, *Astrophys. J.* 403 (1993) 450–465.
- [248] E. Gaztañaga, J.A. Frieman, *Astrophys. J.* 437 (1994) L13–L16.
- [249] E. Gaztañaga, *Mon. Not. R. Astron. Soc.* 268 (1994) 913–924.
- [250] E. Gaztañaga, *Astrophys. J.* 454 (1995) 561–572.
- [251] E. Gaztañaga, C.M. Baugh, *Mon. Not. R. Astron. Soc.* 273 (1995) 1P–6P.
- [252] E. Gaztañaga, P. Mähönen, *Astrophys. J.* 462 (1996) L1–L4.
- [253] E. Gaztañaga, C.M. Baugh, *Mon. Not. R. Astron. Soc.* 294 (1998) 229–244.
- [254] E. Gaztañaga, F. Bernardeau, *Astron. Astrophys.* 331 (1998) 829–837.
- [255] E. Gaztañaga, P. Fosalba, *Mon. Not. R. Astron. Soc.* 301 (1998) 524–534.
- [256] E. Gaztañaga, R.A.C. Croft, *Mon. Not. R. Astron. Soc.* 309 (1999) 885–904.
- [257] E. Gaztañaga, G.B. Dalton, *Mon. Not. R. Astron. Soc.* 312 (2000) 417–428.
- [258] E. Gaztañaga, P. Fosalba, E. Elizalde, *Astrophys. J.* 539 (2000) 522–531.
- [259] E. Gaztanaga, A. Lobo, *Astrophys. J.* 548 (2001) 47–59.
- [260] E. Gaztanaga, R. Juszkiewicz, *Astrophys. J.* 558 (2001) L1–L4.
- [261] E. Gaztanaga, *Mon. Not. R. Astron. Soc.*, in press, astro-ph/0106379, 2001.
- [262] E. Gaztanaga, *Astrophys. J.*, in press, astro-ph/0110126, 2001.
- [263] E. Gaztañaga, P. Fosalba, R.A.C. Croft, *Mon. Not. R. Astron. Soc.* 331 (2002) 13–22.
- [264] S. Gelato, D.F. Chernoff, I. Wasserman, *Astrophys. J.* 480 (1997) 115–131.
- [265] M.J. Geller, M.J. Kurtz, V. de Lapparent, *Astrophys. J.* 287 (1984) L55–L58.
- [266] M.J. Geller, M.J. Kurtz, G. Wegner, J.R. Thorstensen, D.G. Fabricant, R.O. Marzke, J.P. Huchra, R.E. Schild, E.E. Falco, *Astron. J.* 114 (1997) 2205–2211.
- [267] S. Ghigna, S.A. Bonometto, L. Guzzo, R. Giovanelli, M.P. Haynes, A. Klypin, J.R. Primack, *Astrophys. J.* 463 (1996) 395–403.
- [268] R. Giovanelli, M.P. Haynes, *Annu. Rev. Astron. Astrophys.* 29 (1991) 499–541.
- [269] N.Y. Gnedin, *Astrophys. J.* 97 (Suppl.) (1995) 231–257.
- [270] M.H. Goroff, B. Grinstein, S.-J. Rey, M.B. Wise, *Astrophys. J.* 311 (1986) 6–14.
- [271] J.R. Gott, M.J. Rees, *Astron. Astrophys.* 45 (1975) 365–376.
- [272] J.R. Gott, E.L. Turner, *Astrophys. J.* 232 (1979) L79–L81.
- [273] M. Gramann, *Astrophys. J.* 405 (1993) L47–L50.
- [274] B. Grinstein, M.B. Wise, *Astrophys. J.* 320 (1987) 448–453.
- [275] E.J. Groth, P.J.E. Peebles, *Astrophys. J.* 217 (1977) 385–405.
- [276] E.J. Groth, P.J.E. Peebles, *Astrophys. J.* 310 (1986) 507–517.
- [277] E.J. Groth, P.J.E. Peebles, *Astrophys. J.* 310 (1986) 499–506.
- [278] S.N. Gurbatov, A.I. Saichev, S.F. Shandarin, *Mon. Not. R. Astron. Soc.* 236 (1989) 385–402.
- [279] A.H. Guth, *Phys. Rev. D* 23 (1981) 347–356.

- [280] A.H. Guth, S.-Y. Pi, *Phys. Rev. Lett.* 49 (1982) 1110–1113.
- [281] A.H. Guth, S.-Y. Pi, *Phys. Rev. D* 32 (1985) 1899–1920.
- [282] N.W. Halverson, E.M. Leitch, C. Pryke, J. Kovac, J.E. Carlstrom, W.L. Holzapfel, M. Dragovan, J.K. Cartwright, B.S. Mason, S. Padin, T.J. Pearson, M.C. Shepherd, A.C.S. Readhead, *Astrophys. J.* 568 (2002) 38–45.
- [283] T. Hamana, S. Colombi, Y. Suto, *Astron. Astrophys.* 367 (2001) 18–26.
- [284] T. Hamana, S. Colombi, A. Thion, J. Devriendt, Y. Mellier, F. Bernardeau, *Mon. Not. R. Astron. Soc.* 330 (2002) 365–377.
- [285] T. Hamana, *Mon. Not. R. Astron. Soc.* 326 (2002) 326–332.
- [286] A.J.S. Hamilton, *Astrophys. J.* 332 (1988) 67–74.
- [287] A.J.S. Hamilton, J.R. Gott, *Astrophys. J.* 331 (1988) 641–647.
- [288] A.J.S. Hamilton, *Astrophys. J.* 332 (1988) 67–74.
- [289] A.J.S. Hamilton, P. Kumar, A. Matthews, E. Lu, *Astrophys. J.* 374 (1991) L1–L4;
A.J.S. Hamilton, P. Kumar, A. Matthews, E. Lu, *Astrophys. J.* 442 (1995) L73–L73.
- [290] A.J.S. Hamilton, *Astrophys. J.* 385 (1992) L5–L8.
- [291] A.J.S. Hamilton, *Astrophys. J.* 417 (1993) 19–35.
- [292] A.J.S. Hamilton, M. Culhane, *Mon. Not. R. Astron. Soc.* 278 (1996) 73–86.
- [293] A.J.S. Hamilton, *Mon. Not. R. Astron. Soc.* 289 (1997) 285–294.
- [294] A.J.S. Hamilton, *Mon. Not. R. Astron. Soc.* 289 (1997) 295–304.
- [295] A.J.S. Hamilton, in: D. Hamilton (Ed.), *The Evolving Universe*, Kluwer Academic, Dordrecht, 1998, pp. 185–275.
- [296] A.J.S. Hamilton, *Mon. Not. R. Astron. Soc.* 312 (2000) 257–284.
- [297] A.J.S. Hamilton, M. Tegmark, *Mon. Not. R. Astron. Soc.* 312 (2000) 285–294.
- [298] A.J.S. Hamilton, M. Tegmark, *Mon. Not. R. Astron. Soc.* 330 (2002) 506–530.
- [299] A.J.S. Hamilton, M. Tegmark, N. Padmanabhan, *Mon. Not. R. Astron. Soc.* 317 (2000) L23–L27.
- [300] E. Harrison, *Phys. Rev. D* 1 (1970) 2726–2730.
- [301] S. Hatton, S. Cole, *Mon. Not. R. Astron. Soc.* 296 (1998) 10–20.
- [302] S. Hatton, S. Cole, *Mon. Not. R. Astron. Soc.* 310 (1999) 1137–1146.
- [303] M.G. Hauser, J.P.E. Peebles, *Astrophys. J.* 185 (1973) 757–786.
- [304] S.W. Hawking, *Phys. Lett. B* 115 (1982) 295–297.
- [305] D.J. Heath, *Mon. Not. R. Astron. Soc.* 179 (1977) 351–358.
- [306] A.F. Heavens, A.N. Taylor, *Mon. Not. R. Astron. Soc.* 275 (1995) 483–497.
- [307] A.F. Heavens, *Mon. Not. R. Astron. Soc.* 299 (1998) 805–808.
- [308] A.F. Heavens, A. Refregier, C. Heymans, *Mon. Not. R. Astron. Soc.* 319 (2000) 649–656.
- [309] L. Hernquist, *Astrophys. J.* 64 (Suppl.) (1987) 715–734.
- [310] L. Hernquist, F.R. Bouchet, Y. Suto, *Astrophys. J.* 75 (Suppl.) (1991) 231–240.
- [311] P.C. Hewett, *Mon. Not. R. Astron. Soc.* 201 (1982) 867–883.
- [312] C. Heyde, *J. R. Stat. Soc.* 25 (1963) 392–393.
- [313] E. Hivon, F.R. Bouchet, S. Colombi, R. Juszkiewicz, *Astron. Astrophys.* 298 (1995) 643–660.
- [314] R.W. Hockney, J.W. Eastwood, *Computer Simulation using Particles*, McGraw-Hill, New York, 1981.
- [315] H. Hoekstra, H.K.C. Yee, M.D. Gladders, *Astrophys. J.* 558 (2001) L11–L14.
- [316] D.W. Hogg, *astro-ph/9905116*, 1999.
- [317] E.B. Holmberg, A. Lauberts, H.E. Schuster, R.M. West, *Astron. Astrophys.* 18 (Suppl.) (1974) 463–489.
- [318] F. Hoyle, C. Baugh, T. Shanks, A. Ratcliffe, *Mon. Not. R. Astron. Soc.* 309 (1999) 659–671.
- [319] F. Hoyle, I. Szapudi, C. Baugh, *Mon. Not. R. Astron. Soc.* 317 (2000) L51–L56.
- [320] W. Hu, N. Sugiyama, *Astrophys. J.* 444 (1995) 489–506.
- [321] W. Hu, N. Sugiyama, *Astrophys. J.* 471 (1996) 542–570.
- [322] E.P. Hubble, Ph.D. Thesis, 1917, p. 1.
- [323] E.P. Hubble, *Astrophys. J.* 84 (1936) 517–554.
- [324] J. Huchra, M. Davis, D. Latham, J. Tonry, *Astrophys. J.* 52 (Suppl.) (1983) 89–119.
- [325] J.D. Hudon, S.J. Lilly, *Astrophys. J.* 469 (1996) 519–528.
- [326] L. Hui, *Astrophys. J.* 519 (1999) L9–L12.
- [327] L. Hui, E. Bertschinger, *Astrophys. J.* 471 (1996) 1–12.
- [328] L. Hui, E. Gaztañaga, *Astrophys. J.* 519 (1999) 622–636.

- [329] L. Hui, A. Stebbins, S. Burles, *Astrophys. J.* 511 (1999) L5–L8.
- [330] L. Hui, L. Kofman, S. Shandarin, *Astrophys. J.* 537 (2000) 12–20.
- [331] D. Huterer, L. Knox, R.C. Nichol, *Astrophys. J.* 555 (2001) 547–557.
- [332] L. Infante, C.J. Pritchett, *Astrophys. J.* 439 (1995) 565–583.
- [333] A. Jaffe, *Phys. Rev. D* 49 (1994) 3893–3909.
- [334] B. Jain, E. Bertschinger, *Astrophys. J.* 431 (1994) 495–505.
- [335] B. Jain, H.J. Mo, S.D.M. White, *Mon. Not. R. Astron. Soc.* 276 (1995) L25–L29.
- [336] B. Jain, E. Bertschinger, *Astrophys. J.* 456 (1996) 43–54.
- [337] B. Jain, *Mon. Not. R. Astron. Soc.* 287 (1997) 687–698.
- [338] B. Jain, E. Bertschinger, *Astrophys. J.* 509 (1998) 517–530.
- [339] B. Jain, U. Seljak, *Astrophys. J.* 484 (1997) 560–573.
- [340] B.T. Jannuzi, A. Dey, in: R. Weymann, L. Storrie-Lombardi, M. Sawicki, R. Brunner (Eds.), *Photometric Redshifts and the Detection of High Redshift Galaxies*, ASP Conference Series, Vol. 191, p. 111.
- [341] C. Jessop, M. Duncan, W.Y. Chau, *J. Comput. Phys.* 115 (1994) 339–351.
- [342] 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, A.H. Nelson, *Astrophys. J.* 499 (1998) 20–40.
- [343] A. Jenkins, C.S. Frenk, S.D.M. White, J.M. Colberg, S. Cole, A.E. Evrard, H.M.P. Couchman, N. Yoshida, *Mon. Not. R. Astron. Soc.* 321 (2001) 372–384.
- [344] L.G. Jensen, A.S. Szalay, *Astrophys. J.* 305 (1986) L5–L9.
- [345] Y.P. Jing, H. Mo, G. Börner, *Astron. Astrophys.* 252 (1991) 449–457.
- [346] Y.P. Jing, G. Börner, *Astron. Astrophys.* 318 (1997) 667–672.
- [347] Y.P. Jing, G. Börner, *Astrophys. J.* 503 (1998) 37–47.
- [348] Y.P. Jing, G. Börner, *Astrophys. J.* 503 (1998) 502–504.
- [349] Y.P. Jing, H.J. Mo, G. Börner, *Astrophys. J.* 494 (1998) 1–12.
- [350] Y.P. Jing, *Astrophys. J.* 515 (1999) L45–L48.
- [351] Y.P. Jing, G. Börner, Y. Suto, *Astrophys. J.* 564 (2002) 15–22.
- [352] R. Jeannerot, *Phys. Rev. D* 56 (1999) 6205–6216.
- [353] R. Juszkiewicz, *Mon. Not. R. Astron. Soc.* 197 (1981) 931–940.
- [354] R. Juszkiewicz, D.H. Sonoda, J.D. Barrow, *Mon. Not. R. Astron. Soc.* 209 (1984) 139–144.
- [355] R. Juszkiewicz, F.R. Bouchet, S. Colombi, *Astrophys. J.* 412 (1993) L9–L12.
- [356] R. Juszkiewicz, D.H. Weinberg, P. Amsterdamski, M. Chodorowski, F.R. Bouchet, *Astrophys. J.* 442 (1995) 39–56.
- [357] R. Juszkiewicz, K.B. Fisher, I. Szapudi, *Astrophys. J.* 504 (1998) L1–L4.
- [358] R. Juszkiewicz, V. Springel, R. Durrer, *Astrophys. J.* 518 (1999) L25–L28.
- [359] R. Juszkiewicz, P.G. Ferreira, H.A. Feldman, A.H. Jaffe, M. Davis, *Science* 287 (2000) 109–112.
- [360] N. Kaiser, *Astrophys. J.* 284 (1984) L9–L12.
- [361] N. Kaiser, *Mon. Not. R. Astron. Soc.* 219 (1986) 785–790.
- [362] N. Kaiser, *Mon. Not. R. Astron. Soc.* 227 (1987) 1–21.
- [363] N. Kaiser, J.A. Peacock, *Astrophys. J.* 379 (1991) 482–506.
- [364] N. Kaiser, *Astrophys. J.* 388 (1992) 272–286.
- [365] N. Kaiser, G. Wilson, G.A. Luppino, astro-ph/0003338, 2000.
- [366] M. Kamionkowski, A. Buchalter, *Astrophys. J.* 514 (1999) 7–11.
- [367] G. Karakatsanis, T. Buchert, A.L. Melott, *Astron. Astrophys.* 326 (1997) 873–884.
- [368] N. Katz, T. Quinn, J.M. Gelb, *Mon. Not. R. Astron. Soc.* 265 (1993) 689–705.
- [369] N. Katz, L. Hernquist, D.H. Weinberg, *Astrophys. J.* 523 (1999) 463–479.
- [370] G. Kauffmann, S.D.M. White, *Mon. Not. R. Astron. Soc.* 261 (1993) 921–928.
- [371] G. Kauffmann, J.M. Colberg, A. Diaferio, S.D.M. White, *Mon. Not. R. Astron. Soc.* 303 (1999) 188–206.
- [372] M. Kerscher, *Astron. Astrophys.* 343 (1999) 333–347.
- [373] M. Kerscher, I. Szapudi, A. Szalay, *Astrophys. J.* 535 (2000) L13–L16.
- [374] E.J. Kibblewhite, M.T. Bridgeland, P.S. Bunclark, M.J. Irwin, in: D.A. Klinglesmith (Ed.), *Astronomical Microdensitometry Conference*, Vol. 277, Proceedings of a Workshop Held at Goddard Space Flight Center, Greenbelt, MD, 11–13 May 1983, National Aeronautics and Space Administration, Scientific and Technical Information Branch, Washington, DC, 1984.

- [375] C. Kiefer, D. Polarski, A.A. Starobinsky, *Int. J. Mod. Phys. D* 7 (1998) 455–462.
- [376] C. Kiefer, J. Lesgourgues, D. Polarski, A.A. Starobinsky, *Class. Quant. Grav.* 15 (1998) L67–L72.
- [377] R.S. Kim, M.A. Strauss, *Astrophys. J.* 493 (1998) 39–51.
- [378] R.P. Kirshner, A. Oemler, P.L. Schechter, S.A. Shectman, *Astrophys. J.* 248 (1981) L57–L60.
- [379] A.A. Klypin, S.F. Shandarin, *Mon. Not. R. Astron. Soc.* 204 (1983) 891–907.
- [380] L. Kofman, G.R. Blumenthal, H. Hodges, J.R. Primack, in: D.W. Latham, L.A.N. da Costa (Eds.), *Large-Scale Structures and Peculiar Motions in the Universe*, Vol. 339, 1989.
- [381] L. Kofman, D. Pogosyan, S.F. Shandarin, *Mon. Not. R. Astron. Soc.* 242 (1990) 200–208.
- [382] L. Kofman, E. Bertschinger, J.M. Gelb, A. Nusser, A. Dekel, *Astrophys. J.* 420 (1994) 44–57.
- [383] E.W. Kolb, M.S. Turner, *The Early Universe*, Addison-Wesley, Reading, MA, 1993.
- [384] D.C. Koo, A.S. Szalay, *Astrophys. J.* 282 (1984) 390–397.
- [385] K. Koyama, J. Soda, A. Taruya, *Mon. Not. R. Astron. Soc.* 310 (1999) 1111–1118.
- [386] A.V. Kravtsov, A.A. Klypin, A.M. Khokhlov, *Astrophys. J.* 111 (Suppl.) (1997) 73–94.
- [387] A. Kudlicki, M. Chodorowski, T. Plewa, M. Różyczka, *Mon. Not. R. Astron. Soc.* 316 (2000) 464–472.
- [388] C. Lacey, S. Cole, *Mon. Not. R. Astron. Soc.* 262 (1993) 627–649.
- [389] M. Lachièze-Rey, J.P. Luminet, *Phys. Rep.* 254 (1995) 135–214.
- [390] O. Lahav, P.B. Lilje, J.R. Primack, M.J. Rees, *Mon. Not. R. Astron. Soc.* 251 (1991) 128–136.
- [391] O. Lahav, M. Itoh, S. Inagaki, Y. Suto, *Astrophys. J.* 402 (1993) 387–397.
- [392] L.D. Landau, E.M. Lifschitz, *Fluid Mechanics*, Pergamon Press, Oxford, 1987.
- [393] S.D. Landy, A.S. Szalay, *Astrophys. J.* 412 (1993) 64–71.
- [394] S.D. Landy, S.A. Shectman, H. Lin, R.P. Kirshner, A.A. Oemler, D. Tucker, *Astrophys. J.* 456 (1996) L1–L4.
- [395] S.D. Landy, A.S. Szalay, T.H. Broadhurst, *Astrophys. J.* 494 (1998) L133–L136.
- [396] D. Layzer, *Astron. J.* 61 (1956) 383–385.
- [397] A.T. Lee, P. Ade, A. Balbi, J. Bock, J. Borrill, A. Boscaleri, P. De Bernardis, P.G. Ferreira, S. Hanany, V.V. Hristov, A.H. Jaffe, P.D. Mauskopf, C.B. Netterfield, E. Pascale, B. Rabbii, P.L. Richards, G.F. Smoot, R. Stompor, C.D. Winant, J.H.P. Wu, *Astrophys. J.* 561 (2001) L1–L5.
- [398] O. Le Fèvre et al., in: *Mining the Sky*, ESO Astrophysics Symposia, 2001, in press, astro-ph/0101034.
- [399] A.R. Liddle, D.H. Lyth, *Phys. Rep.* 231 (1993) 1–105.
- [400] A.R. Liddle, D.H. Lyth, *Cosmological Inflation and Large-Scale Structure*, Cambridge University Press, Cambridge, UK, 2000.
- [401] S.J. Lilly, O. LeFevre, D. Crampton, F. Hammer, L. Tresse, *Astrophys. J.* 455 (1995) 50–59.
- [402] D.N. Limber, *Astrophys. J.* 117 (1953) 134–144.
- [403] H. Lin, R.P. Kirshner, S.A. Shectman, S.D. Landy, A. Oemler, D.L. Tucker, P.L. Schechter, *Astrophys. J.* 471 (1996) 617–635.
- [404] A. Linde, V. Mukhanov, *Phys. Rev. D* 56 (1997) R535–R539.
- [405] M. Longuet-Higgins, *J. Fluid. Mech.* 17 (1963) 459–480.
- [406] M. Longuet-Higgins, *Radio Sci. J. of Res. D* 68 (1964) 1049.
- [407] E.L. Łokas, R. Juszkiewicz, D.H. Weinberg, F.R. Bouchet, *Mon. Not. R. Astron. Soc.* 274 (1995) 730–744.
- [408] E.L. Łokas, R. Juszkiewicz, F.R. Bouchet, E. Hivon, *Astrophys. J.* 467 (1996) 1–9.
- [409] E.L. Łokas, *Mon. Not. R. Astron. Soc.* 296 (1998) 491–501.
- [410] J. Loveday, G. Efstathiou, B.A. Peterson, S.J. Maddox, *Astrophys. J.* 400 (1992) L43–L46.
- [411] J. Loveday, B.A. Peterson, S.J. Maddox, G. Efstathiou, *Astrophys. J.* 107 (Suppl.) (1996) 201–214.
- [412] J. Loveday, G. Efstathiou, S.J. Maddox, B.A. Peterson, *Astrophys. J.* 468 (1996) 1–16.
- [413] L.B. Lucy, *Astron. J.* 79 (1974) 745–754.
- [414] X. Luo, D.N. Schramm, *Astrophys. J.* 408 (1993) 33–42.
- [415] D.H. Lyth, A. Riotto, *Phys. Rev. D* 59 (1999) 103503.
- [416] C.-P. Ma, E. Bertschinger, *Astrophys. J.* 455 (1995) 7–25.
- [417] C.-P. Ma, *Astrophys. J.* 508 (1998) L5–L8.
- [418] C.-P. Ma, J.N. Fry, *Mon. Not. R. Astron. Soc.* 531 (2000) L87–L90.
- [419] C.-P. Ma, J.N. Fry, *Mon. Not. R. Astron. Soc.* 543 (2000) 503–513.
- [420] T. MacFarland, H.M.P. Couchman, F.R. Pearce, J. Pichlmeier, *New Astron.* 3 (1998) 687–705.

- [421] H.T. MacGillivray, R.J. Dodd, in: D.A. KlingleSmith (Ed.), *Astronomical Microdensitometry Conference*, Vol. 237, Proceedings of a Workshop Held at Goddard Space Flight Center, Greenbelt, MD, 11–13 May 1983, National Aeronautics and Space Administration, Scientific and Technical Information Branch, Washington, DC, 1984.
- [422] S.J. Maddox, G. Efstathiou, W.J. Sutherland, J. Loveday, *Mon. Not. R. Astron. Soc.* 242 (1990) 43P–47P.
- [423] S.J. Maddox, G. Efstathiou, W.J. Sutherland, J. Loveday, *Mon. Not. R. Astron. Soc.* 243 (1990) 692–712.
- [424] S.J. Maddox, G. Efstathiou, W.J. Sutherland, *Mon. Not. R. Astron. Soc.* 246 (1990) 433–457.
- [425] S.J. Maddox, G. Efstathiou, W.J. Sutherland, *Mon. Not. R. Astron. Soc.* 283 (1996) 1227–1263.
- [426] M. Magliocchetti, S.J. Maddox, O. Lahav, J.V. Wall, *Mon. Not. R. Astron. Soc.* 306 (1999) 943–953.
- [427] J. Makino, M. Taiji, T. Ebisuzaki, D. Sugimoto, *Astrophys. J.* 480 (1997) 432–446.
- [428] N. Makino, M. Sasaki, Y. Suto, *Phys. Rev. D* 46 (1992) 585–602.
- [429] P.J. Mancinelli, A. Yahil, *Astrophys. J.* 452 (1995) 75–81.
- [430] R.G. Mann, A.F. Heavens, J.A. Peacock, *Mon. Not. R. Astron. Soc.* 263 (1993) 798–816.
- [431] R.G. Mann, J.A. Peacock, A.F. Heavens, *Mon. Not. R. Astron. Soc.* 293 (1998) 209–221.
- [432] H. Martel, W. Freudling, *Astrophys. J.* 371 (1991) 1–7.
- [433] S. Matarrese, F. Lucchin, L. Moscardini, D. Saez, *Mon. Not. R. Astron. Soc.* 259 (1992) 437–452.
- [434] S. Matarrese, L. Verde, A.F. Heavens, *Mon. Not. R. Astron. Soc.* 290 (1997) 651–662.
- [435] T. Matsubara, *Astrophys. J.* 424 (1994) 30–41.
- [436] T. Matsubara, *Astrophys. J.* 434 (1994) L43–L46.
- [437] T. Matsubara, Y. Suto, *Astrophys. J.* 420 (1994) 497–503.
- [438] T. Matsubara, Y. Suto, *Astrophys. J.* 470 (1996) L1–L5.
- [439] T. Matsubara, Y. Suto, I. Szapudi, *Astrophys. J.* 491 (1997) L1–L5.
- [440] T. Matsubara, *Astrophys. J.* 525 (1999) 543–553.
- [441] T. Matsubara, astro-ph/0006269, 2000.
- [442] T. Matsubara, *Astrophys. J.* 535 (2000) 1–23.
- [443] T. Matsubara, A.S. Szalay, S.D. Landy, *Astrophys. J.* 535 (2000) L1–L4.
- [444] T. Matsubara, A.S. Szalay, *Astrophys. J.* 556 (2001) L67–L70.
- [445] S. Maurogordato, R. Schaeffer, L.N. da Costa, *Astrophys. J.* 390 (1992) 17–33.
- [446] J. McClelland, J. Silk, *Astrophys. J.* 217 (1977) 331–352.
- [447] H.J. McCracken, O. Le Fèvre, M. Brodwin, S. Foucaud, S.J. Lilly, D. Crampton, Y. Mellier, *Astron. Astrophys.* 376 (2001) 756–774.
- [448] P. McDonald, J. Miralda-Escudé, *Astrophys. J.* 518 (1999) 24–31.
- [449] P. McDonald, astro-ph/0108064, 2001.
- [450] C. McGill, *Mon. Not. R. Astron. Soc.* 242 (1990) 428–438.
- [451] A. Meiksin, I. Szapudi, A. Szalay, *Astrophys. J.* 394 (1992) 87–90.
- [452] A. Meiksin, M. White, *Mon. Not. R. Astron. Soc.* 308 (1999) 1179–1184.
- [453] Y. Mellier, *Annu. Rev. Astron. Astrophys.* 37 (1999) 127–189.
- [454] A.L. Melott, *Mon. Not. R. Astron. Soc.* 202 (1983) 595–604.
- [455] A.L. Melott, T. Buchert, A.G. Weiß, *Astron. Astrophys.* 294 (1995) 345–365.
- [456] J. Miralda-Escudé, *Astrophys. J.* 380 (1991) 1–8.
- [457] C.W. Misner, K. Thorne, J.A. Wheeler, *Gravitation*, Freeman, San Francisco, 1973.
- [458] H. Mo, Y.P. Jing, G. Börner, *Astrophys. J.* 392 (1992) 452–457.
- [459] H. Mo, Y.P. Jing, S.D.M. White, *Mon. Not. R. Astron. Soc.* 284 (1997) 189–201.
- [460] H. Mo, S.D.M. White, *Mon. Not. R. Astron. Soc.* 282 (1996) 347–361.
- [461] A.W. Moore, A.J. Connolly, C. Genovese, A. Gray, L. Grone, N. Kanidori II, R.C. Nichol, J. Schneider, A.S. Szalay, I. Szapudi, L. Wasserman, in: A.J. Banday, S. Zaroubi, M. Bartelmann (Eds.), *Mining the Sky*, ESO Astrophysics Symposia, Springer, Berlin, Heidelberg, New York, 2001, pp. 71–82.
- [462] B. Moore, C.S. Frenk, G. Efstathiou, W. Saunders, *Mon. Not. R. Astron. Soc.* 269 (1994) 742–752.
- [463] B. Moore, T. Quinn, F. Governato, J. Stadel, G. Lake, *Mon. Not. R. Astron. Soc.* 310 (1999) 1147–1152.
- [464] L. Moscardini, S. Matarrese, F. Lucchin, A. Messina, *Mon. Not. R. Astron. Soc.* 248 (1991) 424–438.
- [465] F. Moutarde, J.-M. Alimi, F.R. Bouchet, R. Pellat, A. Ramani, *Astrophys. J.* 382 (1991) 377–381.
- [466] V.F. Mukhanov, H.A. Feldman, R.H. Brandenberger, *Phys. Rep.* 215 (1992) 203–333.
- [467] D. Munshi, P. Coles, *Mon. Not. R. Astron. Soc.* 313 (2000) 148–152.

- [468] D. Munshi, B. Jain, *Mon. Not. R. Astron. Soc.* 318 (2000) 109–123.
- [469] D. Munshi, B. Jain, *Mon. Not. R. Astron. Soc.* 322 (2001) 107–120.
- [470] D. Munshi, A.A. Starobinski, *Astrophys. J.* 428 (1994) 433–438.
- [471] D. Munshi, V. Sahni, A.A. Starobinski, *Astrophys. J.* 436 (1994) 517–527.
- [472] D. Munshi, F. Bernardeau, A.L. Melott, R. Schaeffer, *Mon. Not. R. Astron. Soc.* 303 (1999) 433–445.
- [473] D. Munshi, A.L. Melott, P. Coles, *Mon. Not. R. Astron. Soc.* 311 (2000) 149–160.
- [474] V.K. Narayanan, A.A. Berlind, D.H. Weinberg, *Astrophys. J.* 528 (2000) 1–20.
- [475] J.F. Navarro, C.S. Frenk, S.D.M. White, *Astrophys. J.* 490 (1997) 493–508.
- [476] C.B. Netterfield, P.A.R. Ade, J.J. Bock, J.R. Bond, J. Borrill, A. Boscaleri, K. Coble, C.R. Contaldi, B.P. Crill, P. de Bernardis, P. Farese, K. Ganga, M. Giacometti, E. Hivon, V.V. Hristov, A. Iacoangeli, A.H. Jaffe, W.C. Jones, A.E. Lange, L. Martinis, S. Masi, P. Mason, P.D. Mauskopf, A. Melchiorri, T. Montroy, E. Pascale, F. Piacentini, D. Pogosyan, F. Pongetti, S. Prunet, G. Romeo, J.E. Ruhl, F. Scaramuzzi, astro-ph/0104460, 2001.
- [477] J. Neyman, E.L. Scott, *Astrophys. J.* 116 (1952) 144–163.
- [478] P. Nilson, *Uppsala General Catalogue of Galaxies*, Uppsala, 1973.
- [479] R. Nityananda, T. Padmanabhan, *Mon. Not. R. Astron. Soc.* 271 (1994) 976–980.
- [480] P. Norberg, C.M. Baugh, E. Hawkins, S. Maddox, J.A. Peacock, S. Cole, C.S. Frenk, et al., *Mon. Not. R. Astron. Soc.* 328 (2001) 64–70.
- [481] P. Norberg, S. Cole, et al., astro-ph/0111011, 2001.
- [482] A. Nusser, M. Davis, *Astrophys. J.* 421 (1994) L1–L4.
- [483] A. Nusser, M. Haehnelt, *Mon. Not. R. Astron. Soc.* 303 (1999) 179–187.
- [484] J.H. Oort, *Annu. Rev. Astron. Astrophys.* 21 (1983) 373–428.
- [485] T. Padmanabhan, *Mon. Not. R. Astron. Soc.* 278 (1996) L29–L33.
- [486] T. Padmanabhan, R. Cen, J.P. Ostriker, F.J. Summers, *Astrophys. J.* 466 (1996) 604–613.
- [487] N. Padmanabhan, M. Tegmark, A.J.S. Hamilton, *Astrophys. J.* 550 (2001) 52–64.
- [488] C. Park, J.R. Gott, *Mon. Not. R. Astron. Soc.* 249 (1991) 288–299.
- [489] C. Park, M.S. Vogele, M.J. Geller, J.P. Huchra, *Astrophys. J.* 431 (1994) 569–585.
- [490] W.E. Parry, *Phys. Lett. A* 60 (1997) 265.
- [491] J.A. Peacock, A.F. Heavens, *Mon. Not. R. Astron. Soc.* 217 (1985) 805–820.
- [492] J.A. Peacock, D. Nicholson, *Mon. Not. R. Astron. Soc.* 253 (1991) 307–319.
- [493] J.A. Peacock, S.J. Dodds, *Mon. Not. R. Astron. Soc.* 267 (1994) 1020–1034.
- [494] J.A. Peacock, S.J. Dodds, *Mon. Not. R. Astron. Soc.* 280 (1996) L19–L26.
- [495] J.A. Peacock, R.E. Smith, *Mon. Not. R. Astron. Soc.* 318 (2000) 1144–1156.
- [496] J.A. Peacock, et al., *Nature* 410 (2001) 169–173.
- [497] J.A. Peacock, S. Cole, P. Norberg, et al., *Nature* 410 (2001) 169–173.
- [498] F.R. Pearce, A. Jenkins, C.S. Frenk, J.M. Colberg, S.D.M. White, P.A. Thomas, H.M.P. Couchman, J.A. Peacock, G. Efstathiou, *Astrophys. J.* 521 (1999) L99–L102.
- [499] P.J.E. Peebles, J.T. Yu, *Astrophys. J.* 162 (1970) 815–836.
- [500] P.J.E. Peebles, *Astrophys. J.* 185 (1973) 413–440.
- [501] P.J.E. Peebles, *Astrophys. J.* 28 (Suppl.) (1974) 37–50.
- [502] P.J.E. Peebles, *Astron. Astrophys.* 32 (1974) 197–201.
- [503] P.J.E. Peebles, M.G. Hauser, *Astrophys. J.* 28 (Suppl.) (1974) 19–36.
- [504] P.J.E. Peebles, E.J. Groth, *Astrophys. J.* 196 (1975) 1–11.
- [505] P.J.E. Peebles, *Astrophys. J.* 196 (1975) 647–651.
- [506] P.J.E. Peebles, *Astrophys. J.* 205 (1976) 318–328.
- [507] P.J.E. Peebles, *Astrophys. Space Sci.* 45 (1976) 3–19.
- [508] P.J.E. Peebles, *The Large-Scale Structure of the Universe*, Princeton University Press, Princeton, 1980.
- [509] P.J.E. Peebles, *Astrophys. J.* 263 (1982) L1–L5.
- [510] P.J.E. Peebles, *Astrophys. J.* 365 (1990) 27–36.
- [511] P.J.E. Peebles, *Principles of Physical Cosmology*, Princeton University Press, Princeton, 1993.
- [512] P.J.E. Peebles, *Astrophys. J.* 483 (1997) L1–L4.
- [513] P.J.E. Peebles, *Astrophys. J.* 510 (1999) 523–530.
- [514] P.J.E. Peebles, *Astrophys. J.* 510 (1999) 531–540.

- [515] P.J.E. Peebles, in: V.J. Martinez, V. Trimble, M.J. Pons-Borderia (Eds.), *Historical Development of Modern Cosmology*, ASP Conference Series, 2001, astro-ph/0103040.
- [516] U.-L. Pen, *Astrophys. J.* 100 (Suppl.) (1995) 269–280.
- [517] U.-L. Pen, *Astrophys. J.* 504 (1998) 601–606.
- [518] W.J. Percival, *Mon. Not. R. Astron. Soc.* 327 (2001) 1297–1306.
- [519] B.A. Peterson, R.S. Ellis, E.J. Kibblewhite, M.T. Bridgeland, T. Hooley, D. Horne, *Astrophys. J.* 233 (1979) L109–L113.
- [520] S. Phillipps, R. Fong, S.M. Fall, H.T. MacGillivray, *Mon. Not. R. Astron. Soc.* 182 (1978) 673–686.
- [521] C. Pichon, F. Bernardeau, *Astron. Astrophys.* 343 (1999) 663–681.
- [522] E. Pierpaoli, J. García-Bellido, S. Borgani, *J. High Energy Phys.* 10 (1999) 15–48.
- [523] H.D. Politzer, M.B. Wise, *Astrophys. J.* 285 (1984) L1–L3.
- [524] A. Pollo, *Acta Astron.* 47 (1997) 413–429.
- [525] M.-J. Pons-Borderi, V.J. Martínez, D. Stoyan, H. Stoyan, E. Saar, *Astrophys. J.* 523 (1999) 480–491.
- [526] P.A. Popowski, D.H. Weinberg, B.S. Ryden, P.S. Osmer, *Astrophys. J.* 498 (1998) 11–25.
- [527] C. Porciani, P. Catelan, C. Lacey, *Astrophys. J.* 513 (1999) L99–L102.
- [528] C. Porciani, M. Giavalisco, *Astrophys. J.* 565 (2002) 24–29.
- [529] M. Postman, M.J. Geller, *Astrophys. J.* 281 (1994) 95–99.
- [530] M. Postman, T.R. Lauer, I. Szapudy, W. Oegerle, *Astrophys. J.* 506 (1998) 33–44.
- [531] W.H. Press, P. Schechter, *Astrophys. J.* 187 (1974) 425–438.
- [532] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, *Numerical Recipes in FORTRAN*, 2nd Edition, Cambridge University Press, Cambridge, 1992.
- [533] C.J. Pritchett, L. Infante, *Astron. J.* 91 (1986) 1–5.
- [534] Z.A.M. Protogerios, R.J. Scherrer, *Mon. Not. R. Astron. Soc.* 284 (1997) 425–438.
- [535] A. Ratcliffe, T. Shanks, Q.A. Parker, A. Broadbent, F.G. Watson, A.P. Oates, C.A. Collins, R. Fong, *Mon. Not. R. Astron. Soc.* 300 (1998) 417–462.
- [536] B. Ratra, P.J.E. Peebles, *Phys. Rev. D* 37 (1988) 3406–3427.
- [537] B.D. Ripley, *Statistical Inference for Spatial Processes*, Cambridge University Press, Cambridge, 1988.
- [538] J. Robinson, E. Gawiser, J. Silk, *Astrophys. J.* 532 (2000) 1–16.
- [539] N. Roche, T. Shanks, N. Metcalfe, R. Fong, *Mon. Not. R. Astron. Soc.* 263 (1993) 360–368.
- [540] N. Roche, S.A. Eales, *Mon. Not. R. Astron. Soc.* 307 (1999) 703–721.
- [541] H.J. Rood, *Astrophys. J.* 49 (Suppl.) (1982) 111–147.
- [542] L. Ruamsuwan, J.N. Fry, *Astrophys. J.* 396 (1992) 416–429.
- [543] K. Rudnicki, T.Z. Dworak, P. Flin, B. Baranowski, A. Sendrakowski, *Acta Cosmol.* 1 (1973) 7.
- [544] R.K. Sachs, *Proc. R. Soc. London A* 264 (1961) 309–338.
- [545] J.K. Salmon, M.S. Warren, *J. Comput. Phys.* 111 (1994) 136–155.
- [546] J.K. Salmon, M.S. Warren, G.S. Winckelmans, *Int. J. Supercomp. Appl.* 8 (1994) 129–142.
- [547] R.F. Sanford, *Lick Obs. Bull.* 297 (1917) 80–91.
- [548] W.C. Saslaw, *Astrophys. J.* 177 (1972) 17–29.
- [549] W. Saunders, C. Frenk, M. Rowan-Robinson, A. Lawrence, G. Efstathiou, *Nature* 349 (1991) 32–38.
- [550] W. Saunders, et al., *Mon. Not. R. Astron. Soc.* 317 (2000) 55–64.
- [551] R. Schaeffer, *Astron. Astrophys.* 134 (1984) L15–L16.
- [552] R.J. Scherrer, E. Bertschinger, *Astrophys. J.* 381 (1991) 349–360.
- [553] R.J. Scherrer, D.H. Weinberg, *Astrophys. J.* 504 (1998) 607–611.
- [554] R.J. Scherrer, E. Gaztañaga, *Mon. Not. R. Astron. Soc.* 328 (2001) 257–265.
- [555] D.J. Schlegel, D.P. Finkbeiner, M. Davis, *Astrophys. J.* 500 (1998) 525–553.
- [556] P. Schneider, M. Bartelmann, *Mon. Not. R. Astron. Soc.* 273 (1995) 475–483.
- [557] R. Scoccimarro, J.A. Frieman, *Astrophys. J.* 105 (Suppl.) (1996) 37–73.
- [558] R. Scoccimarro, J.A. Frieman, *Astrophys. J.* 473 (1996) 620–644.
- [559] R. Scoccimarro, *Astrophys. J.* 487 (1997) 1–17.
- [560] R. Scoccimarro, S. Colombi, J.N. Fry, J.A. Frieman, E. Hivon, A. Melott, *Astrophys. J.* 496 (1998) 586–604.
- [561] R. Scoccimarro, *Mon. Not. R. Astron. Soc.* 299 (1998) 1097–1118.
- [562] R. Scoccimarro, H.M.P. Couchman, J.A. Frieman, *Astrophys. J.* 517 (1999) 531–540.

- [563] R. Scoccimarro, J.A. Frieman, *Astrophys. J.* 520 (1999) 35–44.
- [564] R. Scoccimarro, M. Zaldarriaga, L. Hui, *Astrophys. J.* 527 (1999) 1–15.
- [565] R. Scoccimarro, *Astrophys. J.* 542 (2000) 1–8.
- [566] R. Scoccimarro, *Astrophys. J.* 544 (2000) 597–615.
- [567] R. Scoccimarro, H. Feldman, J.N. Fry, J.A. Frieman, *Astrophys. J.* 546 (2001) 652–664.
- [568] R. Scoccimarro, H.M.P. Couchman, *Mon. Not. R. Astron. Soc.* 325 (2001) 1312–1316.
- [569] R. Scoccimarro, in: J.M. Fry, R. Buchler, H. Kandrup (Eds.), *The Onset of Nonlinearity in Cosmology*, *Ann. NY. Acad. Sci.* 927 (2001) 13–24, astro-ph/0008277.
- [570] R. Scoccimarro, R.K. Sheth, L. Hui, B. Jain, *Astrophys. J.* 546 (2001) 20–34.
- [571] R. Scoccimarro, R.K. Sheth, *Mon. Not. R. Astron. Soc.* 329 (2002) 629–640.
- [572] R. Scranton, D. Johnston, S. Dodelson, J.A. Frieman, et al., *Astrophys. J.*, 2001, submitted for publication, astro-ph/0107416.
- [573] M.D. Seabone, W. Sutherland, H. Tadros, G. Efstathiou, C.S. Frenk, O. Keeble, S.J. Maddox, R.G. McMahon, S. Oliver, M. Rowan-Robinson, W. Saunders, S.D.M. White, *Mon. Not. R. Astron. Soc.* 309 (1999) 89–99.
- [574] W.L. Sebok, *Astrophys. J.* 62 (Suppl.) (1986) 301–330.
- [575] M. Seldner, P.J.E. Peebles, *Astrophys. J.* 215 (1977) 703–716.
- [576] M. Seldner, P.J.E. Peebles, *Astrophys. J.* 225 (1978) 7–20.
- [577] M. Seldner, P.J.E. Peebles, *Astrophys. J.* 227 (1979) 30–36.
- [578] U. Seljak, M. Zaldarriaga, *Astrophys. J.* 469 (1996) 437–444.
- [579] U. Seljak, *Mon. Not. R. Astron. Soc.* 318 (2000) 203–213.
- [580] N. Seto, *Astrophys. J.* 523 (1999) 24–31.
- [581] C.D. Shane, C.A. Wirtanen, *Publ. Lick Obs.* 22 (Part 1) (1967).
- [582] T. Shanks, R. Fong, R.S. Ellis, H.T. MacGillivray, *Mon. Not. R. Astron. Soc.* 192 (1980) 209–232.
- [583] N.A. Sharp, S.A. Bonometto, F. Lucchin, *Astron. Astrophys.* 130 (1984) 79–85.
- [584] S.A. Shectman, S.D. Landy, A. Oemler, D.L. Tucker, H. Lin, R.P. Kirshner, P.L. Schechter, *Astrophys. J.* 470 (1996) 172–188.
- [585] R.K. Sheth, *Mon. Not. R. Astron. Soc.* 277 (1995) 933–944.
- [586] R.K. Sheth, *Mon. Not. R. Astron. Soc.* 279 (1996) 1310–1324.
- [587] R.K. Sheth, B. Jain, *Mon. Not. R. Astron. Soc.* 285 (1997) 231–238.
- [588] R.K. Sheth, *Mon. Not. R. Astron. Soc.* 300 (1998) 1057–1070.
- [589] R.K. Sheth, G. Lemson, *Mon. Not. R. Astron. Soc.* 304 (1999) 767–792.
- [590] R.K. Sheth, B. Tormen, *Mon. Not. R. Astron. Soc.* 308 (1999) 119–126.
- [591] R.K. Sheth, L. Hui, A. Diaferio, R. Scoccimarro, *Mon. Not. R. Astron. Soc.* 325 (2001) 1288–1302.
- [592] R.K. Sheth, A. Diaferio, L. Hui, R. Scoccimarro, *Mon. Not. R. Astron. Soc.* 326 (2001) 463–472.
- [593] R.K. Sheth, H.J. Mo, B. Tormen, *Mon. Not. R. Astron. Soc.* 323 (2001) 1–12.
- [594] J. Silk, R. Juszkiewicz, *Nature* 353 (1991) 386–388.
- [595] V.A. Smirnov, *Renormalization and Asymptotic Expansions*, Birkhäuser, Berlin, 1991.
- [596] G.F. Smoot, C.L. Bennett, A. Kogut, E.L. Wright, J. Aymon, N.W. Boggess, E.S. Cheng, G. de Amici, S. Gulkis, M.G. Hauser, G. Hinshaw, P.D. Jackson, M. Janssen, E. Kaita, T. Kelsall, P. Keegstra, C. Lineweaver, K. Loewenstein, P. Lubin, J. Mather, S.S. Meyer, S.H. Moseley, T. Murdock, L. Rokke, R.F. Silverberg, L. Tenorio, R. Weiss, D.T. Wilkinson, *Astrophys. J.* 396 (1992) L1–L5.
- [597] M. Sneath, *Metrica* 49 (1999) 245–255.
- [598] R.S. Somerville, J.R. Primack, *Mon. Not. R. Astron. Soc.* 310 (1999) 1087–1110.
- [599] R.S. Somerville, G. Lemson, Y. Sigad, A. Dekel, G. Kauffmann, S.D.M. White, *Mon. Not. R. Astron. Soc.* 320 (2001) 289–306.
- [600] D.N. Spergel, P.J. Steinhardt, *Phys. Rev. Lett.* 84 (2000) 3760–3763.
- [601] V. Springel, N. Yoshida, S.D.M. White, *New Astron.* 6 (2001) 79–117.
- [602] A.A. Starobinsky, *Phys. Lett. B* 117 (1982) 175–178.
- [603] C.C. Steidel, K.L. Adelberger, M. Dickinson, M. Giavalisco, M. Pettini, M. Kellogg, *Astrophys. J.* 492 (1998) 428–438.
- [604] P.R.F. Stevenson, T. Shanks, R. Fong, H.T. MacGillivray, *Mon. Not. R. Astron. Soc.* 213 (1985) 953–969.
- [605] A.J. Stirling, J.A. Peacock, *Mon. Not. R. Astron. Soc.* 283 (1996) L99–L104.

- [606] M.A. Strauss, J.P. Huchra, M. Davis, A. Yahil, K.B. Fisher, J. Tonry, *Astrophys. J.* 83 (Suppl.) (1992) 29–63.
- [607] M.A. Strauss, J.A. Willick, *Phys. Rep.* 261 (1993) 271–431.
- [608] M.A. Strauss, in: A. Dekel, J.P. Ostriker (Eds.), *Formation of Structure in the Universe*, Cambridge University Press, Cambridge, 1999, p. 172, astro-ph/9610033.
- [609] A. Stuart, K. Ord, *Kendall's Advanced Theory of Statistics*, 6th Edition, Vol. 1, Edward Arnold, Paris, 1994.
- [610] A. Stuart, K. Ord, S. Arnold, *Kendall's Advanced Theory of Statistics*, 6th Edition, Vol. 2A, Arnold, Paris, 1999.
- [611] N. Sugiyama, *Astrophys. J.* 100 (Suppl.) (1985) 281–305.
- [612] W. Sutherland, H. Tadros, G. Efstathiou, C.S. Frenk, O. Keeble, S. Maddox, R.G. McMahon, S. Oliver, M. Rowan-Robinson, W. Saunders, S.D.M. White, *Mon. Not. R. Astron. Soc.* 308 (1999) 289–301.
- [613] Y. Suto, M. Sasaki, *Phys. Rev. Lett.* 66 (1991) 264–267.
- [614] Y. Suto, T. Matsubara, *Astrophys. J.* 420 (1994) 504–511.
- [615] A.S. Szalay, *Astrophys. J.* 333 (1998) 21–23.
- [616] A.S. Szalay, T. Matsubara, S.D. Landy, *Astrophys. J.* 498 (1998) L1–L4.
- [617] A.S. Szalay, B. Jain, T. Matsubara, R. Scranton, M.S. Vogeley, et al., *Astrophys. J.* (2001), submitted for publication, astro-ph/0107419.
- [618] I. Szapudi, A.S. Szalay, P. Boschan, *Astrophys. J.* 390 (1992) 350–358.
- [619] I. Szapudi, A.S. Szalay, *Astrophys. J.* 408 (1993) 43–56.
- [620] I. Szapudi, G.B. Dalton, G. Efstathiou, A.S. Szalay, *Astrophys. J.* 444 (1995) 520–531.
- [621] I. Szapudi, S. Colombi, *Astrophys. J.* 470 (1996) 131–148.
- [622] I. Szapudi, A. Meiksin, R.C. Nichol, *Astrophys. J.* 473 (1996) 15–21.
- [623] I. Szapudi, A.S. Szalay, *Astrophys. J.* 481 (1997) L1–L4;
I. Szapudi, A.S. Szalay, *Astrophys. J.* 515 (1999) L43–L43.
- [624] I. Szapudi, A.S. Szalay, *Astrophys. J.* 494 (1998) L41–L44.
- [625] I. Szapudi, *Astrophys. J.* 497 (1998) 16–20.
- [626] I. Szapudi, *Mon. Not. R. Astron. Soc.* 300 (1998) L35–L38.
- [627] I. Szapudi, E. Gaztañaga, *Mon. Not. R. Astron. Soc.* 300 (1998) 493–496.
- [628] I. Szapudi, A.S. Szalay, *Astrophys. Lett. J.* 515 (1999) L43–L43.
- [629] I. Szapudi, T. Quinn, J. Stadel, G. Lake, *Astrophys. J.* 517 (1999) 54–63.
- [630] I. Szapudi, S. Colombi, F. Bernardeau, *Mon. Not. R. Astron. Soc.* 310 (1999) 428–444.
- [631] I. Szapudi, S. Colombi, A. Jenkins, J. Colberg, *Mon. Not. R. Astron. Soc.* 313 (2000) 725–733.
- [632] I. Szapudi, E. Branchini, C.S. Frenk, S. Maddox, W. Saunders, *Mon. Not. R. Astron. Soc.* 318 (2000) L45–L50.
- [633] I. Szapudi, A.S. Szalay, *J. Appl. Probab.*, in press, astro-ph/9903132.
- [634] I. Szapudi, in: J.M. Fry, R. Buchler, H. Kandrup (Eds.), *The Onset of Nonlinearity in Cosmology*, Ann. NY. Acad. Sci. 927 (2001) 94–102, astro-ph/0008224.
- [635] I. Szapudi, M. Postman, T.R. Lauer, W. Oegerle, *Astrophys. J.* 548 (2001) 114–126.
- [636] I. Szapudi, S. Prunet, S. Colombi, *Astrophys. J.* 561 (2001) L11–L14.
- [637] I. Szapudi, J.A. Frieman, R. Scoccimarro, A.S. Szalay, et al., *Astrophys. J.* 570 (2002) 75–85.
- [638] H. Tadros, G. Efstathiou, *Mon. Not. R. Astron. Soc.* 276 (1995) L45–L50.
- [639] H. Tadros, G. Efstathiou, *Mon. Not. R. Astron. Soc.* 282 (1996) 1381–1396.
- [640] H. Tadros, W.E. Ballinger, A.N. Taylor, A.F. Heavens, G. Efstathiou, W. Saunders, C.S. Frenk, O. Keeble, R.G. McMahon, S.J. Maddox, S. Oliver, M. Rowan-Robinson, W. Sutherland, S.D.M. White, *Mon. Not. R. Astron. Soc.* 305 (1999) 527–546.
- [641] A. Taruya, K. Koyama, J. Soda, *Astrophys. J.* 510 (1999) 541–550.
- [642] A.N. Taylor, A.J.S. Hamilton, *Mon. Not. R. Astron. Soc.* 282 (1996) 767–778.
- [643] A.N. Taylor, W.E. Ballinger, A.F. Heavens, H. Tadros, *Mon. Not. R. Astron. Soc.* 327 (2001) 689–696.
- [644] A.N. Taylor, P.I.R. Watts, *Mon. Not. R. Astron. Soc.* 314 (2000) 92–98.
- [645] A.N. Taylor, P.I.R. Watts, *Mon. Not. R. Astron. Soc.* 328 (2001) 1027–1038.
- [646] M. Tegmark, A.N. Taylor, A.F. Heavens, *Astrophys. J.* 480 (1997) 22–35.
- [647] M. Tegmark, *Phys. Rev. D* 55 (1997) 5895–5907.
- [648] M. Tegmark, A.J.S. Hamilton, M.A. Strauss, M.S. Vogeley, A.S. Szalay, *Astrophys. J.* 499 (1998) 555–576.
- [649] M. Tegmark, P.J.E. Peebles, *Astrophys. J.* 500 (1998) L79–L82.
- [650] M. Tegmark, M. Zaldarriaga, A.J.S. Hamilton, *Phys. Rev. D* 63 (2001) 043007.

- [651] M. Tegmark, S. Dodelson, D. Eisenstein, V. Narayanan, R. Scoccimarro, R. Scranton, M. Strauss, et al., *Astrophys. J.* 571 (2002) 191–205.
- [652] M. Tegmark, A.J.S. Hamilton, Y. Xu, astro-ph/0111575, 2001.
- [653] H. Totsuji, T. Kihara, *Publ. R. Astr. Soc. Japan* 21 (1969) 221–229.
- [654] D.L. Tucker, A. Oemler, R.P. Kirshner, H. Lin, S.A. Shectman, S.D. Landy, P.L. Schechter, V. Muller, S. Gottlober, J. Einasto, *Mon. Not. R. Astron. Soc.* 285 (1997) L5–L9.
- [655] N. Turok, D.N. Spergel, *Phys. Rev. Lett.* 66 (1991) 3093–3096.
- [656] H. Ueda, J. Yokoyama, *Mon. Not. R. Astron. Soc.* 280 (1996) 754–766.
- [657] J.M. Uson, D.T. Wilkinson, *Astrophys. J.* 277 (1984) L1–L3.
- [658] P. Valageas, *Astron. Astrophys.* 354 (2000) 767–786.
- [659] P. Valageas, *Astron. Astrophys.* 356 (2000) 771–787.
- [660] P. Valageas, *Astron. Astrophys.* 379 (2001) 8–20.
- [661] P. Valageas, *Astron. Astrophys.* 382 (2002) 412–430.
- [662] P. Valageas, *Astron. Astrophys.* 382 (2002) 431–449.
- [663] P. Valageas, *Astron. Astrophys.* 382 (2002) 450–476.
- [664] L. van Waerbeke, *Astron. Astrophys.* 334 (1998) 1–10.
- [665] L. Van Waerbeke, F. Bernardeau, Y. Mellier, *Astron. Astrophys.* 342 (1999) 15–33.
- [666] 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, P. Schneider, *Astron. Astrophys.* 358 (2000) 30–44.
- [667] L. van Waerbeke, T. Hamana, R. Scoccimarro, S. Colombi, F. Bernardeau, *Mon. Not. R. Astron. Soc.* 322 (2001) 918–926.
- [668] G. Veneziano, *Phys. Lett. B* 406 (1997) 297–303.
- [669] L. Verde, A. Heavens, S. Matarrese, L. Moscardini, *Mon. Not. R. Astron. Soc.* 300 (1998) 747–756.
- [670] L. Verde, L. Wang, A. Heavens, M. Kamionkowski, *Mon. Not. R. Astron. Soc.* 313 (2000) 141–147.
- [671] L. Verde, A. Heavens, S. Matarrese, *Mon. Not. R. Astron. Soc.* 318 (2000) 584–598.
- [672] L. Verde, A. Heavens, *Astrophys. J.* 553 (2001) 14–24.
- [673] G. Vettolani, E. Zucca, R. Merighi, M. Mignoli, D. Proust, G. Zamorani, A. Cappi, L. Guzzo, D. Maccagni, M. Ramella, G.M. Stirpe, A. Blanchard, V. Cayatte, C. Collins, H. MacGillivray, S. Maurogordato, R. Scaramella, C. Balkowski, G. Chincarini, P. Felenbok, *Astron. Astrophys.* 130 (Suppl.) (1998) 323–332.
- [674] A. Vilenkin, E.P.S. Shellard, *Cosmic Strings and Other Topological Defects*, Cambridge University Press, Cambridge, 2000.
- [675] J.V. Villumsen, *Astrophys. J.* 71 (1989) 407–431.
- [676] J.V. Villumsen, *Mon. Not. R. Astron. Soc.* 281 (1996) 369–383.
- [677] J.V. Villumsen, M. Davis, *Astrophys. J.* 308 (1986) 499–509.
- [678] E.T. Vishniac, *Mon. Not. R. Astron. Soc.* 203 (1983) 345–349.
- [679] M.S. Vogeley, C. Park, M.J. Geller, J.P. Huchra, *Astrophys. J.* 391 (1992) L5–L8.
- [680] M.S. Vogeley, A.S. Szalay, *Astrophys. J.* 465 (1996) 34–53.
- [681] G.N. Watson, *A Treatise on the Theory of Bessel Functions*, 2nd Edition, Cambridge University Press, Cambridge.
- [682] P.I.R. Watts, A.N. Taylor, *Mon. Not. R. Astron. Soc.* 320 (2001) 139–152.
- [683] D.H. Weinberg, J.E. Gunn, *Mon. Not. R. Astron. Soc.* 247 (1990) 260–286.
- [684] D.H. Weinberg, S. Cole, *Mon. Not. R. Astron. Soc.* 259 (1992) 652–694.
- [685] S.D.M. White, M. Davis, G. Efstathiou, C.S. Frenk, *Nature* 330 (1987) 451–453.
- [686] S.D.M. White, M.J. Rees, *Mon. Not. R. Astron. Soc.* 183 (1978) 341–358.
- [687] S.D.M. White, *Mon. Not. R. Astron. Soc.* 186 (1979) 145–154.
- [688] S.D.M. White, in: R. Schaeffer, J. Silk, M. Spiro, J. Zinn-Justin (Eds.), *Cosmology and Large Scale Structure*, Les Houches Lectures, Session LX, Elsevier Science, Amsterdam, 1996.
- [689] M. White, *Mon. Not. R. Astron. Soc.* 310 (1999) 511–516.
- [690] S.S. Wilks, *Mathematical Statistics*, 2nd Edition, Wiley, New York, 1963.
- [691] J.A. Willick, *Astrophys. J.* 530 (2000) 80–95.
- [692] M.B. Wise, in: W.G. Unruh, G.W. Semenoff (Eds.), *The Early Universe*, Vol. 215, Reidel, Dordrecht, 1988.
- [693] D.M. Wittman, J.A. Tyson, D. Kirkman, I. Dell’Antonio, G. Bernstein, *Nature* 405 (2000) 143–148.

- [694] E.L. Wright, S.S. Meyer, C.L. Bennett, N.W. Boggess, E.S. Cheng, M.G. Hauser, A. Kogut, C. Lineweaver, J. Mather, G.F. Smoot, R. Weiss, S. Gulkis, G. Hinshaw, M. Janssen, T. Kelsall, P. Lubin, S.H. Moseley, T. Murdock, R.A. Shafer, R.F. Silverberg, D.T. Wilkinson, *Astrophys. J.* 396 (1992) L13–L18.
- [695] G. Xu, *Astrophys. J.* 98 (Suppl.) (1995) 355–366.
- [696] J. Yang, M.S. Turner, D.N. Schramm, G. Steigman, K.A. Olive, *Astrophys. J.* 281 (1984) 493–511.
- [697] T. Yano, N. Gouda, *Astrophys. J.* 487 (1997) 473–481.
- [698] T. Yano, N. Gouda, *Astrophys. J.* 495 (1998) 533–538.
- [699] D. York, et al., *Astron. J.* 120 (2000) 1579–1587.
- [700] N. Yoshida, V. Springel, S.D.M. White, G. Tormen, *Astrophys. J.* 544 (2000) L87–L90.
- [701] N. Yoshida, J. Colberg, S.D.M. White, A.E. Evrard, T.J. MacFarland, H.M.P. Couchman, A. Jenkins, C.S. Frenk, F.R. Pearce, G. Efstathiou, J.A. Peacock, P.A. Thomas, *Mon. Not. R. Astron. Soc.* 325 (2001) 803–816.
- [702] M. Zaldarriaga, R. Scoccimarro, L. Hui, astro-ph/0111230, 2001.
- [703] S. Zaroubi, Y. Hoffman, *Astrophys. J.* 462 (1996) 25–31.
- [704] I. Zehavi, M.R. Blanton, J.A. Frieman, D.H. Weinberg, H.J. Mo, M.A. Strauss, et al., *Astrophys. J.* 571 (2002) 172–190.
- [705] Ya.B. Zel’dovich, *Astron. Astrophys.* 5 (1970) 84–89.
- [706] Ya.B. Zel’dovich, *Mon. Not. R. Astron. Soc.* 160 (1972) 1P–3P.
- [707] I. Zlatev, I. Wand, P.J. Steinhardt, *Phys. Rev. Lett.* 82 (1999) 896–899.
- [708] W.H. Zurek, P.J. Quinn, J.K. Salmon, M.S. Warren, *Astrophys. J.* 431 (1994) 559–568.
- [709] F. Zwicky, *Publ. Astron. Soc. Pac.* 50 (1938) 218–220.
- [710] F. Zwicky, E. Herzog, P. Wild, M. Karpowicz, C.T. Kowal, *Catalogue of Galaxies and Clusters of Galaxies* (6 volumes), California Institute of Technology, Pasadena, 1961–1968.