

Robust Estimation Techniques for the Cosmological Analysis of Large Scale Structure



Oliver Leicht

Department of Applied Mathematics and Theoretical Physics
University of Cambridge

This dissertation is submitted for the degree of
Doctor of Philosophy

Fitzwilliam College

September 2020

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification at the University of Cambridge, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified below and in the text.

Chapter 3 is based on work done in collaboration with Cora Uhlemann, Francisco Villaescusa-Navarro, Sandrine Codis, Lars Hernquist and Shy Genel. I am the first author of the resulting paper [1], made major contributions to the result section and compiled the manuscript. In detail, I extended and improved the existing LSSFAST code by Sandrine Codis and Cora Uhlemann to be able to efficiently construct bias relations and to measure statistical uncertainties. Using the code, I obtained the results concerning the one-point probability density functions and bias relations.

Chapter 4 and 5 are based on work done in collaboration with Tobias Baldauf, James Fergusson and Paul Shellard. I am the first author of the corresponding paper [2], developed and implemented the code for the analysis, wrote the paper and produced all figures and tables.

Oliver Leicht
September 2020

Acknowledgements

First and foremost, I would like to thank my supervisor Paul Shellard for giving me the opportunity to work in his group.

I am grateful to Tobias Baldauf who supported me when I was struggling and from whom I learned a great deal about physics.

I am indebted to my collaborators Cora Uhlemann, Sandrine Codis, James Fergusson and Francisco Villaescusa-Navarro for insightful conversations. I want to thank the Institut d'Astrophysique de Paris for its hospitality.

Special thanks go to friends and colleagues, in particular Muntazir Abidi, Carlos d'Aragona Balhana, Thomas Boehmer, Mark Daniels, Konradin Eigler, Nikzad Falhati, Sascha Feldmann, Pavao Santak and Petr Zapletal for enriching discussions and their support during these years as a PhD student.

I am also grateful to the Cambridge Trust, Science and Technology Facilities Council, Studienstiftung des deutschen Volkes, and Fitzwilliam College for funding me.

Lastly, I want to thank my family and Mariam for continuous and unwavering support.

Abstract

The Λ CDM model of cosmology together with Inflation has had tremendous success over the past 30 years in explaining the increasingly rich data sets of the cosmic microwave background and large-scale structure. The next generation of large-scale structure surveys is expected to answer many open questions about the microscopic description of the Universe. In order to fully leverage those data sets, one needs exquisite theoretical predictions. Here, the main difficulty is the non-linear nature of the large-scale structure observables which, together with the exquisitely small statistical errors, cause real concern of false discoveries. In this thesis, we study two estimators that allow us to extract non-linear information from the large-scale structure while being robust against one of the leading sources of systematic uncertainties: Redshift-space distortions.

By means of a bias relation, we extend the matter counts-in-cells statistic for the first time to neutral hydrogen. Neutral hydrogen is particularly interesting for counts-in-cells statistics because of the vast regions that can be covered by intensity mapping. We find percent-level accuracy when comparing the prediction for the density in spheres probability density function to the IllustrisTNG simulation. The measured density dependent clustering signal, which could be used to break the bias-amplitude degeneracies, matches theoretical expectations. Our bias model is able to capture the effect of redshift-space distortions making the estimator robust.

Based on a separation idea, we present an efficient code to compute projected bispectra. The separation approach is orders of magnitude more efficient than the direct integration. This allows us to investigate the relation between biases in the estimated parameters and inaccurate modelling of non-linear redshift-space distortions for the power spectrum and bispectrum of projected galaxy density fields and lensing convergence. For a toy galaxy survey that resembles the CMASS sample of the baryon oscillation spectroscopic survey, we find that modelling non-linear redshift-space distortion only becomes necessary for galaxy bins thinner than $150 h^{-1}\text{Mpc}$. In case a better radial resolution is available, errors on cosmological parameters can be improved by 20% when including an accurate non-linear RSD model that allows us to use bins of depth $\sim 60 h^{-1}\text{Mpc}$. The separation of projection integrals proves also useful for theoretical uncertainties. We use the separability of Gaussian correlation

functions to develop a consistent model for theoretical uncertainties of the projected power spectrum.

Table of contents

List of figures	xiii
List of tables	xxiii
1 Introduction	1
1.1 Historical overview	1
1.2 Open questions, surveys and methods	2
1.3 Outline	3
2 Background	5
2.1 The homogeneous Universe	5
2.1.1 General Relativity and cosmological solutions	5
2.1.2 Cosmic ingredients and timeline	7
2.1.3 Redshift and distances	9
2.2 Initial conditions from Inflation	11
2.2.1 Inflation	11
2.2.2 Gauge problem in GR	13
2.2.3 Initial conditions from Inflation	15
2.2.4 Primordial non-Gaussianities	17
2.2.5 Connecting early and late times	19
2.3 Newtonian Perturbation Theory	21
2.3.1 Evolution equations in co-moving coordinates	21
2.3.2 Eulerian perturbation theory	22
2.3.3 Redshift-space distortions	27
2.3.4 Lagrangian perturbation theory	31
2.3.5 Redshift-space distortions in LPT	33
2.3.6 Tracer bias	34
2.3.7 Primordial non-Gaussianities	37

2.4	Halo model	39
2.4.1	Spherical collapse	40
2.4.2	Filtering of fields	42
2.4.3	Mass functions	43
2.4.4	Biasing in the halo model	45
2.4.5	Scale-dependent bias from PNGs	45
2.5	Counts-in-cells	48
2.5.1	Counts-in-cells statistics for concentric spheres	49
2.5.2	Counts-in-Cells statistics in the large separation limit	54
3	Extreme Spheres: Counts-in-cells for 21cm intensity mapping	61
3.1	Introduction	62
3.2	Numerical simulation	64
3.3	Statistics of matter densities in spheres	67
3.3.1	One-point PDF of matter density	68
3.3.2	Density-dependent clustering of matter	69
3.4	Statistics of tracer densities in spheres	73
3.4.1	One-point PDF of tracer density	74
3.4.2	Density-dependent clustering of tracers	75
3.4.3	Parametrisation-independent bias functions	75
3.4.4	Polynomial bias model in log-densities	76
3.5	Results	77
3.5.1	Joint density distributions and mean bias model	77
3.5.2	One-point PDF of neutral hydrogen	86
3.5.3	Density-dependent clustering of neutral hydrogen	86
3.6	Conclusions and Outlook	87
4	Computations for projected two- and three-point statistics	91
4.1	Projected power- and bispectra	92
4.1.1	Projections	92
4.1.2	Projected power spectrum	94
4.1.3	Projected bispectrum	95
4.2	Statistics	97
4.2.1	Power spectrum	97
4.2.2	Power spectrum signal-to-noise analysis	100
4.2.3	Bispectrum	102
4.3	Theoretical uncertainties	106

4.3.1	Theoretical uncertainties in three dimensions	106
4.3.2	Projected theoretical uncertainties	109
4.3.3	Empirical validation	111
4.4	Efficient bispectrum projection integrals	111
4.4.1	Separating the bispectrum	113
4.4.2	Analytic results for the CMB-lensing function	114
4.4.3	Analytic results for Gaussian selection functions	117
4.4.4	Symmetries of a survey with identical and thin bins	117
4.5	Outlook: Signal compression	118
5	Projected two- and three-point statistics: Forecasts and mitigation of non-linear RSDs	121
5.1	Introduction	122
5.2	Theory predictions	123
5.2.1	Statistics in 3D	123
5.2.2	Projected statistics and survey specifications	124
5.3	Inference methods	126
5.3.1	Fisher forecasting	126
5.3.2	Parameter estimation	127
5.4	Results	128
5.4.1	Choosing the cut-off scale for projected spectra	128
5.4.2	Signal-to-noise analysis	130
5.4.3	Error - bias tradeoff	132
5.4.4	Optimistic forecasts	139
5.4.5	A simple signal compression approach	140
5.5	Conclusion	142
6	Conclusion	147
	References	151

List of figures

2.1	Hubble diagram from [3]. The logarithm of the (luminosity) distance is shown as a function of the redshift. The black line represents the best-fit Λ CDM cosmology with fixed Hubble parameter today, $H(z = 0) = 70\text{km s}^{-1}\text{Mpc}^{-1}$. The residuals in the bottom panel indicate that the fit, and thus the theory, is fully consistent with the observations.	10
2.2	Left: The Poisson factor at redshift zero is shown. The scale dependence is dominated by the k^2 term from the Poisson equation. In addition, the slope changes around the scale of matter-radiation equality $k_{\text{eq}} \simeq 0.015 h\text{Mpc}^{-1}$, since small scales are suppressed due to the Meszaros effect. Right: Linear power spectrum at redshift zero (blue line) compared to the HALOFIT [4, 5] prediction for the non-linear power spectrum (orange). The non-linearities are strongest on small scales and stem from an enhanced clustering compared to the linear predictions.	20
2.3	Left: Growth factor normalised to one today, in a Λ CDM (blue) and Einstein-deSitter (2.83), (orange) Universe. During matter-domination, they are parallel and only started deviating when the cosmological constant became relevant and the growth of structure started to stall which is not accounted for in Einstein-deSitter. Right: The logarithmic growth rate in a Λ CDM cosmology (blue) is decreasing at late times, while it remains constant in an Einstein-deSitter Universe (orange).	24

- 2.4 Two-point correlation function measured in the 2dF galaxy redshift survey [6]. The colour is related to the clustering strength and increases from black to yellow. π measures radial and σ transverse distances. The black lines show equi-correlation lines and even in the absence of RSDs they would not be circular due to the Alcock-Paczynski effect. One would need perfect knowledge of the underlying cosmology to have consistent transverse and radial distance scales. On large scales, the Kaiser effect leads to a squashing because correlations parallel to the LOS are enhanced. On small scales, the FoG effect suppresses correlations parallel to the LOS and thus the equi-correlation lines appear stretched in the radial direction. 27
- 2.5 The full RSD power spectrum is shown as a function of the wave vector k , and the cosine of the angle with the LOS, μ . On large scales, the Kaiser effects enhance modes parallel to the LOS while on small scales, those modes are damped by the FoG effect. 29
- 2.6 Left: The primordial bispectrum of the three commonly used templates is orders of magnitudes smaller than the gravitational bispectrum at $z = 0$ for $f_{\text{NL}} = 1$. We show the ratio of the local (green curve), equilateral (blue) and orthogonal (orange) templates and the tree-level gravitational bispectrum in the squeezed configurations $(k, k, 0.1k)$. Similarly in equilateral configurations, where the three templates are identical, the ratio is still very small (red curve). Right: The scale-dependent bias generates a characteristic feature at large scales. The equilateral template does not give rise to this effect. 38
- 2.7 Left: Evolution of the radius of a sphere following the background evolution (blue curve), the linear perturbative solution (green) and the fully non-linear spherical collapse (orange). Middle: Time evolution of the density in a sphere assuming linear evolution (blue curve) and spherical collapse (orange). The linear density approaches the critical density δ_c , when the spherical collapse solution becomes singular. Right: Illustration of the mapping between the density in a sphere from spherical collapse and the linear overdensity. We show the exact mapping in blue and the orange points show the approximation (2.178). 42
- 2.8 Lagrangian bias parameters from the Sheth-Tormen mass function. In order to have dimensionless quantities, we scale where necessary with $k_l = 0.01 h\text{Mpc}^{-1}$. All parameters are roughly constant and $O(1)$ for tracers of masses $\leq 10^{14}M_\odot$ and grow rapidly afterwards. 46

- 2.9 Probability density function for the density in spheres of radius $10 h^{-1}\text{Mpc}$ at redshifts $z = 1.36, 0.97, 0.65$ and 0 (from light to dark blue). The points are measurements in an N-Body simulation and one sees that the skewness increases as the redshift decreases. The saddle-point approximation with the density breaks down for $\rho \geq 2$, because the rate function ceases to be convex. In contrast, the saddle-point approximation of the log-densities approximates the exact numerical inverse Laplace transform well and matches the measurements. Figure from [7]. 53
- 2.10 The conceptional steps towards a joint PDF for distant concentric spheres. Starting from the cumulant-generating function for $n + 1$ spheres, a), we derive the cumulant-generating function for n concentric spheres and one distant sphere, b). Using perturbation theory, one can then extend this to the general case with two distant sets of concentric spheres. Figure from [8]. 57
- 3.1 Sketch of simulation measurements. Spatial distribution of neutral hydrogen in a $5 h^{-1}\text{Mpc}$ slice of the TNG100 simulation at redshift $z = 3$. The color is showing the numerical value of the HI column density in co-moving units. We also show a small subset of the spheres with radius $R = 5 h^{-1}\text{Mpc}$ used for counts-in-cells statistics. 64
- 3.2 Dark matter PDF showing percent-level agreement between simulation and prediction. Upper panel: Measured PDF of matter densities in spheres of radius $R = 5 h^{-1}\text{Mpc}$ at redshifts $z = 0, 1, 2, 3, 4, 5$ (red to blue data points) with measured non-linear variance as input (solid lines), compared to the prediction from large-deviation statistics. For clarity, error bars are omitted and only shown in the lower panel. Lower panel: Residuals between the theoretical predictions with measured variance and the measured PDFs. 70
- 3.3 Measured density-dependent matter clustering agrees with theoretical expectations. Upper panel: Density-dependent clustering encoded in the sphere bias function from equation (2.232) for matter at redshifts $z = 0$ to 5 (red to blue) as measured in the IllustrisTNG simulation (data points) and predicted by large-deviation statistics using equation (3.3) (lines). Lower panel: Residuals between the theoretical prediction with measured variance and the measured density-dependent clustering. For better visibility, the abscissas for different redshifts are slightly shifted. 71

- 3.4 HI vs. matter scatter plots in real- and redshift-space. Density scatter plots of the neutral hydrogen density ρ_{HI} in real-space (blue-green) and redshift-space (grey) versus the matter density in real-space ρ_{m} for radius $R = 5 h^{-1}\text{Mpc}$ at redshifts $z = 1, 3, 5$ (left to right). The colour (from light to dark) indicates the magnitude of the joint PDF in logarithmic scale and shows a concentration of probability around the mean bias. The figure also shows the parametrisation-independent bias obtained from the CDF in real-space (solid red line) and redshift-space (dashed red line). The main impact of redshift-space distortions is to increase the scatter while the mean bias relation is almost unchanged for average densities and mostly affected in the positive density tails. 78
- 3.5 HI vs. halo mass scatter plots in real-space. Density scatter plots of the real-space neutral hydrogen density ρ_{HI} versus the mass-weighted halo density ρ_{HM} for radius $R = 5 h^{-1}\text{Mpc}$ at redshifts $z = 1, 3, 5$ (left to right). The colour (from light to dark) indicates the magnitude of the joint PDF in logarithmic scale and shows a concentration of probability around the mean bias. The figure also shows the parametrisation-independent bias obtained from the CDF (solid orange line) and a linear fit in log-densities (dotted orange line). This plot shows that the neutral hydrogen closely traces the mass in halos with an almost linear relationship. 79
- 3.6 Necessity of non-linear bias and beyond log-normal prediction. Comparison of measured density PDFs of matter (black), mass-weighted halos (blue) and neutral hydrogen (red) in spheres of radius $R = 5 h^{-1}\text{Mpc}$ at redshift $z = 3$ in real-space. The abscissa shows scaled log-densities such that for tracers that are linearly biased in log-densities curves would overlap. The lack of overlap between tracers and matter motivates the chosen quadratic bias model (3.13). Also shown are log-normal fits to the PDFs (dashed lines) which show significant deviations from the measurements for underdensities of matter and neutral hydrogen. 80

- 3.7 Comparison of bias functions for fully-predictive and measured matter PDF. Upper panel: Bias functions for spheres of radius $R = 5 h^{-1}\text{Mpc}$ are displayed at redshifts $z = 1, 3, 5$ (orange to blue). The parametrisation-independent bias functions from the cumulative PDFs according to equation (3.10) using the measured HI PDF in real- and redshift-space with the measured real-space matter PDF are displayed in solid and dashed lines, respectively. Dotted and dot-dashed lines plot the bias function when using the fully-predictive matter PDF and the measured HI PDF in real- and redshift-space respectively. Note the axes are swapped compared to the scatter plots (Fig. 3.4). Middle and lower panel: Residuals between the parametrisation-independent bias functions for measured dark matter and measured HI (mm) and fully-predictive matter distribution together with measured HI (pm) in real-space (middle) and redshift-space (lower panel). 82
- 3.8 Accuracy of quadratic parametrisation for the bias function. Upper panel: All bias functions plot matter densities as a function of the neutral hydrogen densities in spheres of radius $R = 5 h^{-1}\text{Mpc}$ at redshifts $z = 1, 3, 5$ (orange to blue). The quadratic bias model (3.13) with the fitted bias parameters from Table 3.3 in real-space (dotted lines) and redshift-space (dot-dashed lines) is compared to the measured bias function from Fig. 3.7 in real-space (solid lines) and redshift-space (dashed lines). Note that the axes are swapped compared to the scatter plots (Fig. 3.4). Middle and lower panel: Residuals between quadratic model (mb_n) and the measured bias function (mm) from the upper panel in real-space (middle) and redshift-space (lower panel). 83
- 3.9 Real-space HI PDF showing a few-percent level agreement between simulation and prediction. Upper panel: Measured PDF of HI densities in spheres of radius $R = 5 h^{-1}\text{Mpc}$ for redshifts $z = 1, 3, 5$ compared to the prediction from large-deviation statistics for matter combined with the quadratic log-bias model (dashed lines) for neutral hydrogen densities in real-space. Lower panel: Residuals between the theoretical predictions and the measured PDFs of HI densities in spheres in real-space. 84

- 3.10 Redshift-space HI PDF showing a few-percent level agreement between simulation and prediction. Upper panel: Measured PDF of HI densities in spheres of radius $R = 5 h^{-1}\text{Mpc}$ for redshifts $z = 1, 3, 5$ compared to the prediction from large-deviation statistics for matter combined with the quadratic log-bias model (dashed lines) for neutral hydrogen densities in redshift-space. Lower panel: Residuals between the theoretical predictions and the measured PDFs of HI densities in spheres in redshift-space. 85
- 3.11 Measured density-dependent HI clustering agrees with theoretical expectations. Upper panel: Density-dependent clustering encoded in the sphere bias at redshifts $z = 1, 3, 5$ (orange to blue) for neutral hydrogen in real-space as predicted from equation (3.7) (dashed lines) and measured (data points). Lower panel: Residuals between the theoretical prediction and simulation measurements. For better visibility, the abscissas for different redshifts are slightly shifted. 87
- 4.1 The projected density field at each wave vector \mathbf{k}_\perp is a superposition of all wave vectors with perpendicular component \mathbf{k}_\perp and the weights set by a window function that depends on the LOS component k_\parallel of the wave vector. 93
- 4.2 Linear power spectrum (blue curve) and its 2D projections with a Gaussian kernel of varying depth (orange to brown). The projections change both the overall amplitude and the power spectrum's shape at low k . Instead of the k^{n_s} scaling at low k , the projected power spectra become constant. 95
- 4.3 The four columns show the four different (cross-)bispectra in the squeezed configuration (first two rows) and the equilateral configuration (bottom row). Where possible, we show the 3D bispectrum (blue) and projections with a Gaussian kernel of different depths (orange to brown). The lensing-lensing-clustering bispectra all lie onto each other. 96
- 4.4 Unmarginalized 1σ error bars without theoretical uncertainties (green line) and with theoretical uncertainties (orange and blue) as a function of the upper cut-off k_{max} . The error bars of the cosmological parameters are decreasing with k_{max} without theoretical uncertainties. Using theoretical uncertainties, the error bars saturate. For a smaller correlation length in the correlation function, the saturation happens at a larger scale (blue curve) than with the larger correlation length (orange). The results come from a power spectrum analysis of a cubic survey of side length $1000 h^{-1}\text{Mpc}$ at redshift $z=0.57$ 107

- 4.5 Relative importance of the one-loop envelope to the linear power spectrum in 3D (blue curve) and for various projected power spectra (orange to brown). The ratio in 3D follows a power law, whereas the ratios of the projected power spectra become constant at low k . At large k , the ratios of the projected power spectra approach the 3D value from above. 110
- 4.6 The ratio of (unmarginalized) error bars in 3D and 2D, $\sigma_x^{3D}/\sigma_x^{2D}$, is shown as a function of k_{\max}^{2D} . We use the saturated value for the 3D uncertainty and refer to Fig. 4.4 where we studied the saturation in detail. For each parameter (row) and projection depth (column), the 2D error bar saturates due to theoretical uncertainties. As the number of tomographic bins increases (from left to right), this saturated value approaches the 3D value. 112
- 5.1 We use a galaxy survey of fixed length d consisting of n identical tomographic bins along the LOS. The centre of each bin is given by (5.8). Beside the Gaussian windows that are displayed here, we also use Top-Hat windows since they cover the survey volume evenly. 125
- 5.2 Overview of the windows used in this chapter. The CMB-lensing kernel (blue) peaks halfway between the observer and the surface of last scattering at χ_s . The orange curve shows the lensing kernel multiplied with the growth factor. We show in green the sum of three galaxy bins with width and distance between their centres of $200 h^{-1}\text{Mpc}$ centred around $z=0.57$. The combined galaxy window shown here is constructed to resemble the CMASS galaxy sample. 125
- 5.3 Relation between k_{\max} and maximal relative bias for a varying number of redshift bins for a CMASS-like survey (see Table 5.1a). The points with the black border are closest to 0.2 for each configuration and thus used as cut-offs in this chapter. Their numerical values can be found in Table 5.1b. The 3D curve was determined from a cubic survey of side length $l = (A \cdot d)^{1/3}$ with the same redshift and shot-noise as the 2D setting. 128
- 5.4 The ratio of unmarginalised error bars in two and three dimensions is shown for the galaxy power spectrum (dashed), galaxy bispectrum (dotted) and combined (solid lines). With an increasing number of tomographic bins, the 2D error bars first decrease due to the increase in resolved radial information. Around 10 bins, the decrease in the cut-off k_{\max}^{2D} takes over and the error increases again. 130

5.5	Signal-to-noise scaling for different power spectra and bispectra as a function of the number of tomographic bins. The dependence has three components: 1) With more tomographic bins, more radial information is resolved. 2) The cut-off, $k_{\text{max}}^{2\text{D}}$, decreases with the number of tomographic bins which in turn decreases the SN. 3) The galaxy selection function is changing when varying the number of Gaussian profiles. This effect is negligible from four bins onward.	131
5.6	Marginalised 1σ error bars on cosmological parameters as a function of the projection depth for a selection of cross-spectra are displayed in solid lines. The Fisher forecasts assume FoG Model 2, but are actually independent of the chosen FoG model. The corresponding dashed lines represent five times the bias due to inaccurate theoretical modelling for the FoG model that misspecifies the velocity dispersion by 50%. All forecasted errors are minimal at $\sim 60 h^{-1}\text{Mpc}$	132
5.7	The 1σ contour plots for four different spectra using FoG Model 2 (velocity dispersion underestimated by 50%). The bins are chosen to minimise the errors controlling for systematic uncertainties. The optimal number of bins are reported in Table 5.2.	135
5.8	The marginalised 1σ error bars on PNGs as a function of the projection depth for a selection of cross-spectra are displayed in solid lines. The Fisher forecasts assume FoG Model 2, but are almost independent of the chosen FoG model. The corresponding dashed lines represent five times the bias due to inaccurate theoretical modelling for the FoG model that misspecifies the velocity dispersion by 50%. Each subplot corresponds to a different forecasts and we marginalised over the bias parameters.	136
5.9	1σ contour plots for an f_{NL} -bias forecast combining all power spectra and bispectra and using FoG Model 2. The bins are chosen to minimise the errors controlling for systematic uncertainties. Each colour corresponds to an independent forecast.	139

- 5.10 Error bars (solid lines) and five times the biases (dashed lines) for joint clustering-lensing spectra as a function of the maximal correlation length using FoG Model 2. The rightmost values in each subplot correspond to the full analysis and the leftmost values to an analysis with only the auto-spectra. The intersection of dashed and solid lines marks the point where the relative bias becomes 20%. The top two rows correspond to a power spectrum analysis and the bottom two rows show an analysis with all power spectra and bispectra. Each subplot shows the scaling for a different parameter. 143
- 5.11 Similar as Fig. 5.10 but for the $\delta_g \delta_g$ (top two rows) and gal combination (bottom two rows) of spectra. 144

List of tables

3.1	The measured non-linear variances σ of the log-density $\mu = \log \rho$ of matter (m) in spheres of $R = 5 h^{-1} \text{Mpc}$ at redshifts $z = 0$ to 5. They give an indication of how well the fully-predictive matter theory works, as it requires $\sigma_{\mu,m} < 1$ to converge.	69
3.2	Statistical properties of neutral hydrogen distribution in spheres of $R = 5 h^{-1} \text{Mpc}$ at redshifts $z = 1, 3, 5$. The measured non-linear variances σ of the log-density $\mu = \log \rho$ of neutral hydrogen (HI) in real and redshift-space, shown in the left two columns, can be compared to the matter variances in Table 3.1 to estimate the linear bias. The linear cross-correlation coefficients (3.12) between matter (m), neutral hydrogen in real-space (HI) and redshift-space (HIz) and mass-weighted halos (HM), shown in the right four columns, indicate a tight relation between the different fields.	73
3.3	Collection of the bias parameters obtained from fitting the quadratic bias model from equation (3.13) to the predicted bias function (measured HI and predicted matter) obtained from the CDF according to equation (3.10) both in real (three left columns) and redshift-space (three right columns). The ratio between matter and HI log-variances in Tables 3.1 and 3.2 allows us to cross-check the linear bias parameter b_1	84
5.1	(a): Characterisation of the CMASS-like survey we use in this chapter. The velocity dispersion parameters were obtained from fits against N -body simulations. (b): Cut-offs for different projection depths that ensure all relative biases in the matter predictions are below 20% for a CMASS-like survey described on the right side. The 3D cut-off comes from a cubic survey with the same volume. The values correspond to the black points in Fig. 5.3.	129

5.2	Optimal forecasted relative errors for three FoG models considered. For each spectra, we give the number of tomographic bins such that the maximal relative bias is below 20% and the statistical uncertainties are minimal. On the left side, we report the marginalised relative 1σ uncertainties. On the right side, we show the corresponding relative biases.	134
5.3	Optimal forecasted errors for a bias- f_{NL} analysis with fixed cosmological parameters. We present separate forecasts for the three FoG models considered. For each spectra, we give the number of tomographic bins such that the maximal relative bias is below 20% and the statistical uncertainties are minimal. On the left side, we report the marginalised 1σ uncertainties. On the right side, we show the corresponding relative biases.	137
5.4	Forecasts for the relative 1σ uncertainties of cosmological parameters in a CMASS-like survey with $k_{\text{max}} = 0.15 h^{-1}\text{Mpc}$ and 6 tomographic bins without (left side) and with (right) CMB prior.	140
5.5	Forecasts for the 1σ uncertainties of f_{NL} and bias parameters in a CMASS-like survey with $k_{\text{max}} = 0.15 h^{-1}\text{Mpc}$ and 6 tomographic bins.	141
5.6	The maximal correlation lengths needed to recover 99% of the Fisher information for all cosmological parameters compared to the full analysis. The results are independent of the FoG model. All entries are in $h^{-1}\text{Mpc}$	142

Chapter 1

Introduction

1.1 Historical overview

Physical cosmology in the modern sense started in 1917 when Einstein proposed a cosmological solution to the field equations of General Relativity (GR). In particular, he introduced a cosmological constant to achieve a static solution. It was later shown that this solution was unstable. Instead, Friedmann found that cosmological solutions of GR are initially expanding [9]. Lemaître confirmed these theoretical findings and provided the first empirical evidence for the expansion in 1927 [10]. Two years later, Hubble measured the distance-redshift relation that confirmed the expanding Universe [11].

In the subsequent forty years, additional evidence further substantiated the observation that the Universe was not only expanding but was initially very ‘small’. This Big Bang picture was in opposition to the paradigm of a static universe, the so-called steady-state Universe. The discovery of the cosmic microwave background (CMB) in 1965 by Bell lab researchers Penzias and Wilson [12] led to a universal acceptance of the Big Bang picture. While being superior to the steady-state picture, Big Bang cosmology had its own shortcomings: Namely, the necessity of very specific initial conditions in order to solve the horizon- and flatness problem. In 1981, Alan Guth proposed a dynamical mechanism, called Inflation, that would solve both problems and as it later turned out, could provide the seeds for the rich non-linear structure we observe today too [13]. Until today, Inflation is the leading paradigm for the first fraction of a second of the Universe.

Studying the Coma cluster led Swiss astronomer Zwicky to postulate the existence of dark matter [14] for which Vera Rubin found convincing evidence in the 1980s [15, 16]. The current data suggests that dark matter is only interacting gravitationally and moves at non-relativistic velocities. This is why it is referred to as cold dark matter. The current evidence suggests that there is four times as much dark as baryonic matter in the Universe.

By the end of the last millennium, it was discovered that the Universe’s expansion has begun to accelerate recently [17, 18]. The cosmological constant, driving the recent accelerated expansion, accounts today for two thirds of the energy density in the Universe. This gave rise to the current standard model of cosmology, the Lambda cold dark matter (Λ CDM) model. Λ CDM has been remarkably successful in explaining a variety of observations ranging from the temperature anisotropies in the CMB [19, 20] to Large Scale Structure of the Universe (LSS) observables like the local expansion history [21] and the Baryon Acoustic Oscillations (BAO) in galaxy surveys [22].

1.2 Open questions, surveys and methods

Despite the success of Λ CDM cosmology, there are many fundamental questions remaining: Is General Relativity the correct framework? What are the microscopic origins of the cosmological constant and dark matter? Did Inflation happen and what dynamics were driving it? What are the masses of neutrinos?

In the case of the neutrino masses and inflation there are clear targets for the size of uncertainties when conclusions can be made. For neutrinos, it is known from flavor oscillations of solar and atmospheric neutrinos that they have a non-zero mass. However, the underlying mechanism that generates the neutrino masses is unclear. Cosmological observables can constrain the sum of the neutrino masses. This measurement will, for instance, help deciding whether the masses follow a ‘normal’ or ‘inverted’ hierarchy [23]. Here, the required sensitivity for the total neutrino masses is $\sigma(\sum m_\nu) \sim 14 \text{ meV}$ [24]. The dynamics of inflation are imprinted in the initial conditions of the classical evolution of the Universe. In particular, primordial non-Gaussianities (PNGs) (or their absence) will shed a light on the dynamics of inflation. Here, the target is whether or not the amplitude of PNGs of the local, equilateral or orthogonal shape is larger or smaller than 1 [25]. We discuss the implications of a (non-)discovery of the different shapes in section 2.2.4.

Future LSS surveys are expected to help answer all the mentioned questions for two reasons: First, they allow cosmologists to test the expansion history and growth of structure during matter and Λ domination directly. Second, due to the three-dimensional nature of the LSS, these surveys will eventually contain more information than the surveys of the two-dimensional CMB that are currently providing the best constraints. There are several ways to observe the LSS. Upcoming galaxy surveys, like the Vera C. Rubin Observatory (formerly known as LSST) [26], SPHEREx [27], Euclid [28] and DESI [29], are going to map the positions of billions of galaxies that trace the LSS. Intensity mapping (of neutral hydrogen) surveys like Tianlai [30], BINGO [31], CHIME [32], FAST [33], HIRAX [34], MeerKAT

[35], SKA [36] and SPHEREx [27] are sampling vast regions of the post-reionisation Universe. In addition, CMB experiments such as the Simons Observatory [37] and the next generation CMB-S4 [38] enable measurement of the integrated effect of the LSS on CMB photons via (weak-) gravitational lensing. Cross-correlations between those probes offers synergies, for instance by breaking degeneracies between amplitude and bias parameters, and enabling use of cosmic variance cancellation techniques.

The LSS is a highly non-linear structure and contains a wealth of information beyond the traditional two-point or power spectrum analyses. There are two approaches to extract this additional information: In the forward modelling, or reconstruction, approach, one is trying to approximate the initial conditions that gave rise to the observations. This offers a route to modelling all non-Gaussian effects directly but in practice one has to rely on powerful numerical tools whose overall behavior is hard to fully understand. This could cause subtle biases in the final estimates. The late time approach, in contrast, develops estimators for certain non-Gaussian aspects of the LSS directly. While they are easier to understand, they do not necessarily capture all available information. We follow the latter approach.

1.3 Outline

In this thesis, we present work on two estimators that allow us to extract non-Gaussian information from the LSS while being robust against redshift-space distortions, a leading source of systematic uncertainty. The thesis is structured as follows:

- In chapter 2, we provide a background to the formalism used in the subsequent chapters. We start by briefly reviewing General Relativity. Then we discuss in more detail a toy model of Inflation and phenomenological templates for primordial non-Gaussianities. Next, we introduce Newtonian Perturbation Theory, the halo model, and counts-in-cells statistics which all capture different (non-linear) aspects of the LSS.
- In chapter 3, we extend the counts-in-cells matter predictions to neutral hydrogen (HI) by means of a bias relation. We compare two ideas for bias functions and propose the use of a second order approximation to a non-parametric bias function based on the cumulative distribution functions. The bias relation enables prediction of the HI probability density function as well as the so called sphere bias. We validate our predictions against the IllustrisTNG simulation. Without further modelling, we are able to show that the bias relation can capture the effect of redshift-space distortions (RSDs) too. This chapter is based on [1]

- In chapter 4, we discuss the efficient implementation of two projection integrals that occur when projecting bispectra and theoretical uncertainties for the projected power spectrum. Using the theoretical uncertainties we quantify the (Fisher) information loss when comparing a 2D power spectrum analysis with a 3D power spectrum analysis as a function of the number of tomographic bins used. The code for the projected bispectrum is flexible and can be used to answer a range of scientific questions. This chapter is based on [2]
- In chapter 5, we investigate the relation between biases in the estimated cosmological parameters and inaccurate modelling of non-linear RSDs for the power spectrum and bispectrum of projected galaxy density fields and lensing convergence for a moderate-sized galaxy survey like CMASS. The numerical work was performed with the code introduced in chapter 4. This chapter is based on [2]
- In chapter 6, we conclude the thesis, reflect on what we have learned and comment on open research questions.

Chapter 2

Background

2.1 The homogeneous Universe

It is widely believed that General Relativity (GR) describes the evolution of the Universe on its largest scales. We thus briefly review GR and its cosmological solutions. Next, we discuss the composition and dynamics of the energy content of the Universe.

2.1.1 General Relativity and cosmological solutions

GR is based on Einstein's Equivalence Principle [39]:

- The trajectory of a freely falling test body depends only on its initial position and velocity, and is independent of its composition.
- In a local inertial frame, the results of all non-gravitational experiments will be indistinguishable from the results of the same experiments performed in an inertial frame in Minkowski spacetime.

The first part is also known as the Weak Equivalence Principle. Einstein's Equivalence Principle can be satisfied by describing gravity as the curvature of a four-dimensional Lorentzian manifold where the dynamic field is the metric tensor g_{ab} . Then the notion of local inertial frames corresponds to normal coordinates. In order to have consistent equations across different coordinate choices, the objects in those equations have to transform like tensors. In particular, this requires partial derivatives to be replaced by covariant derivatives, also called connections. The Levi-Civita connection is used in GR since it is the unique torsion-free, metric-compatible connection on a Lorentzian manifold. Its connection components or

Christoffel symbols are given in abstract tensor notation by

$$\Gamma_{bc}^a = \frac{1}{2}g^{ad}(\partial_b g_{cd} + \partial_c g_{db} - \partial_d g_{bc}). \quad (2.1)$$

The Riemann curvature tensor

$$R_{bcd}^a = \partial_c \Gamma_{db}^a - \partial_d \Gamma_{cb}^a + \Gamma_{ce}^a \Gamma_{db}^e - \Gamma_{de}^a \Gamma_{cb}^e, \quad (2.2)$$

expresses the curvature in terms of the second derivatives of the metric. Its contractions, the Ricci tensor and Ricci scalar,

$$R_{ab} = R^c{}_{acb}, \quad R = R_{ab}g^{ab}, \quad (2.3)$$

are the building blocks for the Einstein Tensor

$$G_{ab} = R_{ab} - \frac{1}{2}g_{ab}R. \quad (2.4)$$

The Einstein Tensor forms the left hand side of the field equations. Einstein related it to the matter content of the Universe via the Einstein equations [40, 41]

$$G_{ab} = 8\pi G T_{ab} + \Lambda g_{ab} \quad (2.5)$$

with G being Newton's gravitational constant and Λ the cosmological constant and we used the convention $c = 1$, i.e. time and space have the same dimension.

Cosmological solutions to the field equations need to satisfy the cosmological principle which says that the spatial distribution of the Universe is homogeneous and isotropic on sufficiently large scales. While one can test isotropy very well, in particular with the cosmic microwave background (CMB), similar tests for homogeneity are much harder. Using the quasar distribution one can infer that the homogeneity scale is of order $100 h^{-1} \text{Mpc}$ [42]. Imposing spatial homogeneity and isotropy on the metric tensor reduces it to the Friedmann-Lemaître-Robertson-Walker (FLRW) family of metrics [9, 43–47],

$$ds^2 = g_{ab}dx^a dx^b = -dt^2 + a(t)^2 \left(\frac{dr^2}{1-kr^2} + r^2(d\theta^2 + d\phi^2 \sin^2 \theta) \right). \quad (2.6)$$

Hence, a cosmological (background) solution is fully specified by one function, the scale factor a , and a number, k , specifying the spatial geometry of the Universe that can be open ($k < 0$), flat ($k = 0$) or closed ($k > 0$).

Imposing the symmetries on the energy-momentum tensor, T_{ab} , forces it to take the form of a perfect fluid

$$T_{ab} = (\rho + P)U_a U_b + P g_{ab} \quad (2.7)$$

with pressure P , energy density ρ and four-velocity U .

2.1.2 Cosmic ingredients and timeline

Having reduced the number of independent functions in the metric tensor from ten to one, one can derive the resulting evolution equations for a homogeneous and isotropic Universe: The Friedmann equation,

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2} + \Lambda, \quad (2.8)$$

and the acceleration equation,

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3P) + \frac{\Lambda}{3}. \quad (2.9)$$

Since $\rho + 3P \geq 0$ for ordinary matter, the acceleration equation shows that \ddot{a} would always be negative without the cosmological constant. This is why Einstein introduced Λ as a way to obtain a static, $\dot{a} = 0$, cosmological solution. He later called it his ‘biggest blunder’ because this static solution is unstable.

Typically, one rewrites the equations above in terms of Hubble function

$$H = \frac{d \ln a}{dt} = \frac{\dot{a}}{a} \quad (2.10)$$

and the critical density

$$\rho_{\text{crit}} = \frac{3H^2}{8\pi G}. \quad (2.11)$$

The critical density derives its name from the fact that the Universe is flat if the density matches the critical density, open if it is smaller, and closed if it is larger.¹ This can be seen by rewriting the Friedmann equation

$$\frac{\rho}{\rho_{\text{crit}}} - 1 = \frac{k}{(aH)^2}. \quad (2.12)$$

¹In order for this statement to hold in the presence of a cosmological constant, one needs to include the cosmological constant to the energy density ρ in (2.8).

The Universe can be filled with different perfect fluids and assuming each component of the stress-energy tensor is separately conserved, i.e. $\nabla_a T_0^a = 0^2$, one finds an evolution equation for each component

$$\dot{\rho}_i = -3\frac{\dot{a}}{a}(\rho_i + P_i). \quad (2.13)$$

For fluids with constant equation of state, $w_i = P_i/\rho_i$, we can integrate this equation and obtain

$$\rho_i = \rho_{i,0} \left(\frac{a}{a_0}\right)^{-3(1+w)}. \quad (2.14)$$

Here, we chose the integration constants to match the observed values today, hence the subscript 0. Rewriting the density in terms of the critical density yields the so called density parameters

$$\Omega_i = \frac{\rho_i}{\rho_{\text{crit}}}. \quad (2.15)$$

The standard model of cosmology assumes the following three components: dust ($w = 0$), radiation ($w = 1/3$) and cosmological constant ($w = -1$). Their energy densities scale as follows

$$\rho \sim \begin{cases} a^{-4} & \text{radiation } (w=1/3) \\ a^{-3} & \text{dust } (w=0) \\ \text{const.} & \text{cosmological constant } (w=-1). \end{cases} \quad (2.16)$$

Due to the different scalings, the early Universe was dominated by radiation. Shortly before the CMB was formed, matter-radiation equality happened and the Universe became matter dominated. During this period, the rich large-scale structure of the Universe (LSS) grew. Fairly recently, the energy density of matter sunk to the level of the cosmological constant and the Universe has entered a period of cosmological constant domination.

The radiation component of the Universe consists of all relativistic particles (including massless photons). The dust, or matter, component is (approximately) made up of all non-relativistic particles, since they have negligible pressure (proportional to the kinetic energy) compared to their energy densities (proportional to the rest mass). Only $\sim 20\%$ of the matter content is made up by known, baryonic matter, and the other matter is considered to be cold dark matter. The microscopic origin of the cosmological constant is unknown. A flat, dust-only Universe is commonly referred to as the Einstein-deSitter (EdS) Universe and is a useful toy model to study structure formation given the LSS formed during matter domination and the cosmological constant became only recently relevant.

²The stress-energy tensor of the entire matter content must be conserved due to the second Bianchi identity.

2.1.3 Redshift and distances

Objects in the sky are detected via their emitted radiation. This is typically electromagnetic radiation but gravitational waves have, for instance, been used to observe some of the most violent events in the Universe such as black hole mergers [48, 49]. The following discussion is provided for photons but also equally applies to gravitons. Photons travel along null geodesics,

$$\frac{d^2 x^a}{d^2 \lambda} + \Gamma_{bc}^a \frac{dx^b}{d\lambda} \frac{dx^c}{d\lambda} = 0, \quad (2.17)$$

where λ is an affine parameter and $\frac{dx^a}{d\lambda}$ is the photon's 4-momentum. Using the Christoffel symbols for the FLRW metric yields for the evolution of the restframe energy, E ,

$$\frac{\dot{E}}{E} = -\frac{\dot{a}}{a} \implies E \sim 1/a. \quad (2.18)$$

The fractional energy loss of photons emitted with wavelength λ_{emit} when the scale factor was a_{emit} and observed with wavelength λ_{obs} is called redshift, z , and written in terms of wavelength one finds

$$z = \frac{\lambda_{\text{obs}} - \lambda_{\text{emit}}}{\lambda_{\text{emit}}} = \frac{a_{\text{obs}}}{a_{\text{emit}}} - 1. \quad (2.19)$$

Normalizing the scale factor today, $a_0 = 1$, then yields the well known result

$$1 + z = \frac{1}{a_{\text{emit}}}. \quad (2.20)$$

Assuming an expansion history, $H(z)$, we can relate redshift to radial co-moving distances of the object via

$$\chi(z) = \frac{1}{H_0} \int_0^z \frac{dz'}{H(z')}. \quad (2.21)$$

For objects of known intensity (Supernovae of Type Ia, Cepheids) or size (Baryonic Acoustic Oscillations), one can infer their distances directly by the size on which they appear on the sky or the flux one measures. This, together with their redshift, allows us to measure the expansion history of the Universe directly. In Fig. 2.1 we show an application of this idea where the expansion history of the late time Universe was used to infer cosmological parameters [3].

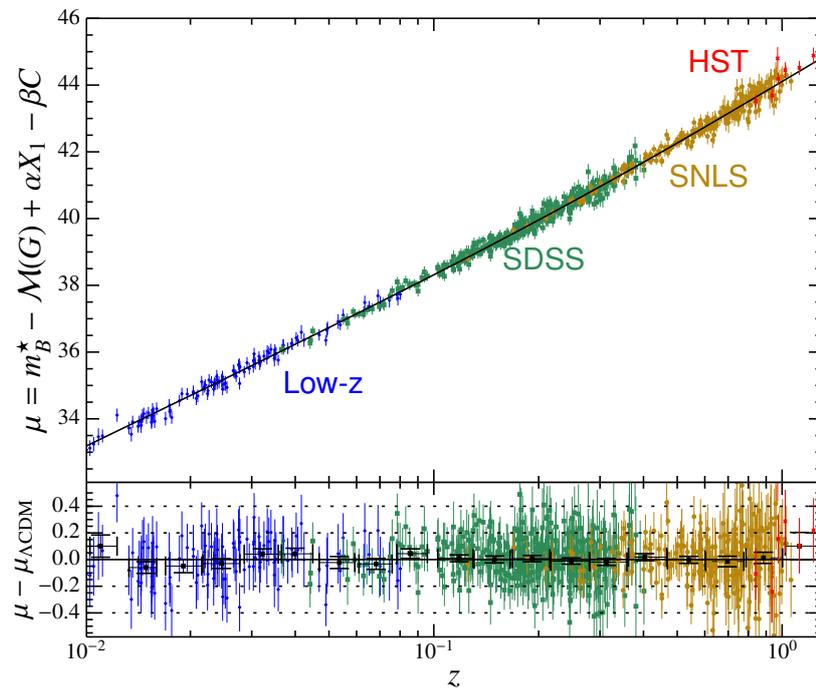


Fig. 2.1 Hubble diagram from [3]. The logarithm of the (luminosity) distance is shown as a function of the redshift. The black line represents the best-fit Λ CDM cosmology with fixed Hubble parameter today, $H(z=0) = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$. The residuals in the bottom panel indicate that the fit, and thus the theory, is fully consistent with the observations.

2.2 Initial conditions from Inflation

Inflation is a postulated phase of accelerated expansion in the very early Universe that would solve some shortcomings of the standard Big Bang picture. Due to the exponential expansion during Inflation, quantum fluctuations in the inflaton field are stretched to cosmological scales where they become classical once they reach the size of the horizon. We will review the calculations that show how the simplest model of Inflation produces Gaussian initial conditions. Next, we introduce phenomenological models for primordial non-Gaussianities. This section was heavily influenced by [50] and we refer the reader to this reference for more details.

2.2.1 Inflation

Despite being very successful, the Big Bang paradigm has at least two shortcomings that require special assumptions about the initial conditions: The CMB consists of approximately 30,000 regions that have not been in causal contact with each other until today, but have all the same mean temperature, and the statistical fluctuations around this mean seem to come from the same distribution too. This is known as the horizon problem. Moreover, the spatial curvature increases during periods of the Universe where $w > -1/3$, i.e. during radiation and matter domination. Given the Universe today is indistinguishable from a flat Universe, it must have been exceptionally flat initially. This is known as the flatness problem. The two problems can be solved by a phase of accelerated expansion during the very early universe (first $\sim 10^{-43}s$), called Inflation [13, 51, 52].

The idea is that the Universe was expanding so rapidly that the co-moving Hubble horizon, $\mathcal{H} = aH$, shrunk

$$\frac{d}{d\tau}\mathcal{H} < 0. \quad (2.22)$$

Here, we take the derivative with respect to conformal time $\tau = \int_0^t dt'/a(t')$. Provided enough expansion took place, the scales we observe today were in causal contact before the onset of the standard evolution, hence solving the horizon problem. Similarly, the rapid expansion would dilute all initial curvature, such that it would be unobservable today for a large range of initial curvatures. The amount of expansion that happened during Inflation is commonly measured in e-folds, N ,

$$N = \int dN = \int d \ln a = \ln \frac{a_{\text{start}}}{a_{\text{end}}}. \quad (2.23)$$

The horizon and flatness problem are both solved if there were at least 60 e-folds of Inflation.

From the continuity equation (2.13), one sees that the strong energy condition ($w > -1/3$) needs to be violated in order to achieve a shrinking co-moving Hubble horizon. For

the remainder of this subsection, we study a toy model which allows us to gain a better understanding of how to achieve this scenario by means of a simple scalar field ϕ described by the action

$$S = \int d^4x \sqrt{-g} (X - V(\phi)) \quad (2.24)$$

where $X = -g^{ab} \partial_a \phi \partial_b \phi$ is the kinetic term and g the determinant of the metric. The expansion during Inflation can then be studied in terms of the potential V . In order to have a shrinking horizon and a sufficient number of e-folds, one commonly requires that the two slow-roll parameters

$$\varepsilon = -\frac{d \ln H}{d \ln a} = -\frac{\dot{H}}{H^2} < 1, \quad \eta = \frac{d \ln(\varepsilon)}{d \ln a} = -\frac{\dot{\varepsilon}}{H \varepsilon} < 1 \quad (2.25)$$

are small. In particular, $\varepsilon < 1$ ensures that the co-moving Hubble is shrinking and $\eta < 1$ ensures that Inflation lasts sufficiently long to solve the mentioned shortcomings. Varying the action (2.24) with respect to the metric tensor yields the stress energy tensor

$$T^{ab} = \frac{2}{\sqrt{-g}} \frac{\delta L}{\delta g_{ab}} \quad (2.26)$$

from which we obtain the energy density

$$\rho_\phi = \frac{1}{4} \left[\sqrt{T_{ab} T^{ab}} - 3T^2 - T \right] = \frac{1}{2} \dot{\phi}^2 + V \quad (2.27)$$

and pressure

$$P_\phi = \frac{1}{12} \left[\sqrt{T_{ab} T^{ab}} - 3T^2 + 3T \right] = \frac{1}{2} \dot{\phi}^2 - V \quad (2.28)$$

of the scalar field. Inserting those into the Friedmann equation (2.8) yields

$$H^2 = \frac{8\pi G}{3} \left(\frac{1}{2} \dot{\phi}^2 + V \right). \quad (2.29)$$

This allows us to write the slow-roll parameters as

$$\varepsilon = 4\pi G \left(\frac{\dot{\phi}}{H} \right)^2, \quad \eta = 2 \left(\varepsilon + \frac{\ddot{\phi}}{H \dot{\phi}} \right). \quad (2.30)$$

In the slow approximation ($\varepsilon \ll 1$), the Friedmann equation simplifies to

$$H^2 \simeq \frac{8\pi G}{3} V. \quad (2.31)$$

The scalar field's evolution is governed by the Klein-Gordon equation that can be simplified in the slow-roll approximation ($\eta \ll 1$) to

$$\ddot{\phi} + 3H\dot{\phi} + V_{,\phi} = 0 \implies 3H\dot{\phi} = -V_{,\phi}. \quad (2.32)$$

The Friedmann equation and Klein-Gordon equation in the slow-roll limit then allow us to rewrite the slow-roll conditions in terms of the inflaton's potential V

$$\epsilon_V = \frac{8\pi G}{2} \left(\frac{V_{,\phi}}{V} \right)^2, \quad \eta_V = 4\epsilon - \eta = 28\phi G \frac{V_{,\phi\phi}}{V}. \quad (2.33)$$

Those equations allow us to quickly check whether slow-roll Inflation takes place given a specific potential and the initial conditions of the scalar field.

2.2.2 Gauge problem in GR

In order to study fluctuations around this spatially homogeneous and isotropic solution, we first need to understand the Gauge problem in GR. A change of coordinates moves some perturbations across different metric components and between the metric and stress-energy tensor as we will see below. Let us derive the transformations of the perturbations in the Einstein equations under a change of coordinates to understand the gauge degrees of freedom. We start by writing the perturbed line element (around a flat cosmological solution) as

$$ds^2 = a(\tau)^2 \left(-(1 + 2A)d\tau^2 + 2B_i dx^i d\tau + (\delta_{ij}^K + h_{ij}) dx^i dx^j \right). \quad (2.34)$$

We use the convention that indices from the beginning of the alphabet such as a, b , etc. are running from 0 to 3 whereas i, j, k , etc. are spatial indices and range only from 1 to 3. Using the Hemholtz theorem, one can decompose the perturbations into scalar, vector and tensor components (under spatial rotations)

$$B_i = \partial_i B + B_i^V, \quad (2.35)$$

$$h_{ij} = 2C\delta_{ij}^K + 2 \left(\partial_{ij} - \frac{1}{3}\nabla^2 \right) E + \partial_i E_j^V + 2E_{ij}^T, \quad (2.36)$$

where the vector and tensor components are divergence-free and the tensor components are symmetric and trace-free:

$$\partial^i B_i^V = \partial^i E_i^V = 0 = \partial^i E_{ij}^T, \quad \delta^{K,ij} E_{ij}^T = 0, \quad E_{ij}^T = E_{ji}^T. \quad (2.37)$$

The decomposition greatly simplifies the analysis, since at linear order those components do not mix on an isotropic background. From now on, we restrict ourselves to scalar perturbations. Let us consider the (scalar) change of coordinates

$$x \rightarrow \tilde{x} = x + \xi \quad (2.38)$$

where

$$\xi = (T, \partial_i L). \quad (2.39)$$

The co- and contravariant components of a tensor, S , in the two different coordinate system transform like

$$\tilde{S}_\mu^\nu(\tilde{x}) = \frac{\partial x^\lambda}{\partial \tilde{x}^\mu} \frac{\partial \tilde{x}^\nu}{\partial x^\rho} S_\lambda^\rho(\tilde{x} - \xi). \quad (2.40)$$

Expanding the right-hand side to linear order in ξ yields the transformation of the scalar perturbations in the metric:

$$\tilde{A} = A - T' - \mathcal{H}T, \quad (2.41a)$$

$$\tilde{B} = B + T - L', \quad (2.41b)$$

$$\tilde{C} = C - \mathcal{H}T - \frac{1}{3}\nabla L, \quad (2.41c)$$

$$\tilde{E} = E - L. \quad (2.41d)$$

Assuming the perturbed energy-momentum tensor is a perfect fluid, we find at linear order

$$\delta T_{ab} = (\delta\rho + \delta P)U_a U_b + \delta P g_{ab} + (\rho + P)(U_a \delta U_b + U_b \delta U_a) + P \delta g_{ab}. \quad (2.42)$$

After ensuring the four-velocity is normalised, we can parametrise the remaining scalar degree of freedom of the four-velocity by the velocity potential v

$$\delta U^a = \frac{1}{a}(1 - A, \partial^i v). \quad (2.43)$$

The scalar perturbations in the energy momentum tensor transform like

$$\delta \tilde{\rho} = \delta\rho - \rho' T, \quad (2.44a)$$

$$\delta \tilde{P} = \delta P - P' T, \quad (2.44b)$$

$$v = v + L'. \quad (2.44c)$$

While the Gauge freedom can be used to simplify some calculations considerably, it is useful to predict gauge-invariant variables. There are two gauge-invariant variables that are particularly relevant. Firstly, the co-moving curvature perturbation

$$\zeta = C - \frac{1}{3}\nabla^2 E + \mathcal{H} \frac{\delta\rho}{\rho'}. \quad (2.45)$$

Using the Einstein equations, one can show that ζ is constant on super Horizon scales. This allows us to map perturbations from the time when they leave the Horizon (during Inflation) to the time when they reenter the Horizon (during radiation or matter domination) despite the fact that some of the physics in between, in particular reheating, are not well known. Secondly, the Bardeen Potential

$$\Phi = -C + \mathcal{H}(B - E') + \frac{1}{3}\nabla^2 E \quad (2.46)$$

is a popular gauge-invariant variable since it corresponds to the Newtonian gravitational potential on sub-horizon scales.

2.2.3 Initial conditions from Inflation

We will now study quantum fluctuations of the scalar field toy model and aim to express those in terms of the co-moving curvature perturbations. Throughout this section, we will ignore background fluctuations and work in deSitter space.³ This is for the sake of clarity and one can show that the results extend to the slow-roll case. Moreover, we work in spatially flat gauge ($C = E = 0$). In deSitter space we have

$$H = \text{const.}, \quad a \sim -1/\tau. \quad (2.47)$$

Decomposing the field of the toy model (2.24) via

$$\phi(\mathbf{x}, \tau) = \bar{\phi}(\tau) + \delta\phi(\mathbf{x}, \tau) = \bar{\phi}(\tau) + f(\mathbf{x}, \tau)/a(\tau) \quad (2.48)$$

yields the Mukhanov-Sasaki as equation of motion for the fluctuations

$$f''(\mathbf{k}, \tau) + f(\mathbf{k}, \tau) \left(k^2 - \frac{2}{\tau^2} \right) = 0. \quad (2.49)$$

³This is the case with $\eta = 0$, where Inflation would never end.

The classical solutions to this equation are given by

$$f(\mathbf{k}, \tau) = c_1 \frac{e^{-ik\tau}}{\sqrt{2k}} \left(1 - \frac{i}{k\tau}\right) + c_2 \frac{e^{ik\tau}}{\sqrt{2k}} \left(1 + \frac{i}{k\tau}\right). \quad (2.50)$$

To quantize the field, we observe that deep inside the horizon, $k \gg aH$, the equation of motion (2.49) reduces to the harmonic oscillator, which we know how to (canonically) quantize by introducing the following commutation relation for the raising- and lowering-operators a and a^\dagger

$$[\hat{a}(\mathbf{k}), \hat{a}^\dagger(\mathbf{k}')] = (2\pi)^3 \delta^D(\mathbf{k} + \mathbf{k}'). \quad (2.51)$$

The general solution of the quantized Mukhanov-Sasaki equation is a linear combination of those raising- and lowering-operators times a solution of the classical equation of motion

$$\hat{f}(\mathbf{x}, \tau) = \int \frac{d^3k}{(2\pi)^3} \left[f(\mathbf{k}, \tau) \hat{a}(\mathbf{k}) e^{-i\mathbf{k}\mathbf{x}} + f^*(\mathbf{k}, \tau) \hat{a}^\dagger(\mathbf{k}) e^{i\mathbf{k}\mathbf{x}} \right]. \quad (2.52)$$

In the initial conditions, $\tau \rightarrow -\infty$, we are only interested in the modes deep inside the horizon where the equation of motions approximates the harmonic oscillator. Hence we require the solution to converge to the positive energy solution, the so-called Bunch-Davies vacuum, i.e. $c_1=1$ and $c_2=0$. From this we can compute the power spectrum:

$$\langle \phi(\mathbf{k}, \tau) \phi(\mathbf{k}', \tau) \rangle = (2\pi)^3 \delta^D(\mathbf{k} + \mathbf{k}') P_\phi(k, \tau), \quad (2.53)$$

with $P_\phi(k, \tau) = \frac{1}{2ka^2} \left(1 + \frac{1}{(k\tau)^2}\right)$. The dimensionless power spectrum of the scalar field, Δ_ϕ , is defined as

$$\Delta_\phi^2 \equiv \frac{k^3}{2\pi^2} P_\phi(k) = \frac{k^3}{2\pi^2} \frac{1}{2ka^2} \left(1 + \frac{1}{(k\tau)^2}\right) = \left(\frac{H}{2\pi}\right)^2 \left(1 + \frac{k^2}{a^2 H^2}\right). \quad (2.54)$$

On superhorizon scales ($k \ll aH$), the second term in the brackets vanishes and we can translate the perturbation into a co-moving curvature perturbation

$$\Delta_\zeta^2(k) = \left(\frac{aH}{\dot{\phi}}\right)^2 \Delta_\phi^2(k) \Big|_{k \leq aH} = \frac{8\pi G}{\epsilon} \Delta_\phi^2(k) \Big|_{k \leq aH}. \quad (2.55)$$

In slow-roll Inflation, we expect a roughly scale-invariant power spectrum since both the Hubble and epsilon are varying slowly. One typically parametrises the dimensionless power

spectrum in terms of an amplitude, A , and spectral index, n_s , as

$$\Delta_\zeta^2(k) = A \left(\frac{k}{k_*} \right)^{n_s-1}. \quad (2.56)$$

The deviation from scale invariance is given by

$$n_s - 1 = \frac{d \ln \Delta_\zeta^2}{d \ln k} = -2\varepsilon - \eta < 0. \quad (2.57)$$

The power spectrum (2.53) is one of the most important statistics in cosmology. Since we work with fluctuations around the mean, it is the first non-vanishing moment of the distribution. Moreover, for mean-zero Gaussian random fields all the information is contained in the power spectrum.

2.2.4 Primordial non-Gaussianities

Intuitively, fields that are described by a harmonic oscillator, give rise to Gaussian fluctuations. Thus, to obtain significant deviations from Gaussianity in the initial conditions, one needs to break some of the assumptions made in the previous calculations. In particular, one could introduce more dynamical fields, couple the kinetic term non-minimally to gravity, violate the slow-roll conditions or start from a non-Bunch-Davies vacuum [53]. In any case, we know that deviations from Gaussianity must be small [54].

In case those deviations still lead to Inflation, one can test their impact on the shape of the power spectrum or by measuring higher order moments. The bispectrum

$$\langle \Phi(\mathbf{k}_1)\Phi(\mathbf{k}_2)\Phi(\mathbf{k}_3) \rangle = (2\pi)^3 \delta^D(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) B_\Phi(k_1, k_2, k_3) \quad (2.58)$$

is a particularly popular tool to test primordial non-Gaussianities because it is the leading non-Gaussian term in a so-called Edgeworth expansion. Spatial homogeneity makes the bispectrum only non-zero when the three wave vectors form a valid triangle. Note that Φ is referring to the Bardeen potential and not the scalar field ϕ . One typically constrains three general shapes that are meant to test broad classes of models. Upon discovery of a shape, one could then refine the search. Another approach is to use a basis expansion in the space of possible bispectrum shapes and constrain those templates [55]. We will follow the former approach and use the following three templates:

Local shape

The simplest model for a non-Gaussian potential, Φ_{NG} , is built from a local interaction of an underlying Gaussian field Φ_{G} [56–58]

$$\Phi_{\text{NG}}(\mathbf{x}) = \Phi_{\text{G}}(\mathbf{x}) + f_{\text{NL}}(\Phi_{\text{G}}^2(\mathbf{x}) - \langle \Phi_{\text{G}}^2 \rangle) \quad (2.59)$$

where f_{NL} determines the strength of the non-Gaussianities. This model yields a non-zero bispectrum

$$B_{\Phi}^{\text{local}}(k_1, k_2, k_3) = 2A^2 f_{\text{NL}}^{\text{local}} \left[\frac{1}{(k_1 k_2)^{4-n_s}} + 2 \text{ perm.} \right] \quad (2.60)$$

where we used $P_{\Phi} = A k^{4-n_s}$. The shape has a characteristic peak in squeezed (also called local) configurations where one side of the triangle is much shorter than the other two. Maldacena showed that for slow-roll Inflation, this signal is order ϵ , thus unobservable [59]. In general, the squeezed limit of the bispectrum for single field models is order $f_{\text{NL}}(n_s - 1) \sim O(\epsilon)$ making it unobservable too [60]. This means, a non-zero local type non-Gaussianity would be a strong indicator for multifield Inflation - the opposite is not necessarily true. The best 1σ constraints are currently [54]

$$f_{\text{NL}}^{\text{local}} = -0.9 \pm 5.1. \quad (2.61)$$

Equilateral shape

The equilateral shape arises in a wide range of non-vanilla inflationary dynamics, for instance with non-standard kinetic terms, (see [61] and references therein). A separable approximation to the equilateral shape was introduced in [62] and is given by

$$B_{\Phi}^{\text{equil}}(k_1, k_2, k_3) = 6A^2 f_{\text{NL}}^{\text{equil}} \left\{ \left[-\frac{1}{(k_1 k_2)^{4-n_s}} + 2 \text{ perm.} \right] + \frac{1}{(k_1 k_2 k_3)^{2(4-n_s)/3}} \right. \\ \left. + \left[\frac{1}{k_1^{(4-n_s)/3} k_2^{2(4-n_s)/3} k_3^{4-n_s}} + 5 \text{ perm} \right] \right\}. \quad (2.62)$$

The best 1σ constraints are currently [54]

$$f_{\text{NL}}^{\text{equil}} = -26 \pm 47. \quad (2.63)$$

Orthogonal shape

The orthogonal shape,

$$B_{\Phi}^{\text{ortho}}(k_1, k_2, k_3) = 6A^2 f_{\text{NL}}^{\text{ortho}} \left\{ \left[-\frac{3}{(k_1 k_2)^{4-n_s}} + 2 \text{ perm.} \right] + \frac{8}{(k_1 k_2 k_3)^{2(4-n_s)/3}} \right. \\ \left. + \left[\frac{3}{k_1^{(4-n_s)/3} k_2^{2(4-n_s)/3} k_3^{4-n_s}} + 5 \text{ perm.} \right] \right\}, \quad (2.64)$$

was introduced as a tool to probe derivative interaction in multifield Inflation effectively [63]. The best 1σ constraints are currently [54]

$$f_{\text{NL}}^{\text{ortho}} = -38 \pm 24. \quad (2.65)$$

2.2.5 Connecting early and late times

When studying the LSS (at late times), we are interested in the matter density contrast δ . It is related to the Bardeen potential, Φ , and the co-moving curvature perturbations, ζ , via the Einstein equations. During matter and Λ domination, this relation is

$$\delta(\mathbf{k}, z) = \mathcal{M}(k, z) \Phi(\mathbf{k}, z_e) = \mathcal{M}(k, z) \frac{3 + 3w(z_e)}{5 + 3w(z_e)} \zeta(\mathbf{k}, z_e). \quad (2.66)$$

where the Poisson factor \mathcal{M} is given by

$$\mathcal{M}(k, z) = \frac{2k^2 c^2}{3\Omega_{m,0} H_0^2} T(k, z). \quad (2.67)$$

The first term in the Poisson factor stems from the Poisson equation that is valid on sub-horizon scales. In addition, the transfer function $T(k, z)$ realises the time evolution from the time of horizon crossing at z_e . The time evolution is given by the Einstein-Boltzmann system of equations for all the relevant components of the Universe. There are efficient solvers for those equations such as CAMB [64, 65] and CLASS [66] that can be used. It is worthwhile discussing the time evolution in the special cases of radiation domination and matter domination briefly. Those two regimes can be roughly separated by the wave vector k_{eq} that corresponds to the size of the modes that entered at matter-radiation equality. Smaller modes entered the horizon during radiation domination and larger modes did so during matter domination. As noted before, the co-moving curvature perturbations are constant on super-horizon scales. During radiation domination, sub-horizon matter perturbations only grow with the logarithm of the scale factor, the so-called Meszaros effect. During matter

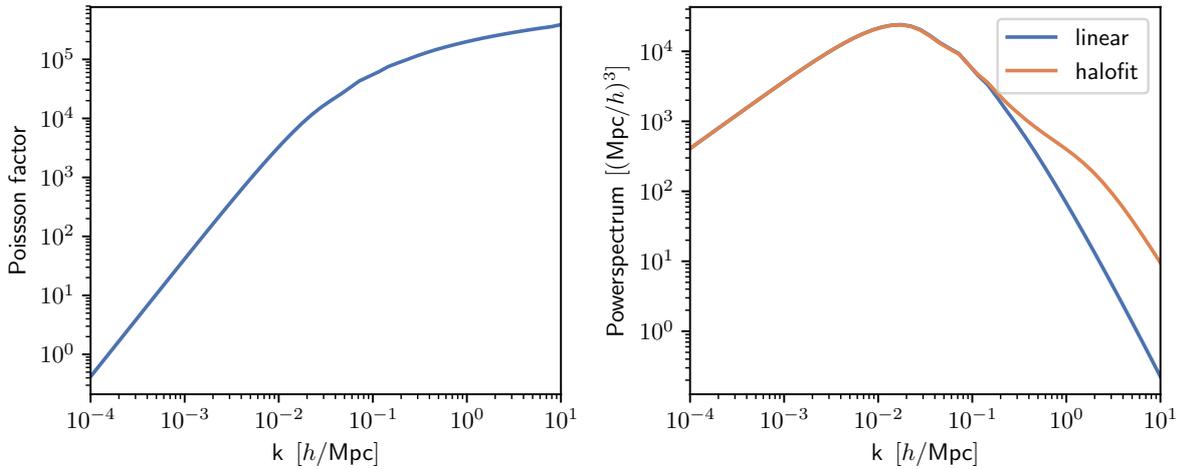


Fig. 2.2 Left: The Poisson factor at redshift zero is shown. The scale dependence is dominated by the k^2 term from the Poisson equation. In addition, the slope changes around the scale of matter-radiation equality $k_{\text{eq}} \simeq 0.015 h\text{Mpc}^{-1}$, since small scales are suppressed due to the Meszaros effect. Right: Linear power spectrum at redshift zero (blue line) compared to the HALOFIT [4, 5] prediction for the non-linear power spectrum (orange). The non-linearities are strongest on small scales and stem from an enhanced clustering compared to the linear predictions.

domination, in contrast, sub-horizon modes grow proportional to the scale factor. This means that small modes, those that entered the horizon during radiation domination, are suppressed relative to those that entered later. This is why transfer function becomes constant for $k < k_{\text{eq}}$ and is decreasing for $k > k_{\text{eq}}$. Finally, the Bardeen potential decays by 10% due to the transition from radiation to matter domination because of the changing equations of state w of the cosmic fluid. Fig. 2.2 shows the Poisson factor at redshift zero on the left side. Around the scale of radiation-matter equality, $k_{\text{eq}} \simeq 0.015 h\text{Mpc}^{-1}$, we see a change of slope that is due to the changing behavior of the transfer function $T(k, z)$. Using the Poisson factor, one can relate the (linear) matter- to the Bardeen potential power spectrum by

$$P_L(k, z) = \mathcal{M}^2(k, z) P_{\Phi}(k, z_e). \quad (2.68)$$

The primordial bispectrum (at late times) is given by

$$B_{\text{prim}}(k_1, k_2, k_3, z) = \mathcal{M}(k_1, z) \mathcal{M}(k_2, z) \mathcal{M}(k_3, z) B_{\Phi}(k_1, k_2, k_3, z_{e,1}, z_{e,2}, z_{e,3}). \quad (2.69)$$

Gravitational collapse introduces non-linearities at late times that lead to a gravitational bispectrum and modify the power spectrum. On the right side of Fig. 2.2, we compare the linear power spectrum (2.68) to a fully non-linear power spectrum model. The non-linear

evolution is generating an increased clustering on small scales which is not captured by the linear prediction.

In the next section, we are going to discuss how one can use the linear solutions to find a perturbative expansion of the full non-linear evolution during matter and Λ domination.

2.3 Newtonian Perturbation Theory

Matter density perturbations grow on sub-horizon scales due to gravitational instability. In this section, we are going to solve the evolution equations perturbatively during matter and Λ domination to capture some of the non-linear effects. On the statistical level, non-linearities change the distribution of the fluctuations by sourcing higher order correlation functions, in particular a gravitational bispectrum. There are two additional non-linear effects that have to be understood in order to accurately describe observations: redshift-space distortions (RSDs) and biasing. Lastly, we will relax the assumptions of Gaussian initial conditions and extend the predictions accordingly.

We are working in an intermediate regime where we are deep inside the horizon but still on much larger scales than the regions with strong gravitational fields or high velocities, so a Newtonian treatment is sufficient. The assumption of a perfect, pressureless fluid coupled to gravity leads us to the Euler-Poisson system that we need to solve (perturbatively). We will review the Standard (or Eulerian) and Lagrangian framework of perturbation theory. Both are reviewed in [67]. Biasing has been reviewed in [68].

2.3.1 Evolution equations in co-moving coordinates

To simplify calculations, we begin by separating the effect of the expanding background from the evolution equations by expressing equations in terms of co-moving coordinates \mathbf{x} . They are related to the physical coordinates \mathbf{r} by the scale factor a

$$\mathbf{r} = a \mathbf{x}. \quad (2.70)$$

Velocities can be split into the Hubble flow [11] and peculiar velocities, \mathbf{u} ,

$$\frac{\partial \mathbf{r}}{\partial \tau} = \mathbf{r}' = \mathbf{v}(\mathbf{x}, \tau) = \mathcal{H} \mathbf{x} + \mathbf{u}(\mathbf{x}, \tau). \quad (2.71)$$

Primes denote derivatives with respect to conformal time τ . Similarly, we are interested in deviations from the average density $\bar{\rho}$, since the average density drives the cosmic expansion

and is captured by the scale factor a . The matter density contrast, δ , is given by

$$\rho(\mathbf{x}, \tau) = \bar{\rho}(\tau) (1 + \delta(\mathbf{x}, \tau)). \quad (2.72)$$

The Poisson equation in co-moving coordinates then reads

$$\nabla_{\mathbf{x}}^2 \Phi / a^2 = 4\pi G \bar{\rho} (1 + \delta). \quad (2.73)$$

One can separate the potential Φ in a component that sources the background expansion and an auxiliary potential ϕ that sources the density contrast via

$$\Phi = \phi + x^2 2a^2 \pi G \bar{\rho}. \quad (2.74)$$

Recalling the definition of the matter density parameter Ω_m , the Poisson equation becomes

$$\nabla^2 \phi = \frac{3}{2} \Omega_m(\tau) \mathcal{H}^2 \delta(\mathbf{x}, \tau). \quad (2.75)$$

Here we assume that the cosmological constant does not cluster. In co-moving coordinates, the continuity equations is

$$\frac{\partial \delta(\mathbf{x}, \tau)}{\partial \tau} + \nabla \cdot [(1 + \delta(\mathbf{x}, \tau)) \mathbf{u}(\mathbf{x}, \tau)] = 0, \quad (2.76)$$

and the Euler equations becomes

$$\frac{\partial \mathbf{u}(\mathbf{x}, \tau)}{\partial \tau} + \mathcal{H} \mathbf{u}(\mathbf{x}, \tau) + \mathbf{u}(\mathbf{x}, \tau) \cdot \nabla \mathbf{u}(\mathbf{x}, \tau) = -\nabla \phi. \quad (2.77)$$

The non-linear term in the Euler equation can be absorbed in a total time derivative of the peculiar velocity: $d\mathbf{u}/d\tau = \mathbf{u}' + \mathbf{u}\nabla\mathbf{u}$.

2.3.2 Eulerian perturbation theory

In Eulerian (or Standard) perturbation theory (SPT), we are solving the evolution equations at late times perturbatively. To simplify the equations, we decompose the peculiar velocity into its divergence θ and curl \mathbf{w}

$$\theta(\mathbf{x}, \tau) = \nabla \cdot \mathbf{u}(\mathbf{x}, \tau), \quad (2.78)$$

$$\mathbf{w}(\mathbf{x}, \tau) = \nabla \times \mathbf{u}(\mathbf{x}, \tau). \quad (2.79)$$

The Euler equation of the curl

$$\mathbf{w}' + \mathcal{H} \mathbf{w} - \nabla \times (\mathbf{u} \times \mathbf{w}) = 0 \quad (2.80)$$

has no source term. Accordingly, if the curl was zero at some time, it would remain zero. In addition, linearizing (2.80) shows $\mathbf{w} \sim 1/a$. Thus, even if there was vorticity initially, when the linear theory was valid, it would decay. For our analysis, it is sufficient to ignore vorticity even though there is evidence that vorticity occurs at late times and small scales, indicating that the fluid is neither perfect nor pressureless. The effective field theory (EFT) of LSS takes this into account. Without curl, the velocity is fully specified by the velocity potential

$$\theta(\mathbf{x}, \tau) = \nabla \mathbf{u}(\mathbf{x}, \tau) \rightarrow \mathbf{u}(\mathbf{k}, \tau) = i \frac{\mathbf{k}}{k^2} \theta(\mathbf{k}, \tau). \quad (2.81)$$

Next, we study the time evolution of the velocity potential at linear order by taking the divergence and a time derivative of the linear Euler equation. Using the Poisson equation for the gravitational potential, we obtain

$$\theta'' + \mathcal{H} \theta' = \frac{3}{2} \Omega_m(\tau) \mathcal{H}^2(\tau) \theta. \quad (2.82)$$

One can solve this equation with a separable Ansatz, $\theta(\mathbf{x}, \tau) = D(\tau) \theta(\mathbf{x})$, and denote the growing (shrinking) solution by D_+ (D_-) respectively. Since only the growing mode is relevant for structure formation, we use $D \equiv D_+$. In an Einstein-deSitter Universe ($a = \tau^2$, $\mathcal{H} = 2/\tau$, $\Omega_m(\tau) = 1$), we find the solutions

$$D \equiv D_+ \sim \tau^2 \sim a, \quad D_- \sim \tau^{-3} \sim a^{-3/2}. \quad (2.83)$$

Lastly, the linear continuity equation relates the first order density contrast and velocity potential via

$$\theta(\mathbf{x}, \tau) = -\frac{\partial D}{\partial \tau} \delta(\mathbf{x}) = -\frac{\partial \ln D}{\partial \ln a} \frac{d \ln a}{d \ln \tau} \delta(\mathbf{x}, \tau) = -f \mathcal{H} \delta(\mathbf{x}, \tau). \quad (2.84)$$

Here f is the logarithmic growth rate, $f = d \ln D / d \ln a$. In Fig. 2.3 we show the growth factor (left side) and the logarithmic growth rate (right side) both for the EdS and Λ CDM cosmology. Up to a normalization constant, the values are identical during matter domination and only when the cosmological constant became relevant do the curves start differing.

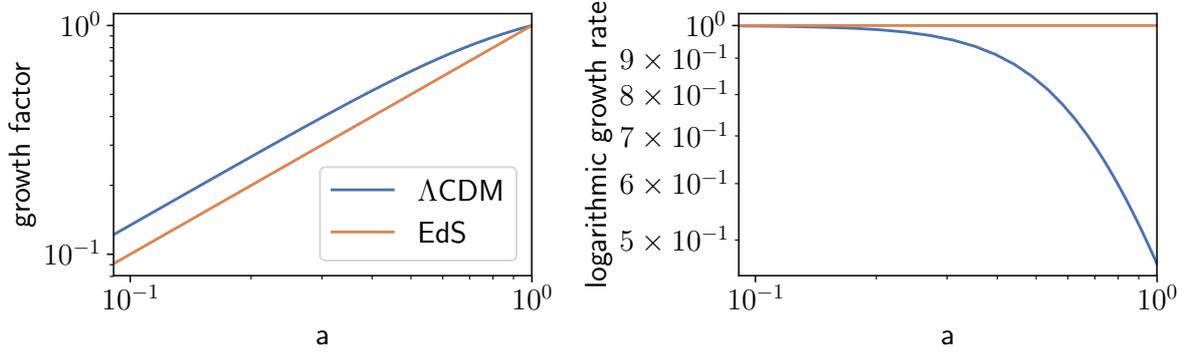


Fig. 2.3 Left: Growth factor normalised to one today, in a Λ CDM (blue) and Einstein-deSitter (2.83), (orange) Universe. During matter-domination, they are parallel and only started deviating when the cosmological constant became relevant and the growth of structure started to stall which is not accounted for in Einstein-deSitter. Right: The logarithmic growth rate in a Λ CDM cosmology (blue) is decreasing at late times, while it remains constant in an Einstein-deSitter Universe (orange).

Higher orders

The perturbative expansion is most easily done in Fourier space. There, the Euler and continuity equation are

$$\frac{\partial \delta(\mathbf{k}, \tau)}{\partial \tau} + \theta(\mathbf{k}, \tau) = - \int \frac{d^3 \mathbf{p}_1}{(2\pi)^3} \frac{d^3 \mathbf{p}_2}{(2\pi)^3} (2\pi)^3 \delta^D(\mathbf{k} - \mathbf{p}_{1:2}) \alpha(\mathbf{p}_1, \mathbf{p}_2) \theta(\mathbf{p}_1, \tau) \delta(\mathbf{p}_2, \tau), \quad (2.85)$$

and

$$\begin{aligned} \frac{\partial \theta(\mathbf{k}, \tau)}{\partial \tau} + \mathcal{H} \theta(\mathbf{k}, \tau) + \frac{3}{2} \Omega_m \mathcal{H}^2 \delta(\mathbf{k}, \tau) = \\ - \int \frac{d^3 \mathbf{p}_1}{(2\pi)^3} \frac{d^3 \mathbf{p}_2}{(2\pi)^3} (2\pi)^3 \delta^D(\mathbf{k} - \mathbf{p}_{1:2}) \beta(\mathbf{p}_1, \mathbf{p}_2) \theta(\mathbf{p}_1, \tau) \theta(\mathbf{p}_2, \tau) \end{aligned} \quad (2.86)$$

where $\mathbf{p}_{1:n} = \mathbf{p}_1 + \dots + \mathbf{p}_n$ and we defined the kernels

$$\alpha(\mathbf{k}_1, \mathbf{k}_2) = \frac{\mathbf{k}_{1:2} \cdot \mathbf{k}_1}{k_1^2} \quad \text{and} \quad \beta(\mathbf{k}_1, \mathbf{k}_2) = \frac{1}{2} \mathbf{k}_{1:2}^2 \frac{\mathbf{k}_1 \cdot \mathbf{k}_2}{k_1^2 k_2^2}. \quad (2.87)$$

To simplify calculations, we assume an EdS Universe where $D^{(n)} = (D^{(1)})^n$ [67]. Then the perturbative expansions of the density potential and velocity divergence are

$$\delta(\mathbf{k}, \tau) = \sum_{n=1}^{\infty} D^{(n)}(\tau) \delta^{(n)}(\mathbf{k}) = \sum_{n=1}^{\infty} a^n(\tau) \delta^{(n)}(\mathbf{k}), \quad (2.88)$$

$$\theta(\mathbf{k}, \tau) = -\mathcal{H} \sum_{n=1}^{\infty} D^{(n)}(\tau) \theta^{(n)}(\mathbf{k}) = -\mathcal{H} \sum_{n=1}^{\infty} a^n(\tau) \theta^{(n)}(\mathbf{k}). \quad (2.89)$$

Inserting those expansions (2.88) and (2.89) into the Euler- and continuity equation allows us to express the n -th order perturbations in terms of lower orders via

$$A_n(\mathbf{k}) = n\delta^{(n)}(\mathbf{k}) - \theta^{(n)}(\mathbf{k}) \quad (2.90)$$

and

$$B_n(\mathbf{k}) = 3\delta^{(n)}(\mathbf{k}) - (2n+1)\theta^{(n)}(\mathbf{k}) \quad (2.91)$$

where

$$A_n(\mathbf{k}) = \int \frac{d^3\mathbf{p}_1}{(2\pi)^3} \frac{d^3\mathbf{p}_2}{(2\pi)^3} (2\pi)^3 \delta^D(\mathbf{k} - \mathbf{p}_{1:2}) \alpha(\mathbf{k}, \mathbf{p}_1) \sum_{m=1}^{n-1} \theta^{(m)}(\mathbf{p}_1) \delta^{(n-m)}(\mathbf{p}_2), \quad (2.92)$$

$$B_n(\mathbf{k}) = -2 \int \frac{d^3\mathbf{p}_1}{(2\pi)^3} \frac{d^3\mathbf{p}_2}{(2\pi)^3} (2\pi)^3 \delta^D(\mathbf{k} - \mathbf{p}_{1:2}) \beta(\mathbf{k}, \mathbf{p}_1) \sum_{m=1}^{n-1} \theta^{(m)}(\mathbf{p}_1) \theta^{(n-m)}(\mathbf{p}_2). \quad (2.93)$$

Solving for the n -th order density contrast and velocity potential, we find

$$\delta^{(n)}(\mathbf{k}) = \frac{1}{(2n+3)(n-1)} [(2n+1)A_n(\mathbf{k}) - B_n(\mathbf{k})], \quad (2.94)$$

$$\theta^{(n)}(\mathbf{k}) = \frac{1}{(2n+3)(n-1)} [3A_n(\mathbf{k}) - nB_n(\mathbf{k})]. \quad (2.95)$$

Applying the identities iteratively, allows us to express each term in the expansions in terms of the linear perturbation $\delta^{(1)}$ as

$$\delta^{(n)}(\mathbf{k}) = \prod_{i=1}^n \left(\int \frac{d^3\mathbf{p}_i}{(2\pi)^3} \delta^{(1)}(\mathbf{p}_i) \right) (2\pi)^3 \delta^D(\mathbf{k} - \mathbf{p}_{1:n}) F_n(\mathbf{p}_1, \dots, \mathbf{p}_n) \quad (2.96)$$

$$\theta^{(n)}(\mathbf{k}) = \prod_{i=1}^n \left(\int \frac{d^3\mathbf{p}_i}{(2\pi)^3} \delta^{(1)}(\mathbf{p}_i) \right) (2\pi)^3 \delta^D(\mathbf{k} - \mathbf{p}_{1:n}) G_n(\mathbf{p}_1, \dots, \mathbf{p}_n). \quad (2.97)$$

The gravitational kernels, F_n , are given by

$$F_n(\mathbf{p}_1, \dots, \mathbf{p}_n) = \sum_{m=1}^{n-1} \frac{G_m(\mathbf{p}_1, \dots, \mathbf{p}_m)}{(2n+3)(n-1)} [(2n+1)\alpha(\mathbf{p}_{1:m}, \mathbf{p}_{m+1:n}) F_{n-m}(\mathbf{p}_{m+1}, \dots, \mathbf{p}_n) \\ + 2\beta(\mathbf{p}_{1:m}, \mathbf{p}_{m+1:n}) G_{n-m}(\mathbf{p}_{m+1}, \dots, \mathbf{p}_n)] \quad (2.98)$$

while the kernels for the velocity potential, G_n , are

$$G_n(\mathbf{p}_1, \dots, \mathbf{p}_n) = \sum_{m=1}^{n-1} \frac{G_m(\mathbf{p}_1, \dots, \mathbf{p}_m)}{(2n+3)(n-1)} [3\alpha(\mathbf{k}_1, \mathbf{k}_2)F_{n-m}(\mathbf{p}_{m+1}, \dots, \mathbf{p}_n) + 2n\beta(\mathbf{k}_1, \mathbf{k}_2)G_{n-m}(\mathbf{p}_{m+1}, \dots, \mathbf{p}_n)]. \quad (2.99)$$

The recursion starts with $F_1 = G_1 = 1$ and the first non-trivial kernels are

$$F_2(k_1, k_2) = \frac{5}{7} + \frac{1}{2}\mu(k_1, k_2) \left(\frac{k_1}{k_2} + \frac{k_2}{k_1} \right) + \frac{2}{7}\mu(k_1, k_2)^2 \quad (2.100)$$

and

$$G_2(k_1, k_2) = \frac{3}{7} + \frac{1}{2}\mu(k_1, k_2) \left(\frac{k_1}{k_2} + \frac{k_2}{k_1} \right) + \frac{4}{7}\mu(k_1, k_2)^2 \quad (2.101)$$

with $\mu(k_1, k_2) = \mathbf{k}_1 \cdot \mathbf{k}_2 / (k_1 k_2)$ being the cosine of the angle between the two vectors.

Statistics

Since velocities in the Universe are not directly observable, correlators of the matter density contrast are the main tool to study the LSS. The goal is to express all correlators in terms of the linear power spectrum, P_L , that defines the initial conditions and can be computed via Einstein-Boltzmann codes (see: section 2.2.5). The leading order power spectrum is simply given by the linearly evolved initial power spectrum

$$\langle \delta^{(1)}(\mathbf{k}_1) \delta^{(1)}(\mathbf{k}_2) \rangle = (2\pi)^3 \delta^D(\mathbf{k}_1 + \mathbf{k}_2) P_L(k_1). \quad (2.102)$$

At leading order, the gravitational bispectrum is zero and the first non-zero order is given by

$$\begin{aligned} \langle \delta(\mathbf{k}_1) \delta(\mathbf{k}_2) \delta(\mathbf{k}_3) \rangle &= \langle \delta^{(1)}(\mathbf{k}_1) \delta^{(1)}(\mathbf{k}_2) \delta^{(2)}(\mathbf{k}_3) \rangle + 2 \text{cyclic} \\ &= (2\pi)^3 \delta^D(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) 2F_2(k_1, k_2) P_L(k_1) P_L(k_2) + \text{cyclic} \end{aligned} \quad (2.103)$$

where we used

$$\delta^{(2)}(\mathbf{k}) = \int \frac{d^3 p}{(2\pi)^3} F_2(\mathbf{k} - \mathbf{p}, \mathbf{p}) \delta^{(1)}(\mathbf{k} - \mathbf{p}) \delta^{(1)}(\mathbf{p}). \quad (2.104)$$

Thus, the tree-level matter bispectrum is given by

$$B_{\text{mmm}}(k_1, k_2, k_3) = 2F_2(k_1, k_2) P_L(k_1) P_L(k_2) + \text{cyclic} \quad (2.105)$$

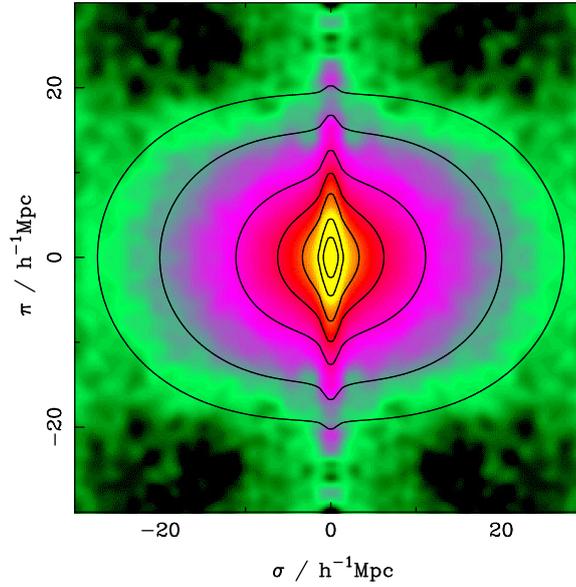


Fig. 2.4 Two-point correlation function measured in the 2dF galaxy redshift survey [6]. The colour is related to the clustering strength and increases from black to yellow. π measures radial and σ transverse distances. The black lines show equi-correlation lines and even in the absence of RSDs they would not be circular due to the Alcock-Paczynski effect. One would need perfect knowledge of the underlying cosmology to have consistent transverse and radial distance scales. On large scales, the Kaiser effect leads to a squashing because correlations parallel to the LOS are enhanced. On small scales, the FoG effect suppresses correlations parallel to the LOS and thus the equi-correlation lines appear stretched in the radial direction.

2.3.3 Redshift-space distortions

Galaxies'⁴ radial distances are measured via their (cosmological) redshifts. However, peculiar velocities parallel to the line-of-sight (LOS) give rise to a Doppler redshift that is degenerate with the cosmological redshift and biases distance measurements. On large scales, RSDs are caused by the coherent infall of galaxies into gravitational potentials and leads to an enhancement of modes parallel to the LOS [69], first described perturbatively in [70]. Fig. 2.4 shows the resulting squashing of the two-point correlation function on large scales.

The observed redshift-space positions, \mathbf{s} , are related to Eulerian space via [71]

$$\mathbf{s} = \mathbf{x} + \frac{u_{\parallel}}{\mathcal{H}} \mathbf{e}_{\parallel} = \mathbf{x} + \frac{\mathbf{e}_{\parallel} \cdot (\nabla \theta(\mathbf{x}))}{\mathcal{H}} \mathbf{e}_{\parallel}. \quad (2.106)$$

⁴For concreteness, we will use the term galaxies for tracers, but the discussion translates directly to other tracers such as neutral hydrogen.

The parallel and perpendicular indices are always referring to directions with respect to the LOS. The mapping from Eulerian to redshift-space preserves the mass locally

$$(1 + \delta^{(s)})d^3s = (1 + \delta)d^3x. \quad (2.107)$$

This allows us to express the density contrast in redshift-space

$$\delta^{(s)}(\mathbf{s}) = \frac{\delta(\mathbf{x}) + 1 - J}{J} \quad (2.108)$$

in terms of the Jacobian, J , of the mapping $\mathbf{x} \rightarrow \mathbf{s}$

$$J = \det \frac{\partial \mathbf{s}}{\partial \mathbf{x}}. \quad (2.109)$$

In Fourier space, we find

$$\begin{aligned} \delta^{(s)}(\mathbf{k}) &= \int d^3s e^{i\mathbf{k}\mathbf{s}} = \int d^3x J e^{i\mathbf{k}\mathbf{x}} e^{ik_{\parallel}u_{\parallel}} \frac{\delta(\mathbf{x}) + 1 - J}{J} \\ &= \int d^3x e^{i\mathbf{k}\mathbf{x}} e^{ifk_{\parallel}u_{\parallel}} (\delta(\mathbf{x}) + f\nabla_{\parallel}u_{\parallel}(\mathbf{x})) \\ &= \sum_{n=1}^{\infty} \prod_{i=1}^n \left(\int \frac{d^3p_i}{(2\pi)^3} \right) (2\pi)^3 \delta^D(\mathbf{k} - \mathbf{p}_{1:n}) \times \\ &\quad \times (\delta(\mathbf{p}_1) + f\mu_1^2 \theta(\mathbf{p}_1)) \frac{(f\mu k)^{n-1}}{(n-1)!} \frac{\mu_2}{p_2} \theta(\mathbf{p}_2) \cdots \frac{\mu_n}{p_n} \theta(\mathbf{p}_n) \end{aligned} \quad (2.110)$$

where $\mu_i = \mu(\mathbf{k}_i, \mathbf{e}_{\parallel}) = k_{i,\parallel}/k_i$ is the cosine of the angle between the wave vector and the LOS. The multitude of additional terms from the expansion of the exponential function sheds a light on the intrinsic non-linearities of RSDs. In practice, those non-linearities lead to a lower accuracy of the perturbative expansion for a given k than the pure matter predictions [71]. Expanding all fields in terms of their series expansions (2.88 and 2.89) allows us to introduce the redshift kernels Z_i

$$\delta^{(s)}(\mathbf{k}, \tau) = \sum_{n=1}^{\infty} a^n(\tau) \prod_{i=1}^n \left(\int \frac{d^3p_i}{(2\pi)^3} \delta^{(1)}(p_i) \right) (2\pi)^3 \delta^D(\mathbf{k} - \mathbf{p}_{1:n}) Z_n(p_1, \dots, p_n). \quad (2.111)$$

The first two redshift kernels are given by [71]

$$Z_1(\mathbf{k}) = 1 + f\mu(\mathbf{k}, \mathbf{e}_{\parallel})^2 \quad (2.112)$$

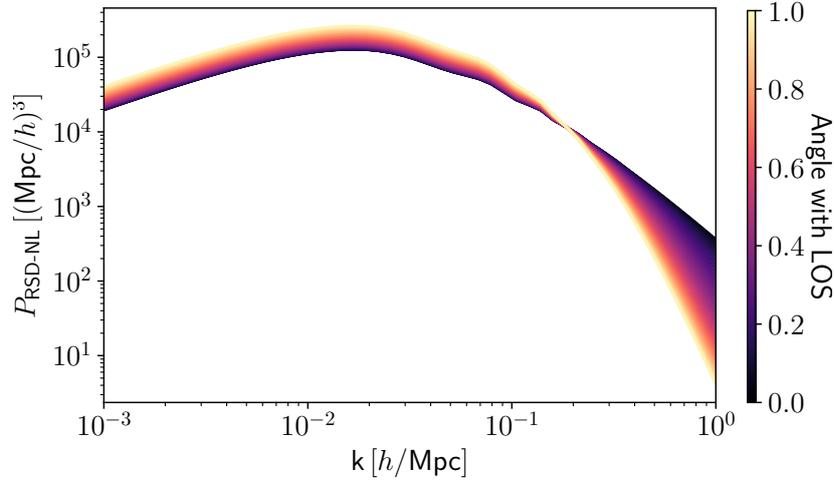


Fig. 2.5 The full RSD power spectrum is shown as a function of the wave vector k , and the cosine of the angle with the LOS, μ . On large scales, the Kaiser effects enhance modes parallel to the LOS while on small scales, those modes are damped by the FoG effect.

and

$$Z_2(\mathbf{k}_1, \mathbf{k}_2) = F_2(\mathbf{k}_1, \mathbf{k}_2) + f\mu_{12}^2 G_2(\mathbf{k}_1, \mathbf{k}_2) + \frac{f\mu_{12}k_{12}}{2} \left(\frac{\mu_1}{k_1} Z_1(\mathbf{k}_1) + \frac{\mu_2}{k_2} Z_1(\mathbf{k}_2) \right) \quad (2.113)$$

where, again, $\mu_i = k_{i,\parallel}/k_i$ and $k_{ij}^2 = (\mathbf{k}_i + \mathbf{k}_j)^2$ and $\mu_{ij}k_{ij} = \mu_i k_i + \mu_j k_j$.

Statistics

Correlating the density contrast in redshift-space, we find the aforementioned enhancement of large modes parallel to the LOS in the power spectrum, first described by Kaiser [70]

$$P_{\text{Kaiser}}(k, \mu) = Z_1(k)^2 P_L(k) = (1 + f\mu^2)^2 P_L(k). \quad (2.114)$$

Here, $\mu = k_{\parallel}/k$ is the cosine of the angle between wave vector and the LOS. Similarly, the tree-level redshift-space bispectrum is given by

$$B_{\text{Kaiser}}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = 2Z_1(\mathbf{k}_1)P_L(k_1)Z_1(\mathbf{k}_2)P_L(k_2)Z_2(\mathbf{k}_1, \mathbf{k}_2). \quad (2.115)$$

non-linear RSDs

On small scales, the large and incoherent velocities of galaxies within potential wells give rise to the Finger of God (FoG) effect, i.e. structures appear elongated along the LOS [72]. In Fig. 2.4, one can see the corresponding suppression of small scale clustering parallel to the LOS in the two-point correlation function. This corresponds to a damping of small-scale modes parallel to the LOS. The theoretical predictions can be extended to smaller scales by modelling the non-linear aspects of RSDs directly.

Typically, the modelling is done by multiplying the perturbative spectra with a damping factor, D_{FoG} , coming from the pairwise-velocity dispersion [73]

$$P_{\text{NL-RSD}}(k, \mu) = D_{\text{FoG}}(k\mu)P_{\text{Kaiser}}(k, \mu). \quad (2.116)$$

Depending on the distributional assumption of the pairwise-velocity dispersion, one can derive a Lorentzian [74, 75]

$$D_{\text{FoG}}(k\mu) = (1 + (k\mu\sigma_P)^2/2)^{-2} \quad (2.117)$$

or Gaussian [76, 77]

$$D_{\text{FoG}}(k\mu) = \exp[-(k\mu\sigma_P)^2] \quad (2.118)$$

damping factor for the power spectrum. In perturbation theory, the pairwise velocity dispersion is given by

$$\sigma_v^2 = \frac{(f\mathcal{H})^2}{6\pi^2} \int dk P(k). \quad (2.119)$$

However, since the model is meant to capture a non-linear effect, one typically treats the velocity dispersion as a free parameter that can be fitted against simulations. In Fig. 2.5, we illustrate the non-linear RSD power spectrum (2.116). On large scales, one sees the Kaiser enhancement while on small scales, modes parallel to the LOS are damped due to the FoG effect. A similar Ansatz was proposed for the redshift-space bispectrum [71]

$$B_{\text{NL-RSD}}(k_1, k_2, k_3, \mu_1, \mu_2, \mu_3) = D_{\text{FoG}}(k_1\mu_1, k_2\mu_2, k_3\mu_3)B_{\text{Kaiser}}(k_1, \dots, \mu_3) \quad (2.120)$$

where the damping was achieved by a generalised Lorentzian term

$$D_{\text{FoG}}(k_1\mu_1, k_2\mu_2, k_3\mu_3) = (1 + ((k_1\mu_1)^2 + (k_2\mu_2)^2 + (k_3\mu_3)^2)^2 \sigma_B^4/2)^{-2}. \quad (2.121)$$

The deviation from the pure Lorentzian form was chosen to achieve a better fit in equilateral configurations. A further (phenomenological) improvement of this model was performed

in [78]. The Gaussian damping generalises to

$$D_{\text{FoG}}(k_1\mu_1, k_2\mu_2, k_3\mu_3) = \exp \left[- \left((k_1\mu_1)^2 + (k_2\mu_2)^2 + (k_3\mu_3)^2 \right) \sigma_B^2 / 2 \right]. \quad (2.122)$$

As we see later, separable bispectra are beneficial for numerical reasons. The modified Lorentzian damping (2.121) is non-separable but using the pure Lorentzian form, the bispectrum becomes separable using

$$\frac{1}{t^2} = \int_0^\infty d\lambda \lambda \exp(-t\lambda). \quad (2.123)$$

The Gaussian template is directly separable and, in addition, allows us to absorb uncertainties in the redshift measurements in the velocity dispersion [79].

2.3.4 Lagrangian perturbation theory

In Lagrangian perturbation theory (LPT), we study the displacement field Ψ that realises the mapping of mass elements from the initial positions \mathbf{q} to the late-time positions $\mathbf{x}(\mathbf{q}, \tau)$,

$$\mathbf{x}(\mathbf{q}, \tau) = \mathbf{q} + \Psi(\mathbf{q}, \tau). \quad (2.124)$$

As with the mapping from Eulerian to redshift-space, mass conservation implies

$$\rho(\mathbf{x}, \tau) d^3x = \bar{\rho}(\tau) d^3q, \quad (2.125)$$

which in turn allows us to express the density contrast at late times in terms of the Jacobian, $J(\mathbf{q}, \tau) = \det(\delta_{ij}^K + \Psi_{i,j})$, of the transformation $\mathbf{x} \rightarrow \mathbf{q}$ via

$$\delta(\mathbf{x}, \tau) = \frac{\rho}{\bar{\rho}} - 1 = \frac{1}{J(\mathbf{q}, \tau)} - 1. \quad (2.126)$$

In order to derive an evolution equation for the displacement field, we start with the Euler equation (2.86)

$$\frac{d^2x}{d\tau^2} + \mathcal{H} \frac{dx}{d\tau} = -\nabla_{\mathbf{x}}\phi$$

and rewrite it in terms of the Lagrangian coordinates \mathbf{q}

$$J \left[\frac{d^2\Psi_{i,j}}{d\tau^2} + \mathcal{H} \frac{d\Psi_{i,j}}{d\tau} \right] = (\delta_{ij} + \Psi_{i,j}) \frac{3}{2} \Omega_m(\tau) \mathcal{H}^2 (J - 1). \quad (2.127)$$

This equation can be solved by means of a series expansion in the displacement field

$$\Psi(\mathbf{k}, \tau) = \sum_{n=1}^{\infty} D^{(n)}(\tau) \Psi^{(n)}(\mathbf{k}). \quad (2.128)$$

Using the expansion for the determinant

$$\begin{aligned} \det(I + A) &= \exp(\text{tr}(\log(I + A))) = \exp \left[\text{tr} \left(\sum_{i=1}^{\infty} \frac{(-A)^i}{i} \right) \right] \\ &= 1 + \text{tr}A + \frac{1}{2} [\text{tr}^2(A) - \text{tr}A^2] + \mathcal{O}(A^3) \end{aligned} \quad (2.129)$$

one finds a closed-form solution for the n -th order displacement field in Fourier space

$$\Psi^{(n)}(\mathbf{k}) = -i \frac{D^n(\tau)}{n!} \prod_{i=1}^n \left(\int \frac{d^3 p_i}{(2\pi)^3} \delta^{(1)}(\mathbf{p}_i) \right) (2\pi)^3 \delta^D(\mathbf{k} - \mathbf{p}_{1:n}) \mathbf{L}^{(n)}(\mathbf{p}_1, \dots, \mathbf{p}_n) \quad (2.130)$$

where we introduce the kernels $\mathbf{L}^{(n)}$. The first two kernels are given as

$$\begin{aligned} L^{(1)}(\mathbf{p}_1) &= \frac{\mathbf{p}_1}{p_1^2} \quad \text{Zel'dovich approximation} \\ L^{(2)}(\mathbf{p}_1, \mathbf{p}_2) &= \frac{3}{7} \frac{\mathbf{p}_{1:2}}{p_{1:2}^2} (1 - \mu^2(\mathbf{p}_1, \mathbf{p}_2)). \end{aligned} \quad (2.131)$$

Using (2.126) allows us to express the density contrast at late times in terms of the displacement field

$$\delta(\mathbf{k}) = \int d^3 x e^{i\mathbf{k}\mathbf{x}} \delta(\mathbf{x}) = \int d^3 q e^{i\mathbf{k}\mathbf{q}} \left(e^{i\mathbf{k}\Psi} - 1 \right). \quad (2.132)$$

Expanding the exponential in Ψ demonstrates the equivalence between SPT and LPT. At linear order, one finds

$$\delta^{(1, \text{LPT})}(\mathbf{k}) = \int d^3 q e^{i\mathbf{k}\mathbf{q}} i\mathbf{k}\Psi^{(1)}(\mathbf{q}) \delta^{(1)}(\mathbf{q}) = \delta^{(1, \text{SPT})}(\mathbf{k}). \quad (2.133)$$

At second order, the computation is slightly more involved

$$\begin{aligned}
\delta^{(2,\text{LPT})}(\mathbf{k}) &= \int \frac{d^3q}{(2\pi)^3} e^{i\mathbf{k}\mathbf{q}} \left(i\mathbf{k}\Psi^{(2)}(\mathbf{q}) + \frac{1}{2}(\mathbf{k}\Psi^{(1)}(\mathbf{q}))^2 \right) \\
&= \frac{1}{2} \int \frac{d^3p_1}{(2\pi)^3} \frac{d^3p_2}{(2\pi)^3} \delta^D(\mathbf{k} - \mathbf{p}_1 + \mathbf{p}_2) \delta^{(1)}(\mathbf{p}_1) \delta^{(1)}(\mathbf{p}_2) \times \\
&\quad \times \left[\frac{3}{7} \left(1 - \frac{(\mathbf{p}_1 \cdot \mathbf{p}_2)^2}{(p_1 p_2)^2} \right) + \frac{\mathbf{k}\mathbf{p}_1 \cdot \mathbf{k}\mathbf{p}_2}{p_1^2 p_2^2} \right] \\
&= \int \frac{d^3p}{(2\pi)^3} F_2(\mathbf{k} - \mathbf{p}, \mathbf{p}) \delta^{(1)}(\mathbf{k} - \mathbf{p}) \delta^{(1)}(\mathbf{p}) = \delta^{(2,\text{SPT})}(\mathbf{k}).
\end{aligned} \tag{2.134}$$

but one nevertheless recovers the SPT result.

2.3.5 Redshift-space distortions in LPT

LPT offers a natural way to include RSDs into the formalism. The displacement field then describes the mapping from Lagrangian to redshift, instead of Eulerian, space

$$\mathbf{s} = \mathbf{x} + \frac{u_{\parallel}}{\mathcal{H}} \mathbf{e}_{\parallel} = \mathbf{q} + \Psi^s. \tag{2.135}$$

The displacement field to redshift-space, Ψ^s , is related to the displacement field to Eulerian space, Ψ , via

$$\Psi^s = \Psi + \frac{\mathbf{e}_{\parallel} \cdot (d\Psi/d\tau)}{\mathcal{H}} \mathbf{e}_{\parallel} \tag{2.136}$$

where we used $\mathbf{u} = d\mathbf{x}/d\tau = d\Psi/d\tau$. Using the expansion of the displacement field (2.130) reveals

$$d\Psi^{(n)}/d\tau = n\mathcal{H}f\Psi^{(n)}. \tag{2.137}$$

So, the n -th order displacement field to redshift-space is given by

$$\Psi^{s(n)} = \Psi^{(n)} + nf \left(\mathbf{e}_{\parallel} \cdot \Psi^{(n)} \right) \mathbf{e}_{\parallel}. \tag{2.138}$$

At leading order, we find

$$\Psi^{s(1)}(\mathbf{k}) = -iD\delta^{(1)}(\mathbf{k}) \left(\frac{\mathbf{k}}{k^2} + f \frac{k_{\parallel} \mathbf{e}_{\parallel}}{k^2} \right) \tag{2.139}$$

which implies

$$\delta^{(1,s)}(\mathbf{k}) = \int \frac{d^3q}{(2\pi)^3} e^{i\mathbf{k}\mathbf{q}} \mathbf{k}\Psi^{(1,s)}(\mathbf{q}) = \delta^{(1)}(\mathbf{k})(1 + f\mu^2). \tag{2.140}$$

This, again, matches the result from SPT with the redshift kernel Z_1 . At second order, one recovers the SPT result too

$$\begin{aligned}
\delta^{(2,s)}(\mathbf{k}, \tau) &= \int \frac{d^3q}{(2\pi)^3} e^{i\mathbf{k}\mathbf{q}} \left(i\mathbf{k}\Psi^{(2,s)}(\mathbf{q}) + \frac{1}{2}(\mathbf{k}\Psi^{(1,s)}(\mathbf{q}))^2 \right) \\
&= \frac{1}{2} \int \frac{d^3p_1}{(2\pi)^3} \frac{d^3p_2}{(2\pi)^3} \delta^D(\mathbf{p}_1 + \mathbf{p}_2 - \mathbf{k}) \delta^{(1)}(\mathbf{p}_1) \delta^{(1)}(\mathbf{p}_2) \times \\
&\quad \times \left[\frac{3}{7} (1 + 2f\mu^2) \left(1 - \frac{(\mathbf{p}_1\mathbf{p}_2)^2}{(p_1p_2)^2} \right) + \mathbf{k} \cdot \left(\frac{\mathbf{p}_1}{p_1^2} + f \frac{p_{1,z}\mathbf{e}_{\parallel}}{p_1^2} \right) \mathbf{k} \cdot \left(\frac{\mathbf{p}_2}{p_2^2} + f \frac{p_{2,z}\mathbf{e}_{\parallel}}{p_2^2} \right) \right] \\
&= \frac{1}{2} \int \frac{d^3p_1}{(2\pi)^3} \frac{d^3p_2}{(2\pi)^3} \delta^D(\mathbf{p}_1 + \mathbf{p}_2 - \mathbf{k}) \delta^{(1)}(\mathbf{p}_1) \delta^{(1)}(\mathbf{p}_2) \left[2F_2(\mathbf{p}_1, \mathbf{p}_2) \right. \\
&\quad \left. + f \left(\frac{6}{7}\mu^2 \left(1 - \frac{(\mathbf{p}_1\mathbf{p}_2)^2}{(p_1p_2)^2} \right) + (p_{1,z} + p_{2,z}) \left[\left(1 + \frac{\mathbf{p}_1\mathbf{p}_2}{p_1^2} \right) \frac{p_{1,z}}{p_1^2} + \left(1 + \frac{\mathbf{p}_1\mathbf{p}_2}{p_1^2} \right) \frac{p_{2,z}}{p_2^2} \right] \right) \right. \\
&\quad \left. + f^2 (p_{1,z} + p_{2,z})^2 \left(\frac{p_{1,z}p_{2,z}}{p_1^2p_2^2} \right) \right] \\
&= \int \frac{d^3p_1}{(2\pi)^3} \frac{d^3p_2}{(2\pi)^3} \delta^D(\mathbf{p}_1 + \mathbf{p}_2 - \mathbf{k}) \delta^{(1)}(\mathbf{p}_1) \delta^{(1)}(\mathbf{p}_2) \\
&\quad \times \left[F_2(\mathbf{p}_1, \mathbf{p}_2) + f \frac{\mu_{12}k_{12}}{2} \left(\frac{\mu_1}{p_1} + \frac{\mu_2}{p_2} \right) + f\mu_{12}^2 G_2(\mathbf{p}_1, \mathbf{p}_2) + f^2 \frac{\mu_{12}k_{12}}{2} \mu_1\mu_2 \left(\frac{\mu_1}{p_2} + \frac{\mu_2}{p_1} \right) \right].
\end{aligned} \tag{2.141}$$

One advantage of the Lagrangian approach is that it allows us to perform a resummation before expanding the exponential. This can be used to refine the modelling on non-linear RSDs, e.g. [80–82].

2.3.6 Tracer bias

In most observations, one cannot observe the smooth dark matter field directly and must work with tracers that are observable. Those tracers do not follow the underlying dark matter directly, a property called bias. Expanding the tracer density contrast, δ_g , in terms of all functions of the gravitational potential allowed by symmetry yields in Fourier space to second order [83–85]

$$\delta_g(\mathbf{k}) = b_1 \delta(\mathbf{k}) + \frac{b_2}{2} \delta^2(\mathbf{k}) + \frac{b_{s^2}}{2} s^2(\mathbf{k}) \tag{2.142}$$

where δ is the matter density contrast and s^2 is the non-local tidal term

$$s^2(\mathbf{k}) = \int \frac{d^3p}{(2\pi)^3} S_2(\mathbf{k} - \mathbf{p}, \mathbf{p}) \delta(\mathbf{k} - \mathbf{p}) \delta(\mathbf{p}). \tag{2.143}$$

The tidal tensor is

$$S_2(\mathbf{p}_1, \mathbf{p}_2) = \frac{\mathbf{p}_1 \cdot \mathbf{p}_2}{p_1 p_2} - \frac{1}{3}. \quad (2.144)$$

A priori, one cannot predict these bias parameters and so they have to be estimated together with the cosmological parameters.

Lagrangian Biasing

Starting with a local bias expansion in Lagrangian space sheds a light on the necessity of the tidal term for a consistent bias expansion. Translating the bias expansion (2.142) into Lagrangian space yields

$$\delta_g^{(L)}(\mathbf{q}) = b_1^{(L)} \delta(\mathbf{q}) + \frac{b_2^{(L)}}{2} (\delta^2(\mathbf{q}) - \langle \delta^2 \rangle) + \frac{b_{s^2}^{(L)}}{2} \tilde{s}^2(\mathbf{q}). \quad (2.145)$$

The mapping from Lagrangian to Eulerian space can then be derived from the continuity equations of the matter and galaxy fields

$$[1 + \delta(\mathbf{x})] d^3x = d^3q \quad \text{and} \quad [1 + \delta_g^{(E)}(\mathbf{x})] d^3x = [1 + \delta_g^{(L)}(\mathbf{q})] d^3q. \quad (2.146)$$

Combining those equations yields

$$\delta_g^{(E)}(\mathbf{x}) = \delta_g^{(L)}(\mathbf{q}) + \delta(\mathbf{x}) \delta_g^{(L)}(\mathbf{q}) + \delta(\mathbf{x}). \quad (2.147)$$

Expanding to first order, we find

$$\delta_g^{(E,1)}(\mathbf{x}) = \delta_g^{(L,1)}(\mathbf{x}) + \delta^{(1)}(\mathbf{x}) = (b_1^{(L)} + 1) \delta^{(1)}(\mathbf{x}). \quad (2.148)$$

At second order, we find for the galaxy density contrast in Eulerian space

$$\begin{aligned} \delta_g^{(E,2)}(\mathbf{x}) &= \delta_g^{(L,2)}(\mathbf{x}) - \Psi^{(1)}(\mathbf{x}) \nabla_{\mathbf{x}} \delta_g^{(L,1)}(\mathbf{x}) + \delta_g^{(L,1)}(\mathbf{x}) \delta^{(1)}(\mathbf{x}) + \delta^{(2)}(\mathbf{x}) \\ &= \delta_g^{(L,2)} - b_1^{(L)} \left[\frac{17}{21} \delta^2 + \frac{2}{7} s^2 - \delta^{(2)} \right] + \delta_g^{(L,1)} \delta^{(1)} + \delta^{(2)} \\ &= (b_1^{(L)} + 1) \delta^{(2)} + \frac{1}{2} \left(b_2^{(L)} + \frac{8}{21} b_1^{(L)} \right) \delta^2 + \frac{1}{2} \left(b_{s^2}^{(L)} - \frac{4}{7} b_1^{(L)} \right) \tilde{s}^2, \end{aligned} \quad (2.149)$$

where we used

$$\delta^{(2)}(\mathbf{x}) = \frac{17}{21} \delta^2(\mathbf{x}) - \Psi(\mathbf{x}) \nabla_{\mathbf{x}} \delta(\mathbf{x}) + \frac{2}{7} s^2(\mathbf{x}). \quad (2.150)$$

This allows us to summarise the relation between bias parameters at early and late times

$$b_1^{(E)} = b_1^{(L)} + 1, \quad b_2^{(E)} = b_2^{(L)} + \frac{8}{21}b_1^{(L)}, \quad b_{s^2}^{(E)} = b_{s^2}^{(L)} - \frac{4}{7}b_1^{(L)}. \quad (2.151)$$

In particular, we see that a local bias model in Lagrangian space introduces a tidal component at late times. In addition to this contribution from the evolution, there is now evidence for an initial non-local bias [86, 87].

Stochastic bias

In most observations the tracers are discrete objects. This introduces further biases often referred to as shot-noise. Whether or not a discrete tracer forms at some position is modelled by a random process since it is unfeasible to try to measure and understand the small scale physics that would explain whether or not a tracer forms. The probability to have n tracers within some volume is normally modelled by an (inhomogeneous) Poisson process [88, 89, 68].

To gain an intuitive understanding of stochastic biases, we present here a calculation from [90]. We start by writing the overdensity of discrete tracers, $\delta^{(d)}$, in a given realization

$$\delta^{(d)}(\mathbf{x}) = \frac{n(\mathbf{x})}{\bar{n}} - 1 = \frac{1}{\bar{n}} \sum_i \delta^D(\mathbf{x} - \mathbf{x}_i) - 1 \quad (2.152)$$

where \bar{n} is the average number density of tracers and the \mathbf{x}_i are the positions of the tracers. Using the Poissonian assumption, one can compute correlation functions. For the two-point correlation function of the discrete tracers, one finds

$$\begin{aligned} \langle \delta^{(d)}(\mathbf{x}) \delta^{(d)}(\mathbf{x}') \rangle &= \frac{1}{\bar{n}^2} \left\langle \sum_{i,j} \delta^D(\mathbf{x} - \mathbf{x}_i) \delta^D(\mathbf{x}' - \mathbf{x}_j) \right\rangle - \frac{2}{\bar{n}} \left\langle \sum_i \delta^D(\mathbf{x} - \mathbf{x}_i) \right\rangle + 1 \\ &= \frac{1}{\bar{n}^2} \delta^D(\mathbf{x} - \mathbf{x}') \left\langle \sum_i \delta^D(\mathbf{x}' - \mathbf{x}_i) \right\rangle + \frac{1}{\bar{n}^2} \left\langle \sum_{i \neq j} \delta^D(\mathbf{x} - \mathbf{x}_i) \delta^D(\mathbf{x}' - \mathbf{x}_j) \right\rangle - 1 \\ &= \frac{1}{\bar{n}} \delta^D(\mathbf{x} - \mathbf{x}') + \frac{1}{\bar{n}^2} \left\langle \sum_{i \neq j} \delta^D(\mathbf{x} - \mathbf{x}_i) \delta^D(\mathbf{x}' - \mathbf{x}_j) \right\rangle - 1 \\ &= \frac{1}{\bar{n}} \delta^D(\mathbf{x} - \mathbf{x}') + \xi(|\mathbf{x} - \mathbf{x}'|). \end{aligned} \quad (2.153)$$

Beside the two-point correlation function of the smooth matter field, ξ , one finds a shot-noise contribution. This contribution leads in Fourier space to a constant off-set of the discrete

power spectrum, $P^{(d)}$, compared to the power spectrum of the smooth field P

$$P^{(d)}(k) = \frac{1}{\bar{n}} + \int d^3x \xi(|\mathbf{x}|) \exp[i\mathbf{k} \cdot \mathbf{x}] = \frac{1}{\bar{n}} + P(k). \quad (2.154)$$

A similar calculation shows, for the three-point correlation function of discrete tracers,

$$\begin{aligned} \langle \delta^{(d)}(\mathbf{x}) \delta^{(d)}(\mathbf{x}') \delta^{(d)}(\mathbf{x}'') \rangle &= \zeta^{(3)}(\mathbf{x}, \mathbf{x}', \mathbf{x}'') \\ &+ \frac{1}{\bar{n}} \left(\delta^D(\mathbf{x} - \mathbf{x}') \xi(\mathbf{x}' - \mathbf{x}'') + 2 \text{ cyclic} \right). \\ &+ \frac{1}{\bar{n}^2} \delta^D(\mathbf{x} - \mathbf{x}') \delta^D(\mathbf{x}' - \mathbf{x}''). \end{aligned} \quad (2.155)$$

Here, we find two additional terms beside the three-point correlation function $\zeta^{(3)}$ of the continuous field. Taking the Fourier transform yields

$$B^{(d)}(k_1, k_2, k_3) = B(k_1, k_2, k_3) + \frac{1}{\bar{n}} \left(P^{(d)}(k_1) + P^{(d)}(k_2) + P^{(d)}(k_3) - \frac{3}{\bar{n}} \right) + \frac{1}{\bar{n}^2} \quad (2.156)$$

with the last term being the shot-noise of the discrete bispectrum.

To capture those effects, one includes stochastic terms to the bias expansion (2.142). Accounting for stochastic bias, the complete bias expansion to second order reads [83]

$$\delta_g(\mathbf{k}) = b_1 \delta(\mathbf{k}) + \frac{b_2}{2} \delta^{(2)}(\mathbf{k}) + \frac{b_{s^2}}{2} s^2(\mathbf{k}) + \varepsilon(\mathbf{k}) + (\varepsilon_\delta * \delta)(\mathbf{k}). \quad (2.157)$$

By construction, the mean of the bias terms, ε and ε_δ , and cross-correlation of the bias terms with the (long wavelength) matter fields are zero. The only non-vanishing correlators up to second order are then [88, 89, 68]

$$P_{\varepsilon\varepsilon} = \frac{1}{\bar{n}}, \quad P_{\varepsilon\varepsilon_\delta} = \frac{b_1}{2\bar{n}}, \quad B_{\varepsilon\varepsilon\varepsilon} = \frac{1}{\bar{n}^2}. \quad (2.158)$$

2.3.7 Primordial non-Gaussianities

The assumption of Gaussian initial conditions can be tested by allowing deviations from them and constraining their amplitude f_{NL} . This is typically done by adding primordial non-Gaussianities (PNGs) of known shape. Those give rise to two types of new terms: Firstly, the added PNGs lead to a non-zero matter bispectrum at all times. This primordial bispectrum grows with D^3 whereas the gravitational bispectrum grows with D^4 . In practice this means that at late times even for $f_{\text{NL}} = 1$, the primordial bispectrum of the local, equilateral and orthogonal shape are in most configurations significantly smaller than the gravitational

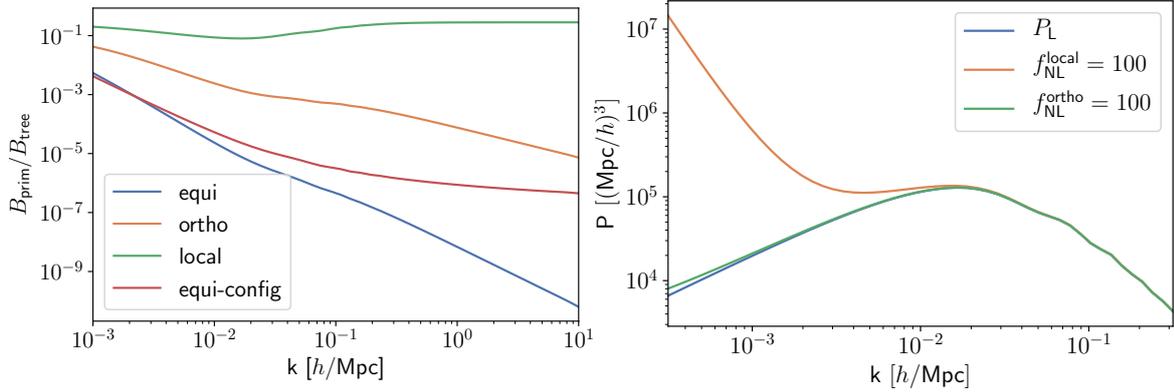


Fig. 2.6 Left: The primordial bispectrum of the three commonly used templates is orders of magnitude smaller than the gravitational bispectrum at $z = 0$ for $f_{\text{NL}} = 1$. We show the ratio of the local (green curve), equilateral (blue) and orthogonal (orange) templates and the tree-level gravitational bispectrum in the squeezed configurations $(k, k, 0.1k)$. Similarly in equilateral configurations, where the three templates are identical, the ratio is still very small (red curve). Right: The scale-dependent bias generates a characteristic feature at large scales. The equilateral template does not give rise to this effect.

bispectrum, see Fig. 2.6. Secondly, the bias expansion (2.157) has to be carried out both in the density and gravitational potential adding new bias terms [91–95]. Those scale-dependent bias terms offer for some shapes, particularly the local shape, the opportunity to constrain PNGs on large scales. See Fig. 2.6.

For generic, quadratic PNGs, one can perform a complete bias expansion in terms of the non-local transformation

$$\Psi(\mathbf{k}) = \int \frac{d^3 p}{(2\pi)^3} \left(\frac{k}{p}\right)^\alpha \Phi(\mathbf{p}) \quad (2.159)$$

of the gravitational potential, Φ , rather than the potential directly [96]. The parameter α is 0, 2, 1 for the local, equilateral and orthogonal shape respectively [97]. Adding the following terms to the bias expansion in Lagrangian space makes it complete (up to second order in the density contrast and first order in f_{NL}) [96]

$$\delta_g^{(NG)}(\mathbf{q}) = f_{\text{NL}} \left(b_\Psi^{(L)} \Psi(\mathbf{q}) + b_{\Psi\delta}^{(L)} \Psi(\mathbf{q}) \delta(\mathbf{q}) + \varepsilon_\Psi(\mathbf{q}) \Psi(\mathbf{q}) \right). \quad (2.160)$$

ε_Ψ captures the additional stochasticity due to the PNGs and its only non-zero correlator with the other fields at second order is [68]

$$P_{\varepsilon\varepsilon_\Psi}(k) = \frac{b_\Psi}{2\bar{n}}. \quad (2.161)$$

Translating the expansion into Eulerian space (up to second order), the convective term gives rise to a second non-local term, \tilde{n}^2 , [91–95, 98]

$$\begin{aligned}\delta_g^{(NG)}(\mathbf{x}) &= f_{\text{NL}} \left(b_{\Psi}^{(L)} \Psi - b_{\Psi}^{(L)} \Psi^{(LPT)}(\mathbf{x}) \nabla \Psi(\mathbf{x}) + b_{\Psi\delta}^{(L)} \Psi(\mathbf{x}) \delta(\mathbf{x}) + \varepsilon_{\Psi}(\mathbf{x}) \Psi(\mathbf{x}) \right) \\ &= f_{\text{NL}} \left(b_{\Psi}^{(E)} \Psi(\mathbf{x}) + b_{n^2}^{(E)} \tilde{n}^2(\mathbf{x}) + b_{\Psi\delta}^{(E)} \Psi(\mathbf{x}) \delta(\mathbf{x}) + \varepsilon_{\Psi}(\mathbf{x}) \Psi(\mathbf{x}) \right)\end{aligned}\quad (2.162)$$

Using the Zel'dovich approximation (2.131), the new term is in Fourier space given by [98]

$$n^2(\mathbf{k}) = \int \frac{d^3 p}{(2\pi)^3} N_2(\mathbf{p}, \mathbf{k} - \mathbf{p}) \frac{|\mathbf{k} - \mathbf{p}|^\alpha}{\mathcal{M}(|\mathbf{k} - \mathbf{p}|)} \delta(\mathbf{p}) \delta(\mathbf{k} - \mathbf{p}) \quad (2.163)$$

with

$$N_2(\mathbf{p}_1, \mathbf{p}_2) = \frac{\mathbf{p}_1 \mathbf{p}_2}{p_1^2}. \quad (2.164)$$

Comparing bias parameters in (2.162), we find

$$b_{\Psi}^{(E)} = b_{\Psi}^{(L)}, \quad b_{\Psi\delta}^{(E)} = b_{\Psi\delta}^{(L)} + b_{\Psi}^{(L)}, \quad b_{n^2}^{(E)} = b_{\Psi}^{(L)}. \quad (2.165)$$

In section 2.4.5, we are going to derive the scale dependence of the bias parameters b_{Ψ} and $b_{\Psi\delta}$ in the different PNG models. There, we also discuss in more detail the origin of the exponent α used in the non-local transformation of the gravitational potential (2.159).

The additional bias terms can be included into the redshift kernels, Z . The first two become [98, 99]

$$Z_1(\mathbf{k}) = b_1 + f\mu^2 + \frac{b_{\Psi}(k)}{M(k)}, \quad (2.166a)$$

$$\begin{aligned}Z_2(\mathbf{k}_i, \mathbf{k}_j) &= b_1 F_2 + \frac{b_2}{2} + \frac{b_{s^2}}{2} S_2 + f\mu_{ij} G_2 + f \frac{\mu_{ij} k_{ij}}{2} \left[\frac{\mu_i}{k_i} Z_1(\mathbf{k}_j) + \frac{\mu_j}{k_j} Z_1(\mathbf{k}_i) \right] \\ &\quad + \frac{1}{2} \left(\frac{(b_{\Psi\delta}(k_i) - b_{\Psi}(k_i) N_2(\mathbf{k}_j, \mathbf{k}_i))}{M(k_i)} + \frac{(b_{\Psi\delta}(k_j) - b_{\Psi}(k_j) N_2(\mathbf{k}_i, \mathbf{k}_j))}{M(k_j)} \right).\end{aligned}\quad (2.166b)$$

2.4 Halo model

A second approach to solve the non-linear evolution equations of the matter fields is the use of symmetries. Here, the idea is to identify the regions in the initial conditions/Lagrangian

space that will collapse into high-density objects, so called halos. Using spherical symmetry, we are then able to solve for the dynamics of those regions exactly.

2.4.1 Spherical collapse

Newton's spherical mass theorem and Birkhoff's theorem in GR guarantee that a homogeneous sphere evolves independently from its surroundings. This is why the evolution equations of a spherically symmetric, overdense region of radius R , within the initial conditions of a flat universe, are tractable. In fact, such a region evolves like a closed Universe (because it is overdense). We are now going to solve the evolution equation of such a region in an Einstein-deSitter Universe. The Friedman equation,

$$H^2 = H_0^2 \frac{\Omega_{m,0}}{a^3} - \frac{k}{a^2}, \quad (2.167)$$

can be integrated analytically. First, we change the time coordinate to conformal time $d\tau = dt/a$ and substitute $a = \frac{k}{H_0^2 \Omega_{m,0}} y^2$

$$\left(\frac{dy}{d\tau}\right)^2 = \frac{1}{4}k(1 - y^2). \quad (2.168)$$

This equation is integrated by

$$\int \frac{dy}{\sqrt{1 - y^2}} = \arcsin x + C. \quad (2.169)$$

Thus, the scale factor of the overdense region evolves like

$$a(\tau) = \frac{\Omega_{m,0}}{2(\Omega_{m,0} - 1)} \left(1 - \cos(\sqrt{k}\tau)\right) = A(1 - \cos \theta) \quad (2.170)$$

where $\theta = \sqrt{k}\tau$ and A is implicitly defined in the last equality. The integration constant was chosen so that the 'Big Bang' is at $\theta = 0$. The solution indicates that a overdense region first expands, then reaches a maximum at $\theta = \pi$ and finally collapses in a 'Big Crunch' at $\theta = 2\pi$. Since this 'Big Crunch' is unphysical, one instead assumes in the halo model that the overdense region forms a virialised object.

The cosmic time is recovered from the conformal time by

$$t = \int_0^t dt = \int_0^\tau a d\tau = \frac{\Omega_{m,0}}{H_0 2(\Omega_{m,0} - 1)^{3/2}} \left(\sqrt{k}\tau - \sin(\sqrt{k}\tau) \right) \\ = B(\theta - \sin \theta). \quad (2.171)$$

It is instructive to expand the parametric solutions of the overdense region to better understand the behavior at early times

$$a_{\text{SC}}(\theta) \simeq A \left(\theta^2/2 \left[1 - \frac{\theta^2}{12} \right] \right), \quad t_{\text{SC}} \simeq B\theta^3/6 \left[1 - \frac{\theta^2}{20} \right]. \quad (2.172)$$

Inverting the second expression and inserting into the first, we find that the scale factor of the overdense region evolves (up to next-to-leading order) like

$$a_{\text{SC}}(t) = \frac{A}{2} \left(\frac{6t}{B} \right)^{2/3} \left[1 - \frac{1}{20} \left(\frac{6t}{B} \right)^{2/3} \right] = a_{\text{EdS}}(t)(1 + \delta_R). \quad (2.173)$$

The leading order matches the background evolution of the Einstein-deSitter universe where $a_{\text{EdS}}(t) = (\frac{3}{2}H_0 t)^{2/3}$. δ_R is the radial perturbation, the relative amount by which the overdense sphere's radius deviates from the radius of a sphere in the background. We can compare the (linear) growth of structure too. Using the conservation of mass,

$$\rho_i R_i^3 = \rho R^3 = \rho_i (1 + \delta^{(1)}) R_i^3 (1 + \delta_R)^3, \quad (2.174)$$

we find that the linear overdensity evolves as

$$\delta^{(1)} = -3\delta_R = \frac{3}{20} [6(\theta - \sin \theta)]^{2/3}. \quad (2.175)$$

At the time of collapse, it approaches the so-called critical density, δ_c ,

$$\delta_{\text{collapse}}^{(1)} = -3\delta_R(t = 2\pi B) = \frac{3}{20} (12\pi)^{2/3} \simeq 1.69 \equiv \delta_c. \quad (2.176)$$

The fully non-linear density in the spherical collapse model, in contrast, is given by

$$\rho_{\text{SC}} = \left(\frac{a_{\text{EdS}}}{a_{\text{SC}}} \right)^3 = \frac{9}{2} \frac{(\theta - \sin \theta)^2}{(1 - \cos \theta)^3}, \quad (2.177)$$

where the normalization is chosen such that for $\theta \rightarrow 0$ the non-linear density approaches the background density. Fig. 2.7 illustrates the scale factor and density evolution in the linear

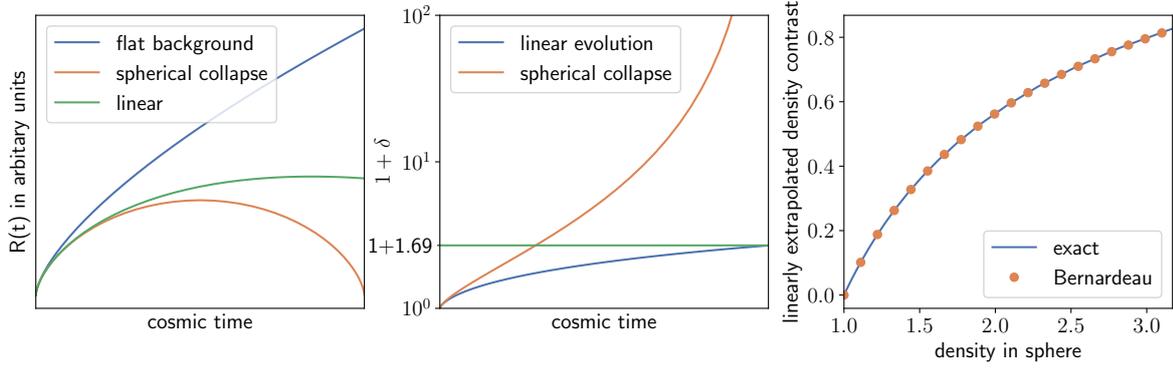


Fig. 2.7 Left: Evolution of the radius of a sphere following the background evolution (blue curve), the linear perturbative solution (green) and the fully non-linear spherical collapse (orange). Middle: Time evolution of the density in a sphere assuming linear evolution (blue curve) and spherical collapse (orange). The linear density approaches the critical density δ_c , when the spherical collapse solution becomes singular. Right: Illustration of the mapping between the density in a sphere from spherical collapse and the linear overdensity. We show the exact mapping in blue and the orange points show the approximation (2.178).

and non-linear scenario as a function of time. Solving (2.175) for θ allows us to express the density in the spherical collapse picture (2.177) in terms of the linear overdensity. This mapping can be approximated by [100]

$$\rho_{\text{SC}}(\delta^{(1)}) \simeq (1 - \delta^{(1)}/\nu)^{-\nu} \implies \delta^{(1)}(\rho_{\text{SC}}) \simeq \nu(1 - \rho_{\text{SC}}^{-1/\nu}). \quad (2.178)$$

Choosing $\nu = 21/13$ matches the high-redshift skewness obtained from perturbation theory. The mapping works for underdense regions too. The choice is correct to 0.5% from $\delta^{(1)} = 0.3 - 2.5$ [101, 7] and we show the mapping in the right subplot of Fig. 2.7.

2.4.2 Filtering of fields

There are situations, like the one just discussed, in which one is only interested in a coarse-grained version of a (stochastic) field. Mathematically, the coarse-grained field corresponds in real space to a convolution of the field with some window function. The convolution becomes a product in Fourier space and the power spectrum of the filtered field is thus given by

$$\delta_R(\mathbf{k}) = \delta(\mathbf{k})W(kR) \implies P_R(k) = |W(kR)|^2 P(k). \quad (2.179)$$

In many situations, a Top-Hat in real space window function,

$$\tilde{W}_{\text{TH}}(x/R) = \frac{3}{4\pi} \theta^H(1 - x/R), \quad (2.180)$$

is physically most meaningful. θ^H refers here to the Heavyside step function. Its Fourier transform in 3D is given by

$$W_{\text{TH}}(kR) = 3 \frac{\sin(kR) - kR \cos(kR)}{(kR)^3}. \quad (2.181)$$

Filtering can also be used to separate a field, like the density contrast δ , into small and large scales

$$\delta_L(\mathbf{k}) = \delta(\mathbf{k})W_L(k/k_{\text{max}}), \quad \delta_S(\mathbf{k}) = \delta(\mathbf{k})(1 - W_L(k/k_{\text{max}})). \quad (2.182)$$

In this situation a Top-Hat in Fourier space window function

$$W_L(k/k_{\text{max}}) = \theta^H(1 - |\mathbf{k}|/k_{\text{max}}), \quad (2.183)$$

simplifies many calculations, even though it does not have a clear interpretation in real space. The separation of perturbations into small and large scales is called peak-background split.

2.4.3 Mass functions

Spherical collapse provides a non-linear mapping of overdense regions in the initial conditions to virialised objects at late times. This allows us to compute the co-moving number density of virialised objects of mass M , called mass function $n(M)$, which is an important ingredient for many cosmological models. Assuming spherical symmetry and mass conservation, a virialised object of mass M corresponds to an initial overdensity of size R in Lagrangian space such that

$$M = \bar{\rho} \frac{4\pi}{3} R^3 = \bar{\rho} V_M. \quad (2.184)$$

Since the density fluctuations in the initial conditions are tiny, one can safely ignore them when converting masses and length scales. We will use the mass M and associated length scale $R(M)$ synonymously from now on. In Lagrangian space, the density in spheres of size $R(M)$ is well described by a Gaussian distribution with variance

$$\sigma_R^2(\tau) = \langle \delta_R^2 \rangle = \frac{1}{2\pi^2} \int P(k, \tau) W^2(kR) k^2 dk. \quad (2.185)$$

Press & Schechter's seminal idea was to linearly evolve the initial conditions and replace regions whose evolved linear density contrast exceeds the critical density δ_c by virialised objects. The probability that a given sphere in the initial conditions will be collapsed at time

τ is thus

$$\begin{aligned} P(\delta_R > \delta_c) &= \frac{1}{\sqrt{2\pi}\sigma_R} \int_{\delta_c}^{\infty} \exp\left(-\frac{x^2}{2\sigma_R(\tau)^2}\right) dx \\ &= \frac{1}{2} \operatorname{erfc}\left(\frac{\delta_c}{\sqrt{2}\sigma_R}\right) = \frac{1}{2} \operatorname{erfc}\left(\sqrt{\frac{v}{2}}\right). \end{aligned} \quad (2.186)$$

with the peak-height v . The model neatly separates the spherical collapse dynamics of the mapping that defined δ_c , from the statistical distribution of the initial conditions given by σ_R . Since the variance of the distribution increases with smaller scales, small objects form first and they subsequently merge into larger halos. However, we see that even in the limit of $\sigma_R \rightarrow \infty$ only half of the spheres of a given size will collapse. To match the expectation that all mass sits in virialised objects, the original paper introduced a fudge factor of 2.

The mass function, $n(M)$, is given by the proportion of Lagrangian space that collapses into objects of mass M . Identifying the proportion of Lagrangian space with $\delta_R > \delta_c$ with the probability that a given sphere exceeds the thresholds, allows us to write

$$n(M) = -\frac{1}{V_M} \frac{dP(\delta_M > \delta_c)}{dM} = -\frac{\bar{\rho}}{M} \frac{dP}{dv} \frac{dv}{dM}. \quad (2.187)$$

In the Press-Schechter model, we obtain

$$n(M) = \sqrt{\frac{v}{2\pi}} \frac{\bar{\rho}}{M^2} \exp\left(-\frac{v}{2}\right) \frac{d \ln v}{d \ln M}. \quad (2.188)$$

This model works surprisingly well, but in the 1990s it was discovered that it failed to accurately predict the number of small halos in N-body simulations [102, 103]. One of the reasons was that collapse is intrinsically triaxial [104, 105]. Based on triaxial collapse dynamics, Sheth & Tormen devised an improved mass function [103, 106]

$$n(M) = A \sqrt{\frac{qv}{2\pi}} \left(1 + \frac{1}{(qv)^p}\right) \frac{\bar{\rho}}{M^2} \exp\left(-\frac{qv}{2}\right) \frac{d \ln v}{d \ln M} \quad (2.189)$$

where the parameters $q = 0.707$, $p = 0.3$ are obtained from simulations and the normalisation A ensures that all mass in the Universe is contained in halos.

The two mentioned mass functions belong to a broader class of universal mass functions parametrised by

$$n(M) = f(v) \frac{\bar{\rho}}{M^2} \frac{d \ln v}{d \ln M}. \quad (2.190)$$

They are universal in the sense that they only depend on the peak height and have no direct cosmology or redshift dependence. They are accurate to $\sim 10 - 20\%$ and for better accuracy one needs to introduce an explicit redshift and/or cosmology dependence [107].

2.4.4 Biasing in the halo model

The mass-function offers a way to compute local bias parameters in Lagrangian space. They are the response coefficients of the mass function to long-wavelength perturbations, δ_L ,

$$n(\mathbf{q}, M) = \bar{n}(M) + \sum_i \frac{\partial^i \bar{n}}{\partial \delta_L^i} (\delta_L^i(\mathbf{q}) - \langle \delta_L^i \rangle). \quad (2.191)$$

The second term in the brackets ensures that we maintain the correct average density. A long-wavelength perturbation increases the local mean density and thus lowers the critical density like

$$\tilde{\delta}_c \simeq \delta_c - \delta_L. \quad (2.192)$$

The tracer overdensity is given by

$$\delta(\mathbf{q}) = \frac{n(\mathbf{q}, M)}{\bar{n}(M)} - 1, \quad (2.193)$$

which allows us to compute the local bias parameters in Lagrangian space as

$$b_i^{(L)}(M) = \frac{1}{\bar{n}} \frac{\partial^i \bar{n}}{\partial \delta_L^i} = -\frac{1}{\bar{n}} \frac{\partial^i \bar{n}}{\partial \delta_c^i}. \quad (2.194)$$

For the Sheth-Tormen mass function (2.189), the first two Lagrangian bias parameters are given by

$$b_1^{(L)}(M) = \frac{2p}{\delta_c [(vq)^p + 1]} + \frac{vq - 1}{\delta_c}, \quad (2.195)$$

$$b_2^{(L)}(M) = \frac{2(2p^2 + 2vpq - p)}{\delta_c^2 [(vq)^p + 1]} + \frac{vq(vq - 3)}{\delta_c^2}. \quad (2.196)$$

We show their mass dependence in Fig. 2.8. The bias parameters for the Press-Schechter mass function correspond to $q = 1, p = 0$.

2.4.5 Scale-dependent bias from PNGs

The presence of primordial non-Gaussianities violates the distributional assumption of the initial conditions used in the previous sections. In this thesis, we only work with weak,

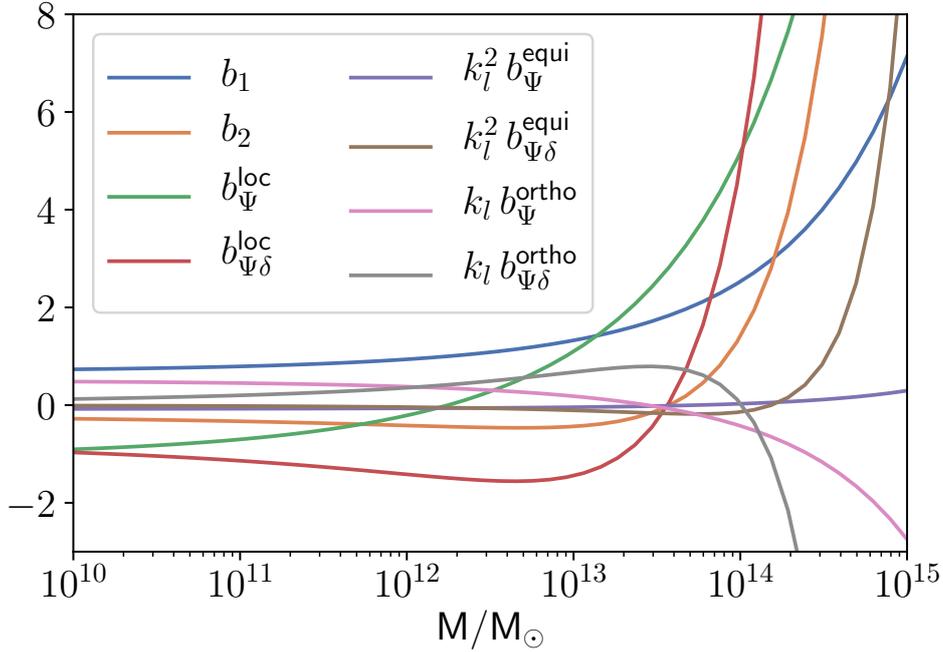


Fig. 2.8 Lagrangian bias parameters from the Sheth-Tormen mass function. In order to have dimensionless quantities, we scale where necessary with $k_l = 0.01 \text{ hMpc}^{-1}$. All parameters are roughly constant and $O(1)$ for tracers of masses $\leq 10^{14} M_\odot$ and grow rapidly afterwards.

quadratic PNGs that can be incorporated into the formalism by a scale-dependent rescaling of the variance. This gives rise to new bias terms and, since we assume weak non-Gaussianities, we limit ourselves to the leading order terms in f_{NL} . In the case of more general PNGs, the deviations from Gaussianity can be captured by means of an Edgeworth expansion and a rescaling of higher moments. We refer the interested reader to [108].

We now derive the leading order additional, scale-dependent bias terms caused by weak, quadratic PNGs. Scale-dependent bias terms were first observed for the local model in [91]. Our discussion here largely follows [109, 110]. We start by writing the gravitational potential, Φ , in terms of an auxiliary Gaussian potential, ϕ ,

$$\Phi(\mathbf{k}) = \phi(\mathbf{k}) + f_{\text{NL}} \int \frac{d^3 p_1}{(2\pi)^3} \frac{d^3 p_2}{(2\pi)^3} K_{\text{NL}}(\mathbf{p}_1, \mathbf{p}_2) \phi(\mathbf{p}_1) \phi(\mathbf{p}_2) \delta_D(\mathbf{k} - \mathbf{p}_{1:2}). \quad (2.197)$$

The kernel K_{NL} incorporates the specifics of the considered (quadratic) PNG model and the amplitude f_{NL} ensures the PNGs are indeed weak. We then use the peak-background split (2.182) and separate the Gaussian potential into short and long wavelength components, ϕ_L and ϕ_S . The associated (Gaussian) density contrasts are δ_L and δ_S . Applying the Poisson

equation to the short scales of the non-Gaussian potential yields

$$\delta_S^{(\text{NG})}(\mathbf{k}) = \delta_S(\mathbf{k}) + 2f_{\text{NL}} \int \frac{d^3p}{(2\pi)^3} K_{\text{NL}}(\mathbf{p}, \mathbf{k} - \mathbf{p}) [\delta_L(\mathbf{p}) \phi_S(\mathbf{k} - \mathbf{p}) + \delta_S(\mathbf{p}) \phi_L(\mathbf{k} - \mathbf{p})]. \quad (2.198)$$

Due to the non-Gaussianities, long and short modes are coupled. This translates to a modulation of the non-Gaussian variance on small scales by the long (Gaussian) mode

$$\sigma_{S,\text{NG}}^2(\mathbf{k}) = \langle \delta_S^{(\text{NG})} \delta_S^{(\text{NG})} \rangle = \sigma_{R,0}^2 + 4f_{\text{NL}} \left[\phi_L(\mathbf{k}) \sigma_{R,\text{NL}}^2(k) + \delta_L(\mathbf{k}) \sigma_{R,\text{NL}-\phi}^2(k) \right]. \quad (2.199)$$

Here $\sigma_{R,0}^2$ is the variance of the (short scale) Gaussian field and we use the definitions:

$$\sigma_{R,\text{NL}}^2(k) = \int \frac{d^3k_s}{(2\pi)^3} K_{\text{NL}}(k, k_s) P_L(k_s) W^2(k_s R), \quad (2.200)$$

$$\sigma_{R,\text{NL}-\phi}^2(k) = \int \frac{d^3k_s}{(2\pi)^3} K_{\text{NL}}(k, k_s) \mathcal{M}^{-1}(k_s) P_L(k_s) W^2(k_s R). \quad (2.201)$$

The window functions come from the peak-background split (2.179). The $4f_{\text{NL}} \sigma_{R,\text{NL}-\phi}^2 \delta_L$ term can be ignored since it only introduces a very small ($\leq 10^{-4} f_{\text{NL}}$) scale-independent correction to the bias terms [110].

To be able to obtain analytical results and since the scale-dependent bias is most relevant on large scales,⁵ we approximate K_{NL} as

$$K_{\text{NL}}(\mathbf{k}, \mathbf{p}) \simeq A \left(\frac{k}{p} \right)^\alpha \quad (2.202)$$

with (A, α) obtained in the squeezed limit as (1,0), (3,1), (-3,2) for local, equilateral and orthogonal PNGs respectively and we approximated $n_s = 1$ [97]. This leads to the following expression for the non-linear variance (2.200)

$$\sigma_{R,\alpha}^2(k) = \frac{1}{2\pi^2} \int dp p^2 \left(\frac{k}{p} \right)^\alpha P(p) W^2(pR). \quad (2.203)$$

Due to the scale-dependent rescaling of the variance, the bias expansion of the mass function (2.191), has to be generalised. For universal mass functions, the variance enters via the

⁵This is known from the local model, where no approximation is needed.

peak-height, $v = \delta_c / \sigma_R$, which transforms like

$$v(\phi_L) = v(\phi_L = 0) / \left((1 + 4f_{\text{NL}}\phi_L \frac{\sigma_{R,\alpha}^2}{\sigma_{R,0}^2}) \right). \quad (2.204)$$

From this, we derive the logarithmic derivative with respect to the mass scale,

$$\frac{d \ln v(\phi_L)}{d \ln M} = \frac{d \ln v(\phi_L = 0)}{d \ln M} \left[1 + 4f_{\text{NL}}\phi_L \frac{\sigma_{R,\alpha}^2}{\sigma_{R,0}^2} \left(\frac{d \ln \sigma_{R,\alpha}^2}{d \ln \sigma_{R,0}^2} - 1 \right) \right]. \quad (2.205)$$

To second order in the potentials and first order in f_{NL} , we find two new bias parameters. For universal mass functions the first order term is given by [97, 110–112]

$$b_{\Psi}^{(L)} = \left. \frac{\partial \ln n(M)}{\partial \phi_L} \right|_{\phi_L=0} = 2A f_{\text{NL}} \frac{\sigma_{R,\alpha}^2}{\sigma_{R,0}^2} \left[b_1 \delta_c + 2 \left(\frac{\partial \ln \sigma_{R,\alpha}^2}{\partial \ln \sigma_{R,0}^2} - 1 \right) \right]. \quad (2.206)$$

Similarly, one finds at second order [94, 112]

$$b_{\Psi\delta}^{(L)} = \frac{1}{n(M)} \left. \frac{\partial^2 \bar{n}_h}{\partial \delta_L \partial \phi_L} \right|_{\phi_L, \delta_L=0} = A f_{\text{NL}} \left[2\delta_c b_2^{(L)} + b_1^{(L)} \left(4 \frac{d \ln \sigma_{R,\alpha}^2}{d \ln \sigma_{R,0}^2} - 6 \right) \right] \frac{\sigma_{R,\alpha}^2}{\sigma_{R,0}^2}. \quad (2.207)$$

In practice, we implement the derivatives using the chain rule and the monotonic relation between the variance and the smoothing scale R . The scale dependence enters the bias parameters through K_{NL} and the Poisson factor when converting the potential Ψ to the (observed) density contrast δ . On large scales, the overall scale dependence behaves like $k^{\alpha-2}$. In Fig. 2.8 we show the mass dependence of those bias terms.

2.5 Counts-in-cells

In this section, we introduce the counts-in-cells formalism that predicts the matter distribution in spheres at late times. Since some of the derivations are rather technical, we summarise the main results here:

- Combining spherical collapse and large deviation theory, one can derive a very accurate probability density function (PDF) for the matter density in concentric spheres at late times. It is given in (2.223) together with (2.225).

- This result can be generalised to the case of distant spheres (2.255). The joint PDF is expressed in terms of the PDF for concentric spheres, the two-point correlation function and the density dependent bias (2.254).

Throughout this section, we assume $\langle \rho \rangle = 1$ and work with filtered fields using a Top-Hat filter (2.181). We introduced filtering in section 2.4.2.

2.5.1 Counts-in-cells statistics for concentric spheres

In the initial conditions, the PDF of densities, ρ_i , in concentric spheres of radii R_i is a multivariate Gaussian distribution with mean one and covariance

$$\Sigma(R_i, R_j) = \Sigma_{ij} = \int \frac{dk k^2}{2\pi^2} P_L(k, z_{\text{ini}}) W(kR_i) W(-kR_j). \quad (2.208)$$

The linearly evolved density field, $\delta^{(1)}$, also follows a mean zero multivariate Gaussian distribution with the covariance evolved accordingly. It is easier to work with the linearly evolved density fields since the linear power spectrum at late times is provided by Boltzmann codes (see section 2.2.5) and one does not explicitly need to define the redshift of the initial conditions.

Beside the PDF, there are (at least) two more ways to characterise probability distributions: the moment-generating function, M , and the cumulant-generating function, ϕ . They are related to the PDF via a Laplace transformation

$$M(\lambda_1, \dots, \lambda_n) = e^{\phi(\lambda_1, \dots, \lambda_n)} = \int \prod_{i=1}^n d\rho_i P(\rho_1, \dots, \rho_n) e^{\rho_i \lambda_i}. \quad (2.209)$$

Here we used $(\lambda_1, \dots, \lambda_n) = \{\lambda_i\}$ to shorten the notation. Hence, one can recover the PDF from both the moment-generating function, M , and the cumulant-generating function, ϕ via an inverse Laplace transform of (2.209), e.g.

$$P(\{\rho_i\}) = \prod_i \left(\int_{-i\infty}^{i\infty} \frac{d\lambda_i}{2\pi i} \right) \exp[-\rho_i \lambda_i + \phi(\{\lambda_i\})]. \quad (2.210)$$

For a multivariate Gaussian, we find for the moment-generating function and cumulant-generating function

$$M(\{\lambda_i\}) = \exp \left[\langle \rho_i \rangle \lambda_i + \frac{1}{2} \lambda_i \Sigma_{ij} \lambda_j \right], \quad \phi(\{\lambda_i\}) = \langle \rho_i \rangle \lambda_i + \frac{1}{2} \lambda_i \Sigma_{ij} \lambda_j. \quad (2.211)$$

The non-linear evolution of the Universe changes the functional form of the PDF. Traditionally, the late-time PDF was approximated by log-normal models [113] whose PDF for the density in a single sphere is

$$P(\rho) = \frac{1}{\rho \sigma_\mu \sqrt{2\pi}} \exp\left(-\frac{(\log(\rho) - \bar{\mu})^2}{2\sigma_\mu^2}\right). \quad (2.212)$$

Requiring a unit mean fixes $\bar{\mu} = -\sigma_\mu^2/2$ and the non-linear variance σ_μ^2 remains as a free parameter that can be fitted against the observations. The idea behind the log-normal model is that a log-transformation reduces the skewness (which is zero in the Gaussian case) of a random variable and extends its range from positive values to positive and negative values.

It is possible to obtain significantly better predictions using the spherical collapse model and large deviation theory. The following discussion is based on [114, 7]. Large deviation theory describes the decay of a PDF in its tails by studying its behavior as some natural parameter goes to zero. In our setting, this natural parameter is the variance at some length scale, R and redshift, z . Importantly, if the variance on one length scale and redshift goes to zero, it implies that the variance at all length scales and redshifts goes to zero. In this limit, the rate function I , gives the leading order behavior of the PDF in its tails

$$I(\{\delta_i^{(1)}\}) = -\lim_{\sigma^2(R) \rightarrow 0} \sigma^2(R) \log P(\{\delta_i^{(1)}\} | \Sigma_{ij}) \implies P(\{\delta_i^{(1)}\}) \sim e^{-I(\{\delta_i^{(1)}\})/\sigma^2(R)}. \quad (2.213)$$

Here we defined $\Sigma(R, R) = \sigma^2(R)$. Following the argument from above, we are free to choose the reference scale (and redshift) which use the reference variance. For concreteness, let us choose $R = \max_i \{R_i\}$. One can compute the rate function of a set of concentric spheres of (linear) density contrasts, $\delta_i^{(1)}$, to

$$I(\{\delta_i^{(1)}\}) = \frac{\sigma^2(R)}{2} \delta_i^{(1)} \Sigma^{-1}(\{R_k\}, \{R_k\})_{ij} \delta_j^{(1)}, \quad (2.214)$$

where the covariance is given by (2.208). The rate function of linear densities is a simple quadratic function. This changes as we move to non-linear densities. The contraction principle allows us to calculate the rate function, ψ , of another random variable, ρ , that is related via a continuous transformation,

$$f : \delta_i^{(1)}(\mathbf{q}) \rightarrow \rho_{R_i}(\mathbf{x}), \quad (2.215)$$

via

$$\psi(\{\rho_i\}) = \inf_{\{\delta_i^{(1)}\}: f(\delta_i^{(1)}) = \rho_i} I(\{\delta_i^{(1)}\}). \quad (2.216)$$

In case the transformation f is not injective, only the least unlikely path is relevant. All other contributions are exponentially suppressed by the (infinitesimally) small variance in (2.213). This is particularly relevant for cosmological mappings where one could imagine many initial conditions resulting in the same final configuration. Beside physical mappings, the contraction principle can also be applied to variable transformations. We will make use of this too during the derivation of the late-time density in spheres PDF.

Let us start with the mapping from early to late times, or maybe more accurately, from linear to non-linear densities. One can argue that the most likely mapping from the initial conditions to a late-time spherical overdensity should respect the high degree of symmetry in the final configuration and is hence given by the spherical collapse [115]. This makes the problem tractable because we can identify, via mass conservation, the spheres in the initial conditions that are mapped on a given set of concentric spheres at late times. The most likely origin of each sphere at late times with radius R_i and mass $(\frac{4}{3}\pi R_i^3 \rho_i)$ is a sphere around the same co-moving coordinate with the same mass and hence radius $R_{\text{ini},i} = \rho_i^{1/3} R_i$. Using then the spherical collapse solution (2.178) as continuous mapping (2.215) allows us to use the contraction principle (2.216). We find for the decay rate function at late times, ψ

$$\psi(\{\rho_k\}) = \frac{\sigma^2(R)}{2} \delta_i^{(1)}(\{\rho_k\}) \Sigma^{-1}(\{R_k \rho_k^{(1/3)}\}, \{R_k \rho_k^{(1/3)}\})_{ij} \delta_j^{(1)}(\{\rho_k\}). \quad (2.217)$$

Using a path integral approach, the result was first obtained in [101]. Under some technical conditions, one can recover the scaled cumulant-generating function, φ , via Varadhan's lemma from the rate function [116, 117, 114]

$$\varphi(\{\lambda_i\}) = \lim_{\sigma^2 \rightarrow 0} \sigma^2 \phi(\{\lambda_k / \sigma^2\}) = \sup_{\{\delta_i\}} (\lambda_i \rho_{\text{SC}}(\delta_i) - I(\{\delta_i\})). \quad (2.218)$$

Inverting the relation (2.178) allows us to take the supremum over the non-linear density ρ instead. Moreover, if the rate function is convex, the last expression simplifies to a Legendre transform with stationarity conditions,

$$\rho_i = \frac{\partial}{\partial \lambda_i} \varphi(\{\lambda_k\}). \quad (2.219)$$

The idea is now to extend the scaled cumulant-generating function, which describes the zero-variance limit, to finite variances via

$$\phi_R(\{\lambda_k\}) = \frac{1}{\sigma_{\text{NL}}^2(R)} \varphi(\{\lambda_k \sigma_{\text{NL}}^2(R)\}). \quad (2.220)$$

It is expected that this approach is valid as long as the variance remains small, i.e. for large enough spheres. Using the inverse Laplace transform of the cumulant-generating function (2.210) one can then obtain an (approximate) matter-density-in-spheres PDF at late times. However, the inverse Laplace transform is numerically challenging. One way to solve the integral (2.210) analytically is the saddle-point approximation. It was shown in [7] that one can approximate very accurately the inverse Laplace transform with the saddle-point approximation in the log-densities.⁶ Applying the contraction principle (2.216) to the log-transform simply amounts to substituting $\rho = \exp(\mu)$ in (2.222). In addition, one then extends the scaled cumulant-generating function to finite values of the variance of the log-density (2.220). It is worth pointing out that the log-transformation merely improves the saddle-point approximation but does not change the integral per se. Applying the saddle-point approximation, one finds for the PDF of log-densities in n concentric spheres,

$$P_\mu(\{\mu_i\}) \simeq \frac{1}{(2\pi)^{n/2}} \sqrt{\det \left[\frac{\partial^2 \Psi_R}{\partial \mu_i \partial \mu_j} \right]} \exp[-\Psi_R(\{\mu_i\})]. \quad (2.221)$$

Here, we used that the maximum of the exponent in the saddle-point approximation corresponds to the Legendre transform of the cumulant-generating function⁷ and introduced the decay rate function, Ψ_R

$$\Psi_R(\{\rho_k\}) \equiv \frac{\psi(\{\rho_k\})}{\sigma_{\text{NL}}^2(R)} = \frac{\sigma^2(R)}{2\sigma_{\text{NL}}^2(R)} \delta_i^{(1)}(\{\rho_k\}) \Sigma^{-1}(\{R_k \rho_k^{(1/3)}\}, \{R_k \rho_k^{(1/3)}\})_{ij} \delta_j^{(1)}(\{\rho_k\}). \quad (2.222)$$

The resulting PDF for the density is obtained via conservation of probability, $P(\rho)d\rho = P_\mu(\log \rho)d \log \rho$. Thus, the density in a single sphere is given by

$$P(\rho) = \sqrt{\frac{\Psi_R''(\rho) + \Psi_R'(\rho)/\rho}{2\pi}} \exp[-\Psi_R(\rho)] \quad (2.223)$$

where primes denote derivatives with respect to the density and we used

$$\frac{\partial^2 \Psi_R(\rho)}{\partial (\log \rho)^2} = \rho^2 \Psi_R''(\rho) + \rho \Psi_R'(\rho). \quad (2.224)$$

⁶We are log-transforming the filtered densities. Applying the log-transform before filtering would mean we are modelling the geometric mean density in spheres.

⁷This is the inverse transform of (2.218).

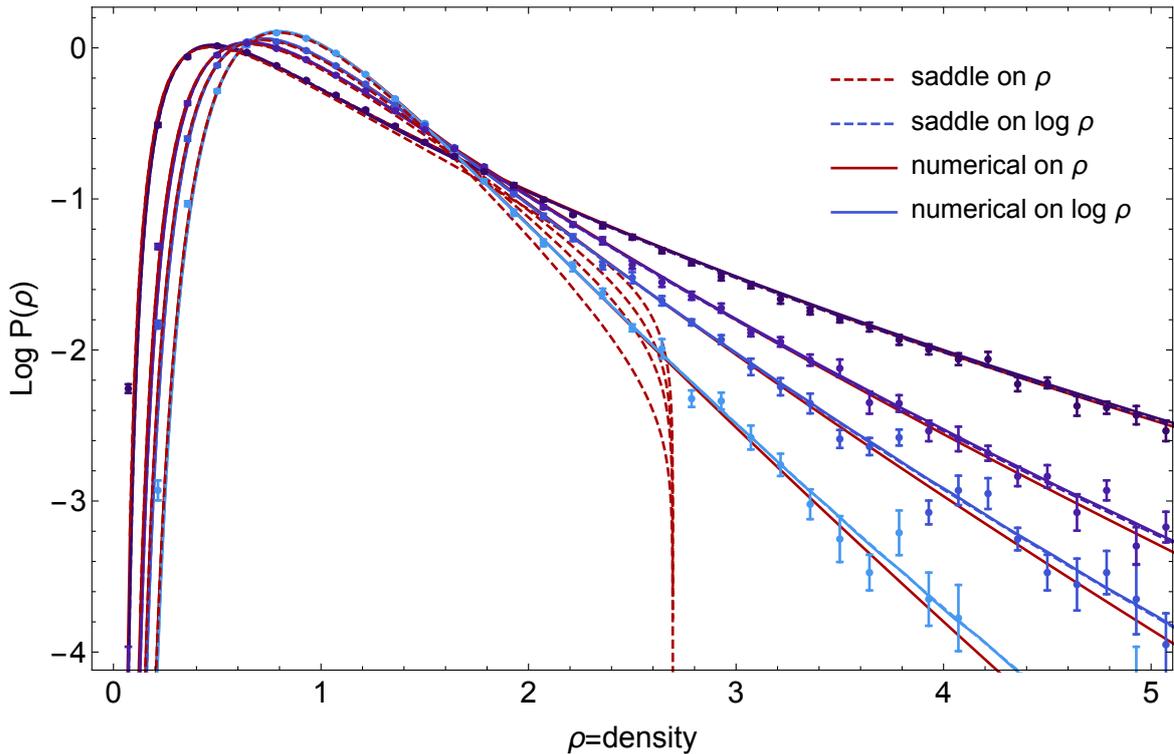


Fig. 2.9 Probability density function for the density in spheres of radius $10 h^{-1} \text{Mpc}$ at redshifts $z = 1.36, 0.97, 0.65$ and 0 (from light to dark blue). The points are measurements in an N-Body simulation and one sees that the skewness increases as the redshift decreases. The saddle-point approximation with the density breaks down for $\rho \geq 2$, because the rate function ceases to be convex. In contrast, the saddle-point approximation of the log-densities approximates the exact numerical inverse Laplace transform well and matches the measurements. Figure from [7].

Lastly, one has to rescale the PDF

$$\mathcal{P}(\rho) = P \left(\rho \frac{\langle \rho \rangle_{\text{SPA}}}{\langle 1 \rangle_{\text{SPA}}} \right) \frac{\langle \rho \rangle_{\text{SPA}}}{\langle 1 \rangle_{\text{SPA}}^2} \quad (2.225)$$

since the saddle-point approximation neither ensures that the approximated PDF is (perfectly) normalised nor that it has unit mean. The average with respect to the saddle-point approximation ‘PDF’ is defined as

$$\langle f(\rho) \rangle_{\text{SPA}} = \int_0^\infty P(\rho) f(\rho) d\rho. \quad (2.226)$$

Like the log-normal model, the counts-in-cells PDF has a non-linear variance at some length scale as a free parameter. [7] showed that the model for the PDF of the real-space matter density field $\mathcal{P}(\rho)$ for a single sphere of radius R with the (non-linear) variance of the log-density $\sigma_\mu^2(R)$ as a driving parameter is accurate at the percent level for standard deviations $\sigma_\mu \lesssim 0.5$, greatly improving over the accuracy that can be obtained from log-normal models [113]. We illustrate the accuracy of the saddle-point approximation in Fig. 2.9 by comparing the saddle-point predictions with both measurements and the predictions from the numerical integration for a wide range of densities and redshifts. In addition, the Figure shows the necessity to perform the saddle-point approximation in the log-densities rather than the densities.

2.5.2 Counts-in-Cells statistics in the large separation limit

This section was heavily influenced by [8, 118] and we refer the reader to those references for more details.

Summary

The joint distribution of the densities in two spheres of radii R, R' at distance r can be written as the product of their marginal distributions and a correlation function, ξ_\circ , that is implicitly defined through

$$\mathcal{P}(\rho, \rho'; r) = \mathcal{P}(\rho) \mathcal{P}(\rho') (1 + \xi_\circ(r, \rho, \rho')). \quad (2.227)$$

In the large separation limit ($r \gg R, R'$), the correlation function separates [8],

$$\xi_\circ(r, \rho, \rho') \sim \xi_\circ(r) b_\circ(\rho) b_\circ(\rho'), \quad (2.228)$$

into a distance-dependent correlation function,

$$\xi_{\circ}(r) = \langle \rho \rho'; r \rangle_c, \quad (2.229)$$

and a modulation of the average correlation, called density-dependent bias or called sphere bias, $b_{\circ}(\rho)$, that satisfies (see: 2.252 and 2.253)

$$\langle b_{\circ} \rangle = 0, \quad \langle \rho b_{\circ} \rangle = 1. \quad (2.230)$$

Using these properties, the density-dependent bias can be measured and computed as conditional expectation since the conditional distribution is

$$\mathcal{P}(\rho|\rho'; r) = \frac{\mathcal{P}(\rho, \rho'; r)}{\mathcal{P}(\rho')} \simeq \mathcal{P}(\rho)(1 + \xi_{\circ}(r)b_{\circ}(\rho)b_{\circ}(\rho')). \quad (2.231)$$

Computing the conditional expectation then yields

$$\langle \rho|\rho'; r \rangle = 1 + b_{\circ}(\rho')\xi_{\circ}(r) \implies b_{\circ}(\rho') = \frac{\langle \rho|\rho'; r \rangle - 1}{\xi_{\circ}(r)}. \quad (2.232)$$

Physically speaking, the density-dependent bias quantifies how much the density of a sphere influences the expected density of a distant sphere.

Gaussian bias

For linear overdensities, we are able to evaluate (2.232). The conditional distribution of a multivariate Gaussian (X, Y) is given by

$$X|Y = y \sim \mathcal{N}(\mu_X + \Sigma_{XY}\Sigma_{YY}^{-1}(y - \mu_Y), \Sigma_{XX} - \Sigma_{XY}\Sigma_{YY}^{-1}\Sigma_{YX}). \quad (2.233)$$

Assuming two spheres of radii (R_i, R_j) sitting at $(\mathbf{r}_i, \mathbf{r}_j)$ yields the covariance,

$$\Sigma_{ij} = \int \frac{d^3k}{(2\pi)^3} P_L(k) W(R_i k) W(-R_j k) \exp[i\mathbf{k} \cdot (\mathbf{r}_i - \mathbf{r}_j)]. \quad (2.234)$$

The phase factor arises from translating the window function in real space since

$$W_{\mathbf{r}}(\mathbf{k}) = \int d^3x e^{i\mathbf{k}\mathbf{x}} W(\mathbf{x} - \mathbf{r}) = e^{i\mathbf{k}\mathbf{r}} \int d^3x e^{i\mathbf{k}\mathbf{x}} W(\mathbf{x}) = e^{i\mathbf{k}\mathbf{r}} W_0(\mathbf{k}). \quad (2.235)$$

Given $\mu_X = \mu_Y = \langle \rho \rangle = 1$ we find

$$b_{\circ}(\rho) = \frac{\langle \rho' | \rho; r \rangle - 1}{\Sigma_{12}} = \frac{\rho - 1}{\Sigma_{22}} = 1_i \frac{\partial I / \sigma^2(R)}{\partial \delta_i} \quad (2.236)$$

for the so-called Kaiser bias [119]. This agrees with the intuition that an overdense sphere makes it more likely that nearby spheres are overdense too. Moreover, the strength of the effect depends on the significance of the deviation from the mean. In fact, the sphere bias can be expressed in terms of the first derivative of the decay-rate function I / σ^2 (2.213). This offers another intuition: the sphere bias quantifies how sensitive the probability of a configuration is to a small change in the density (caused by a second sphere, for instance).

non-Gaussian PDF and late-time bias

Now, we are going to derive the joint PDF of two sets of concentric spheres via the cumulant-generating function. For two sets of distant ($r \gg \max(\{R_i\}, \{R_j\})$) concentric spheres ($\{(\rho_i, R_i)\}, \{(\rho'_j, R'_j)\}$) it is given by

$$\begin{aligned} \phi(\{\lambda_i\}, \{\lambda_j\}; r) &= \sum_{\{p_i\}, \{q_j\}} \langle \rho_1^{p_1} \cdots \rho_n^{p_n} \rho'_{n+1}{}^{q_1} \cdots \rho'_{n+m}{}^{q_m} \rangle_c \prod_{i,j} \frac{\lambda_i^{p_i} \lambda_j^{q_j}}{p_i! q_j!} \\ &\simeq \sum_{\{p_i\}} \langle \rho_1^{p_1} \cdots \rho_n^{p_n} \rangle_c \prod_i \frac{\lambda_i^{p_i}}{p_i!} + \sum_{\{q_j\}} \langle \rho'_{n+1}{}^{q_1} \cdots \rho'_{n+m}{}^{q_m} \rangle_c \prod_j \frac{\lambda_j^{q_j}}{q_j!} \\ &+ \frac{1}{\xi_{\circ}(r)} \sum_{\substack{\{p_i\}, \{q_j\} \text{ s.t.} \\ p_b > 0 \& q_a > 0}} \langle \rho'_{n+a}{}^{p_1} \cdots \rho_b^{p_b-1} \cdots \rho_n^{p_n} \rangle_c \langle \rho_b \rho'_{n+1}{}^{q_1} \cdots \rho'_{n+a}{}^{q_a-1} \cdots \rho'_{n+m}{}^{q_m} \rangle_c \prod_{i,j} \frac{\lambda_i^{p_i} \lambda_j^{q_j}}{p_i! q_j!}. \end{aligned} \quad (2.237)$$

The reason for using the cumulant-generating function is that, in the large separation limit, it has a perturbative expansion that contains, up to first non-trivial order, all cumulants with densities from one set of the concentric spheres and interaction cumulants that have exactly one ‘external’ density [120]. From now on, we use the convention that $\delta^{(1)}$ refers to (linear) over densities, while ρ refers to the non-linear density in spheres. Fig. 2.10 gives an overview of the strategy to compute the cumulant-generating function: We have already computed the cumulants for the PDF of n concentric spheres (setting a)). This allows us to extract the cumulant-generating function for n concentric spheres and one distant sphere (setting b)). Using the above expression, we are then able to compute the cumulant-generating function for two distant sets of concentric spheres (setting c)).

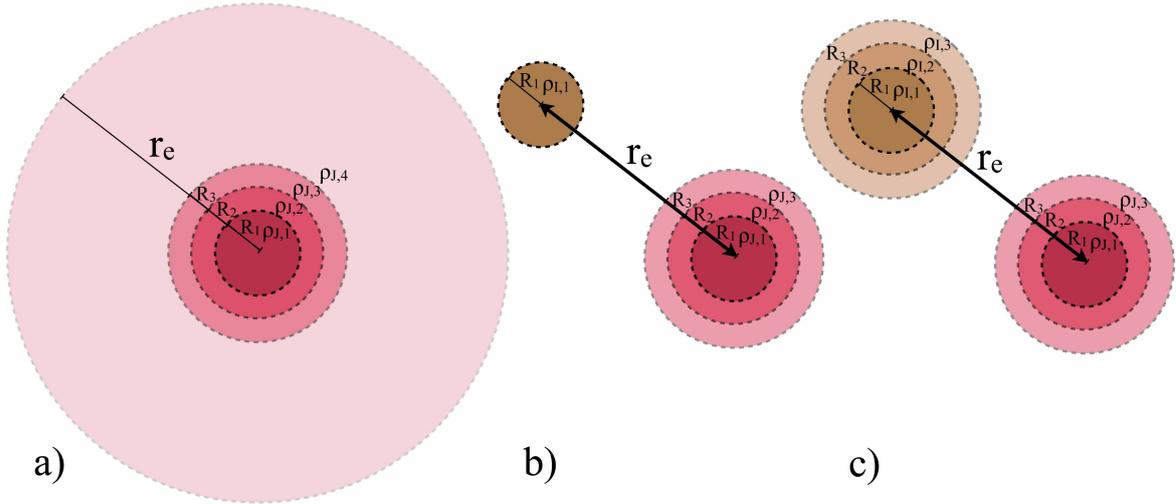


Fig. 2.10 The conceptual steps towards a joint PDF for distant concentric spheres. Starting from the cumulant-generating function for $n + 1$ spheres, a), we derive the cumulant-generating function for n concentric spheres and one distant sphere, b). Using perturbation theory, one can then extend this to the general case with two distant sets of concentric spheres. Figure from [8].

In order to compute the cumulant-generating function with one external density, we start by considering $n + 1$ concentric spheres where the $n + 1$ th sphere has radius $r \gg \max_{i \leq n} R_i$. This cumulant generating function is given by

$$\phi_b(\{\lambda_i\}; < r) = \sum_{p_i=0}^{\infty} \langle \rho_{n+1} \prod_{i=1}^n \rho^{p_i}(R_i) \rangle_c \prod_{i=1}^n \frac{\lambda_i^{p_i}}{p_i!}. \quad (2.238)$$

The one external density can also be captured by a derivative of the general $n + 1$ density cumulant-generating function,

$$\phi_b(\{\lambda_i\}; < r) = \left. \frac{\partial}{\partial \lambda_{n+1}} \phi(\lambda_1, \dots, \lambda_n, \lambda_{n+1}) \right|_{\lambda_{n+1}=0} = \rho_{n+1}(\lambda_1, \dots, \lambda_n, 0) = \rho_{\text{SC}}(\delta_{n+1}^{(1)}), \quad (2.239)$$

where we used the stationary condition (2.219) for the second equality and in the last step the spherical collapse mapping. The expected value of the $(n + 1)$ th density conditional on knowing the first n can be computed by choosing $\delta_{n+1}^{(1)}$ such that

$$\phi_n(\lambda_1, \dots, \lambda_n) = \phi_{n+1}(\lambda_1, \dots, \lambda_{n+1} = 0). \quad (2.240)$$

The idea is that the conditional expectation of the $n + 1$ th density, given we know the first n , is such that if the $n + 1$ th density had indeed this value, knowing it would not change the joint distribution as it contains no new information.

After a lengthy but straightforward calculation (see Appendix A of [8]), one obtains a similar result for the conditional mean as in the Gaussian case (2.233):

$$\delta_{n+1}^{(1)} = \Sigma_{n+1,1:n}(r\rho^{1/3}, \{R_i\rho_i^{1/3}\})(\Sigma_{1:n,1:n})^{-1}\delta_{1:n}^{(1)}. \quad (2.241)$$

The notations $i:j$ means we select the indices from i through j (both included) and interpret the resulting object as vector or matrix. In contrast to the Gaussian case, the covariance matrices now depend on the densities $\rho_{1:n}$. In the large-separation limit, the influence of the inner n concentric spheres on the density within a much larger radius will be small, so $\delta_{n+1}^{(1)} \approx \langle \delta_{n+1}^{(1)} \rangle = 0$. We can thus Taylor expand (2.239)

$$\begin{aligned} \phi_b(\{\lambda_i\}, < r) &= \rho_{\text{sc}}(\delta_{n+1}^{(1)}) \simeq 1 + \delta_{n+1}^{(1)} \\ &= 1 + \Sigma_{n+1,1:n}(r\rho^{1/3}, \{R_i\rho_i^{1/3}\})\Sigma_{1:n,1:n}^{-1}\delta_{1:n}^{(1)} \\ &= 1 + \sigma^2(< r)1_{1:n}(\Sigma_{1:n,1:n})^{-1}\delta_{1:n}^{(1)}. \end{aligned} \quad (2.242)$$

where $1_{1:n}$ is the vector of ones in every entry and we used that for $r \gg R_i$, the window function describing the smaller sphere appears constant in (2.234), so

$$\Sigma(r\rho^{1/3}, R_i\rho_i^{1/3}) \simeq \Sigma(r, 0) \equiv \sigma^2(< r) \quad \text{for } 1 \leq i \leq n. \quad (2.243)$$

Identifying the cumulant-generating function with the spherical collapse solution, we see

$$\phi_b(\{\lambda_i\}; r) = \frac{1}{r^2} \frac{d}{dr} \left(\frac{r^3}{3} \phi_b(\{\lambda_i\}, < r) \right). \quad (2.244)$$

This allows us to write

$$\phi_b(\{\lambda_i\}; r) = 1 + \xi_o(r)b_\phi(\{\lambda_i\}), \quad (2.245)$$

where we used

$$\xi_o(r) = \frac{1}{r^2} \frac{d}{dr} \left(\frac{r^3}{3} \sigma^2(< r) \right) \quad (2.246)$$

and defined the bias generating function, b_ϕ ,

$$b_\phi(\{\lambda_i\}) = 1_{1:n}(\Sigma_{1:n,1:n})^{-1}\delta_{1:n}^{(1)}. \quad (2.247)$$

Inserting this into the general cumulant-generating function (2.237) yields

$$\begin{aligned}\phi(\{\lambda_i\}, \{\lambda_j\}; r) &= \phi(\{\lambda_i\}) + \phi(\{\lambda_j\}) + \frac{1}{\xi_\circ(r)} (\phi_b(\{\lambda_i\}; r) - 1)(\phi_b(\{\lambda_j\}; r) - 1) \\ &= \phi(\{\lambda_i\}) + \phi(\{\lambda_j\}) + \xi_\circ(r) b_\phi(\{\lambda_i\}) b_\phi(\{\lambda_j\}).\end{aligned}\quad (2.248)$$

This equation reveals that the joint PDF takes indeed the form assumed above

$$\mathcal{P}(\{\rho\}, \{\rho'\}; r) = \mathcal{P}(\{\rho\}) \mathcal{P}(\{\rho'\}) (1 + \xi_\circ(r) b_\circ(\{\rho\}) b_\circ(\{\rho'\})) \quad (2.249)$$

where the terms are given by

$$\mathcal{P}(\{\rho\}) = \int \prod_j \frac{d\lambda_j}{2\pi i} \exp[\lambda_j \rho_j + \phi(\{\lambda_j\})] \quad (2.250)$$

and

$$\mathcal{P}(\{\rho\}) b_\circ(\{\rho\}) = \int \prod_j \frac{d\lambda_j}{2\pi i} b_\phi(\{\lambda_k\}) \exp[\lambda_j \rho_j + \phi(\{\lambda_j\})]. \quad (2.251)$$

From (2.247) it follows that $b_\phi(0) = 0$ and $b'_\phi(0) = 1$ which allows us to compute the following two correlators

$$\langle b_\circ(\{\rho\}) \rangle = \int_0^\infty \prod_{i=1}^n d\rho_i P(\{\rho_i\}) b_\circ(\{\rho_i\}) = b_\phi(\{\lambda_i\}) e^{\phi(\{\lambda_i\})} \Big|_{\{\lambda_i\}=0} = 0 \quad (2.252)$$

and for all i we find

$$\langle b_\circ(\{\rho_k\}) \rho_i \rangle = \int_0^\infty \prod_{j=1}^n d\rho_j P(\{\rho_j\}) b_\circ(\{\rho_j\}) \rho_i = \frac{\partial}{\partial \lambda_i} b_\phi(\{\lambda_j\}) e^{\phi(\{\lambda_j\})} \Big|_{\{\lambda_j\}=0} = 1. \quad (2.253)$$

Performing a log-transformation of the density field, allows us to accurately compute the bias parameters $b_\circ(\rho)$ using the saddle-point approximation. It results in mapping the result in the initial condition (2.236) to late times via the spherical collapse dynamics [118]

$$b_\circ(\{\rho_k\}) \approx b_\phi(\{\lambda_k = \partial_k \Psi_R(\{\rho_i\})\}) = 1_{1:n} \Sigma^{-1}(\{\rho_k^3 R_k\}, \{\rho_k^3 R_k\})_{1:n, 1:n} \delta_{1:n}^{(1)}(\{\rho_k\}). \quad (2.254a)$$

As with the marginal PDF (2.225), we need to rescale the bias parameter obtained from the saddle-point approximation to match the above relations. For the n -sphere bias we do so via

$$\hat{b}_\circ(\{\rho\}) = \frac{b_\circ(\{\rho\}) - \langle b_\circ(\{\rho\}) \rangle}{\frac{1}{n} \sum_{i=1}^n (\langle \rho_i b_\circ(\{\rho\}) \rangle - \langle b_\circ(\{\rho\}) \rangle)}. \quad (2.254b)$$

Fig. 3.3 shows the predictive power of the density-dependent bias obtained with the saddle-point approximation for a wide range of redshifts. In contrast to the Kaiser bias (2.236) which predicts straight line, the non-Gaussian bias is curved.

The result (2.249) is in fact more general, as one can extend the cumulant expansion (2.237) to n sets of concentric spheres. Then the joint PDF in the large separation limit is given by

$$\mathcal{P}(\{\rho_1\} \cdots \{\rho_n\}; R_{\max} \ll r_{IJ}) = \prod_{I=1}^n \mathcal{P}(\{\rho_I\}) \left(1 + \sum_{I < J} \xi_{\circ}(r_{IJ}) b_{\circ}(\{\rho_I\}) b_{\circ}(\{\rho_J\}) \right). \quad (2.255)$$

where $R_{\max} = \max_I(\{R_I\})$ and the pairwise separations are r_{IJ} .

Chapter 3

Extreme Spheres: Counts-in-cells for 21cm intensity mapping¹

Summary

Intensity mapping surveys will provide access to a coarse view of the cosmic large-scale structure at high redshifts. Given the large fraction of the sky that can be efficiently scanned using emission from cosmic neutral hydrogen (HI), intensity mapping is ideally suited to probe a wide range of density environments and hence to constrain cosmology and fundamental physics. To efficiently extract information from 21cm intensities beyond average, one needs non-Gaussian statistics that capture large deviations from mean HI density. Counts-in-cells are ideally suited for this purpose, as their statistics can be predicted accurately. We use a large state-of-the-art magneto-hydrodynamic simulation from the IllustrisTNG project to determine the relation between neutral hydrogen and matter densities in cells. We demonstrate how our theoretical knowledge about the matter probability density function (PDF) for a given cosmology can be used to extract a parametrisation-independent HI bias function from a measured HI PDF. Inspired by the shape of this bias function, we employ a simple quadratic approximation that reproduces the measured bias function at a few percent level. Combining this quadratic bias function with the predicted matter PDF yields a fully-predictive model for the HI PDF that matches the measured HI PDF at a few percent accuracy at scale $R = 5 h^{-1} \text{Mpc}$ from redshift $z = 5$ to $z = 1$. Furthermore, we find a density-dependent HI clustering signal that is consistent with theoretical expectations and could allow for joint constraints of HI bias and the amplitude of matter fluctuations or the growth of structure.

¹This chapter is based on [1].

Outline

This chapter is organised as follows: We motivate the application of the counts-in-cells (CIC) framework to intensity mapping in section 3.1. Section 3.2 describes the IllustrisTNG simulation and how we extract the CIC statistics. In section 3.3 we briefly recap the theoretical formalism, introduced in section 2.5, that allows us to obtain the PDF and the density-dependent correlation of matter densities in spheres. Section 3.4 discusses how one can relate these results to the tracer PDF and density-dependent clustering using a bias model. We present the results for the bias relation between matter and neutral hydrogen along with the combined predictions for the neutral hydrogen PDF and density-dependence of clustering in section 3.5. Section 3.6 presents our conclusions and provides an outlook on the potential applications of our findings.

3.1 Introduction

Upcoming large-scale, post-reionisation intensity mapping surveys like Tianlai [30], BINGO [31], CHIME [32], FAST [33], HIRAX [34], MeerKAT [35], SKA [36] and SPHEREx [121] will sample the spatial distribution of cosmic matter through tracers of it, such as neutral hydrogen, at redshifts $0 < z < 6$. The advantages of those surveys with respect to traditional optical methods to map galaxies is that they can sample very large cosmological volumes in a very efficient manner. Following early ideas of intensity mapping [122–125], the first detection of the 21cm cosmological signal was achieved by cross-correlating 21cm intensity maps from the Green Back Telescope with the DEEP2 optical galaxy survey [126–128]. While we have not yet detected the 21cm cosmological signal in auto-correlation in the post-reionisation era² [130], upcoming surveys will have enough sensitivity to allow us to study cosmology at an unprecedented precision with both auto- and cross-correlations [131–139]. Furthermore, we can probe dark energy through baryonic acoustic oscillations [140, 141] or approach weak lensing of intensity mapping [142–145] by using the background as a source image.

It is well known that the non-linear evolution of matter in the Universe introduces a leakage of information from the matter two-point correlation function (or the power spectrum) into higher-order terms [146]. Thus, in order to extract the maximum information from large-scale structure surveys at low redshifts, we need to consider quantities beyond the two-point correlation or to attempt to reconstruct the linear fields [147]. In this chapter, we

²see [129] for a detection claim at high-redshift.

focus on the former approach and consider one-point statistics as complementary source of cosmological information compared to traditional two-point statistics.

For the epoch of reionisation, one-point statistics and higher-order moments have been proposed as sources of information about the physics of reionisation and the nature of ionising sources [148–152]. At later times, counts-in-cells statistics can capture essential non-Gaussian information from the 21cm intensity (and hence HI density) field that is lost in common two-point statistics and add information about the density-dependence of clustering. Furthermore, the underlying matter statistics in real-space can be analytically predicted [101] from first principles for scales at which the variance of the smoothed matter density is below unity. While, the PDF is obtained from an inverse Laplace transform that requires an integration in the complex plane in general, [7] has shown that one can perform an analytical saddle-point approximation in the log-density to obtain a closed-form expression that is valid for a wide range of densities. This approximation relies on an expansion in the non-linear variance and is hence valid at scales with variance below unity. Those scales are typically above $10 h^{-1}\text{Mpc}$ at redshift $z = 0$ which means that the typical low-angular resolution inherent to intensity mapping is not a major limiting factor.

The formalism, based on large-deviation statistics, allows us to access the rare event tails probing large density fluctuations that contain valuable information about fundamental physics (such as neutrino masses, primordial non-Gaussianity and modified gravity) that are inaccessible to common perturbative methods. The advantage of the predictive model for dark matter counts-in-cells statistics lies in its explicit dependence on cosmology through a) the statistics of the initial conditions, which enter through the linear matter power spectrum, b) the non-linear amplitude of fluctuations, which probes the expansion history including the dark energy equation of state [153], and c) the spherical collapse dynamics, which would be sensitive to departures from general relativity. To tap the potential of this probe for cosmology, a counts-in-cell specific bias function has to be established in order to translate predictions (PDF and density-dependent clustering) from dark matter to the various tracers.

Until recently, the matter-tracer relationship for counts-in-cells has been mainly investigated for dark matter halos and galaxies [154, 155, 68, 156–158]. Building upon a previous study of dark matter halos [159], we quantify the bias function that relates matter and neutral hydrogen counts-in-cells on scales where the bias function is non-linear and distinct from the bias measured in two-point clustering [160, 161].

While we focus on cosmology here, counts-in-cells statistics are also used to constrain important astrophysical ingredients such as luminosity functions [162]. Those could potentially be improved by predictions from large-deviation statistics, which are more accurate than phenomenological log-normal models that are currently used.

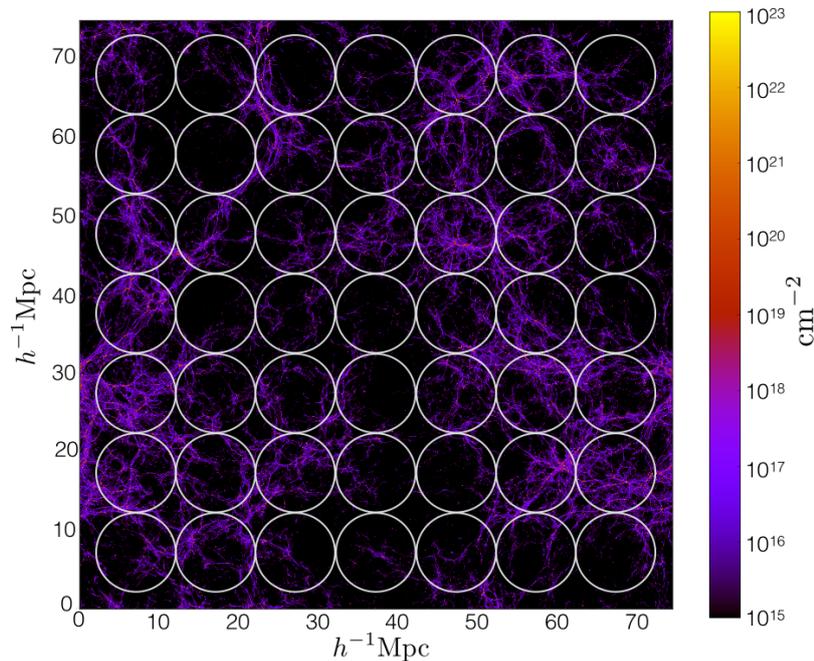


Fig. 3.1 Sketch of simulation measurements. Spatial distribution of neutral hydrogen in a $5 h^{-1}\text{Mpc}$ slice of the TNG100 simulation at redshift $z = 3$. The color is showing the numerical value of the HI column density in co-moving units. We also show a small subset of the spheres with radius $R = 5 h^{-1}\text{Mpc}$ used for counts-in-cells statistics.

We use the magneto-hydrodynamic simulation IllustrisTNG to compare counts-in-cells of neutral hydrogen, halos (as galaxy proxies) and matter. Extending recent results for the clustering statistics of neutral hydrogen [161], we quantify the counts-in-cells bias between the different tracers in IllustrisTNG and determine the effect of redshift-space distortions, which are caused by peculiar velocities that affect the observed redshift. Based on this, we assess the promise of mock catalogs that rely on simplified semi-analytical recipes to generate matter and tracer fields to efficiently access exquisitely sampled counts-in-cells while mitigating inaccuracies from neglecting fully non-linear effects that appeared as deal breakers for the matter power spectrum.

3.2 Numerical simulation

The simulation used in this work is part of the IllustrisTNG project [see 163–167, where stellar mass and assembly, clustering, colours, magnetic fields and chemical enrichment are discussed]. We employ here the TNG100 cosmological box (containing the same volume as the original Illustris simulation; 168–170) that has been evolved down to $z = 0$, with a co-moving box side length of $75 h^{-1}\text{Mpc}$.

The simulation was run with the AREPO code [171], which evolves the initial conditions accounting for gravity (using a TreePM method), magneto-hydrodynamics [172, through a Godunov approach on a moving Voronoi mesh] and a range of astrophysical processes described by subgrid models. These processes include primordial and metal-line cooling, assuming a time-dependent uniform UV background radiation, star and supermassive black hole formation, stellar population evolution that enriches surrounding gas with heavy elements or metals, galactic winds, and several modes of black hole feedback. The numerical methods and subgrid physics models are an update of the Illustris galaxy formation model [173, 174] and specified in detail in [175] and [176]. Importantly, where uncertainty and freedom exist for the implementation of these subgrid models, they are parametrised and tuned to obtain a reasonable match to a small set of observational properties [163]. These include the galaxy stellar mass function, the stellar-to-halo mass relation, and the stellar size-mass relation, all at $z = 0$. Additional aspects including feedback effects of black holes, the evolution of galaxy and halo sizes and metallicities are discussed in [177–179] and [180].

HI modelling. The spatial distribution of neutral hydrogen accounting for ionisation equilibrium with the UV background, HI self-shielding and the presence of molecular hydrogen is modelled following the method depicted in [161], to which we refer the reader for further details. For a study on the impact of the molecular hydrogen (H_2) model on HI properties, see [181].

The dependency between neutral hydrogen and the surrounding matter is highly non-linear as can be seen in Figures 14 and 15 in [161]. The corresponding halo-HI mass function shows a cut-off for small halo masses that is explained in detail in Appendix B of [161]. The mass fraction of hydrogen in the neutral phase was studied as a function of the mass and force resolution in [181]. To analyze the convergence of the neutral hydrogen fraction in the IllustrisTNG100 simulation, the authors use predictions from the lower resolution IllustrisTNG100-1 and IllustrisTNG100-2 simulations for comparison. They find that the convergence of the neutral hydrogen fraction is strongly mass dependent: At low stellar masses ($10^9 - 10^{10} M_\odot$), the comparison of IllustrisTNG100-1 and IllustrisTNG100 reveals an almost perfect convergence whereas mass fractions can differ by up to a factor of 4 at higher stellar masses. [181] does not study the impact of the box size on simulation outcomes since the larger IllustrisTNG300 simulation has an insufficient resolution (similar to IllustrisTNG100-2) and the smaller IllustrisTNG50 simulation was not available at the time. Moreover, we will argue in the following paragraph about counts-in-cells that the 8x smaller volume of the IllustrisTNG50 simulation would drastically reduce the statistical power of our counts-in-cells estimator to the point where one probably would not be able to

disentangle inaccuracies in the simulation and statistical uncertainties. The existence of the cut-off and the convergence of the neutral hydrogen fraction for low masses, indicates that our HI modelling is robust and does not strongly depend on the chosen simulation.

Halo identification. In this chapter, we work with halos identified by the Friends-of-Friends (FoF) algorithm with a linking length of $b = 0.2$ [182]. The halo centre is identified as the position of the most bound particle in the halo. The minimum halo mass we consider is around $2 \times 10^8 M_\odot/h$.

Counts-in-cells. We extract counts-in-cells (i.e. mean densities in spheres) of the matter, neutral hydrogen, and mass-weighted halo field in overlapping spheres of co-moving radius $R = 5 h^{-1} \text{Mpc}$ on a regular grid of size 128^3 yielding approximately 2 million density samples. Fig. 3.1 shows a snapshot of the neutral hydrogen distribution in the TNG100 simulation at redshift $z = 3$ along with a subsample of the spheres used for our counts-in-cells analysis. In the redshift range of interest, $z = 1 - 5$, the size of the spheres corresponds to angular scales between $\theta = 14$ and 4 arcmin.

When we write ρ_{HI} , ρ_{m} and ρ_{HM} , we refer to the density in spheres (hence smoothed at radius R) of neutral hydrogen, matter and mass-weighted halos, respectively. Each gas Voronoi cell carries information on the state of the gas, such as density, pressure, mass and metallicity. The HI mass associated to each gas cell is computed as described in section 2.2 of [161]. We distribute the HI mass of each gas cell in a regular grid of 2048^3 grid cells using the nearest grid point (NGP) interpolation mass scheme. A similar procedure is used for the total matter field, where we use the gas, dark matter, stars and black hole particles. Finally, the same procedure is used to estimate the halo density field. All densities are measured in units of mean density and are hence dimensionless and related to the density contrast δ as $\rho = 1 + \delta$. Note that when measuring this for discrete tracers such as galaxies that are a coarse and biased sampling of the underlying field, one needs to include shot-noise contributions, for example through Poisson sampling, see e.g. [154, 157]. The objects considered in this work are expected to have very low shot-noise amplitudes³, so we neglect, for simplicity, their contribution to counts-in-cells.

The PDFs, encoding the probability density of finding a certain density in a randomly drawn sphere of fixed radius, were estimated using kernel density estimation. In cases where discretisation is needed, we use histograms with a logarithmic binning in densities such that each bin contains approximately $1/75$ of the probability mass. Error bars at those sampling

³The very high-resolution of the IllustrisTNG simulation guarantees a very low shot-noise amplitude for the matter field. [160, 161] have shown that the amplitude of the shot-noise is negligible for neutral hydrogen in the post-reionisation era. Mass-weighted halos are also expected to have a very low shot-noise amplitude [183].

points are determined by means of a jackknife estimator⁴ with 30 random subsamples. The random selection ensures more independent spheres within each subsample in comparison to splitting the volume into 30 regular subboxes. It mitigates effects from long range modes too, even though there will still be super-sample variance effects⁵ due to the small box volume. The random selection might underestimate the error bars, as the different subsamples are correlated. Due to the small cosmological box with a side length of only $75 h^{-1}\text{Mpc}$, we chose spheres of radius $R = 5 h^{-1}\text{Mpc}$. This scale is a compromise between two contrasting requirements: a) due to the expected reach of the theory, we need a small variance and hence sufficiently large spheres, b) in order to have good statistics, we require a large enough number of independent spheres and hence sufficiently small spheres. Since our spheres of radius $R = 5 h^{-1}\text{Mpc}$ are about a factor of 2 smaller than the scales one would want to probe at the lowest redshift $z = 0$ to ensure a variance below unity, we will focus on results for $z > 0$. Since the shape of the counts-in-cells PDF is driven by the non-linear matter variance $\sigma^2(R, z)$ at radius R and redshift z , which is related to the non-linear power spectrum according to Eq. (3.2), our formalism can be used to relate different radii and redshifts; for example, the PDF for $R = 10 h^{-1}\text{Mpc}$ at $z = 0$ closely corresponds to that for $R = 5 h^{-1}\text{Mpc}$ at $z = 1$, as the amplitude of fluctuations is almost the same in those two cases.

Redshift-space mapping. To assess the impact of redshift-space distortions, a mapping from real-space to redshift-space is done by converting the co-moving positions (of matter, halos and neutral hydrogen) \mathbf{x} to the redshift-space ones, \mathbf{s} , by shifting them along the fictitious line-of-sight (chosen in x -direction here) according to their peculiar velocity along that direction via (2.106).

3.3 Statistics of matter densities in spheres

Before moving on to tracers in the next section, let us start with the statistics of the matter density field that is the first building block in our modelling of the statistics of neutral hydrogen densities. We have introduced the CIC formalism in section 2.5 and refer readers there for more details.

⁴A subsampling technique estimating the variance of an estimator based on its variation over the subsamples.

⁵Effects from modes larger than the survey.

3.3.1 One-point PDF of matter density

In section 2.5 we have derived the PDF for the matter density in a sphere of radius R , for small values of the variance (2.223)

$$P_R(\rho_m) = \sqrt{\frac{\Psi_R''(\rho_m) + \Psi_R'(\rho_m)/\rho_m}{2\pi}} \exp(-\Psi_R(\rho_m)),$$

where the prime denotes a derivative with respect to ρ_m and the exponential decay of the PDF is given by the rate function (2.222)

$$\Psi_R(\rho_m) = \frac{\sigma_L^2(R) \left(\delta^{(1)}(\rho_m) \right)^2}{\sigma_\mu^2 2\sigma_L^2(R\rho_m^{1/3})}.$$

Here $\delta^{(1)}(\rho_m)$ is the linear density contrast, averaged within the initial Lagrangian radius $R_{\text{ini}} = R\rho_m^{1/3}$, which can be mapped to the non-linearly evolved density ρ_m within radius R using the spherical collapse model (2.178), σ_L^2 is the linear variance from equation (2.208).

Thus, the functional form of the PDF is fully specified by the scale-dependence of the linear variance computed via the linear power spectrum according to equation (2.208) and the spherical collapse model (2.178). The only free parameter enters through $\sigma_\mu^2 \equiv \sigma_\mu^2(R, z)$, which is the non-linear variance of the log-density (because the formula has been derived from an analytic approximation based on the log-density $\mu_m = \log \rho_m$).

To ensure a unit mean density and the correct normalization of the PDF, one has to evaluate the PDF obtained from equation (2.223) according to (2.225)

$$\mathcal{P}_R(\rho_m) = P_R \left(\rho_m \frac{\langle \rho_m \rangle_{\text{SPA}}}{\langle 1 \rangle_{\text{SPA}}} \right) \cdot \frac{\langle \rho_m \rangle_{\text{SPA}}}{\langle 1 \rangle_{\text{SPA}}^2},$$

with the shorthand notation $\langle f(\rho_m) \rangle_{\text{SPA}} = \int_0^\infty d\rho_m f(\rho_m) P_R(\rho_m)$. This step is necessary as equation (2.223) ensures the correct tree-level cumulants of order 3 and above, the right non-linear variance of μ_m and zero mean for μ_m . Since we, instead, want the density ρ_m to have unit mean, it is necessary to correct for the non-zero value of the mean of μ_m using equation (2.225).

[7] showed that the above model for the PDF of the real-space matter density field $\mathcal{P}_R(\rho_m | \sigma_\mu)$ with the variance of the log-density $\sigma_\mu^2(R)$ as a driving parameter was accurate at the percent level for standard deviations $\sigma_\mu \lesssim 0.5$, greatly improving over the accuracy that can be obtained from log-normal models [113]. To give an impression of the exquisite accuracy, we show a comparison between the measurement from the IllustrisTNG simulation

Table 3.1 The measured non-linear variances σ of the log-density $\mu = \log \rho$ of matter (m) in spheres of $R = 5 h^{-1} \text{Mpc}$ at redshifts $z = 0$ to 5. They give an indication of how well the fully-predictive matter theory works, as it requires $\sigma_{\mu,m} < 1$ to converge.

z	0	1	2	3	4	5
$\sigma_{\mu,m}$	0.851	0.602	0.438	0.338	0.275	0.231

and the theoretical prediction from equations (2.223) for the matter PDF in Fig. 3.2. More precisely, the prediction used the linear variance from equation (2.208) computed with the input power spectrum, and the measured non-linear variances σ_{μ}^2 from the simulation (listed in Table 3.1) as input parameter to the PDF in equation (2.223). As expected, the PDF is close to Gaussian at high redshift and becomes more and more skewed at lower redshifts, as voids are occupying most of the volume and density peaks are exceedingly accreting matter.

When aiming at a fully-predictive model for the matter PDF, one can use the non-linear variance from a non-linear power spectrum P_{NL} , as predicted, for instance, by the phenomenologically motivated HALOFIT emulator [4, 5], in analogy to the linear variance from equation (2.208),

$$\sigma^2(r) = \int \frac{d^3k}{(2\pi)^3} P_{\text{NL}}(k) W^2(kr). \quad (3.2)$$

The non-linear variance of the log-density σ_{μ}^2 that enters the PDF is then chosen such that the variance of the PDF in equation (2.223) matches the non-linear density variance σ^2 from HALOFIT. The HALOFIT non-linear variance agrees with the measured value to typically better than 1% which propagates to an additional 1-2% error on the PDF. At $z = 0$, our method of estimating the logarithmic variance from HALOFIT no longer works and causes significant discrepancies. This is not unexpected, given that we are probing smaller scales than would be desirable at this redshift, and on smaller scales baryonic effects become important, which are not captured in the fit to N-body simulations. Note that in general one can treat σ_{μ} as a free parameter of the theory, and we only rely on an approximate predicted value of σ_{μ} to infer the functional form of the bias relation between matter and neutral hydrogen.

3.3.2 Density-dependent clustering of matter

Apart from the one-point statistics of density in spheres, one can also extract a density-dependent clustering signal that quantifies the difference in clustering of regions with high/low densities compared to regions with average densities. This is encoded in the density-dependent correlation function, which is the ratio of the joint two-point PDF of

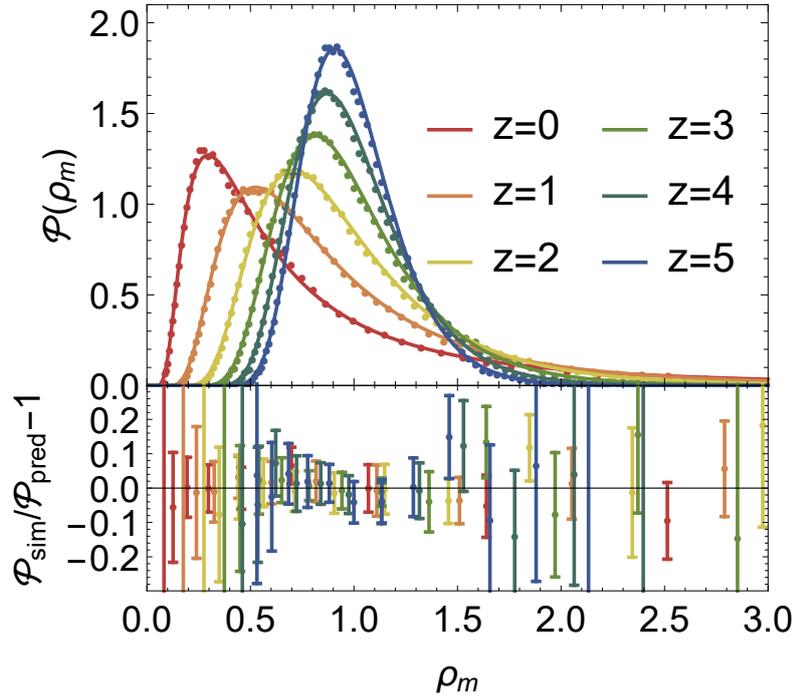


Fig. 3.2 Dark matter PDF showing percent-level agreement between simulation and prediction. Upper panel: Measured PDF of matter densities in spheres of radius $R = 5 h^{-1} \text{Mpc}$ at redshifts $z = 0, 1, 2, 3, 4, 5$ (red to blue data points) with measured non-linear variance as input (solid lines), compared to the prediction from large-deviation statistics. For clarity, error bars are omitted and only shown in the lower panel. Lower panel: Residuals between the theoretical predictions with measured variance and the measured PDFs.

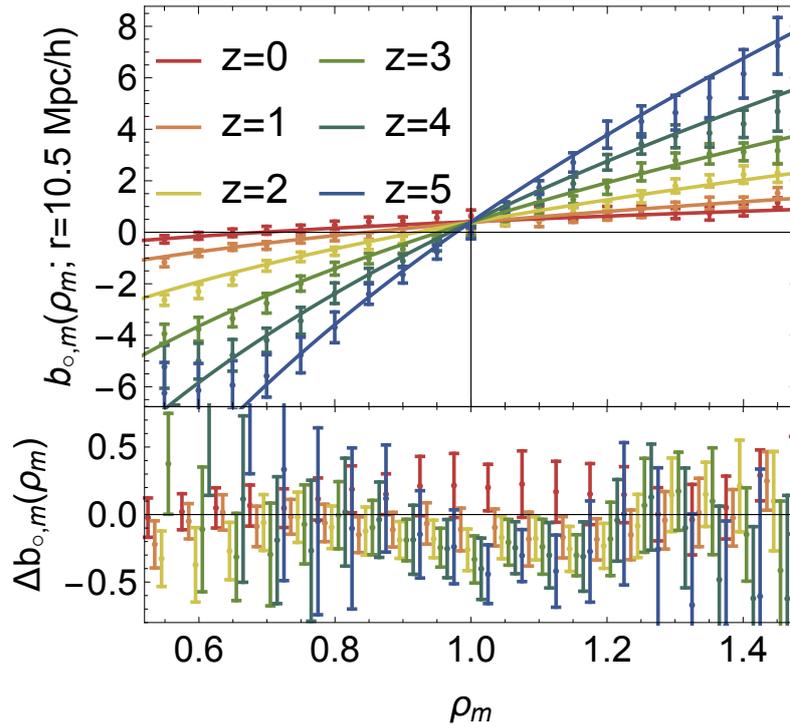


Fig. 3.3 Measured density-dependent matter clustering agrees with theoretical expectations. Upper panel: Density-dependent clustering encoded in the sphere bias function from equation (2.232) for matter at redshifts $z = 0$ to 5 (red to blue) as measured in the IllustrisTNG simulation (data points) and predicted by large-deviation statistics using equation (3.3) (lines). Lower panel: Residuals between the theoretical prediction with measured variance and the measured density-dependent clustering. For better visibility, the abscissas for different redshifts are slightly shifted.

matter densities at separation r and the marginal one-point PDFs (2.227)

$$\xi_{\circ,m}(\rho_m, \rho'_m, r) = \frac{\mathcal{P}_R(\rho_m, \rho'_m; r)}{\mathcal{P}_R(\rho_m) \mathcal{P}_R(\rho'_m)} - 1.$$

At large separation $r \gg 2R$, this correlation function factorises into an average separation-dependent correlation $\xi_{\circ,m}(r)$, and a density-dependent modulation called sphere bias $b_{\circ,m}(\rho_m)$ (2.228) [8]

$$\xi_{\circ,m}(\rho_m, \rho'_m, r) \simeq \xi_{\circ,m}(r) b_{\circ,m}(\rho_m) b_{\circ,m}(\rho'_m).$$

The mean sphere two-point correlation, $\xi_{\circ,m}(r)$, is the standard matter correlation function of densities smoothed at the sphere radius R . The sphere bias measures the excess correlation induced by a matter density ρ_m at separation r and therefore can be defined as a ratio of the conditional mean of the sphere density ρ'_m given a density ρ_m at separation r (2.232) and (2.229)

$$b_{\circ,m}(\rho_m) = \frac{\langle \rho'_m | \rho_m; r \rangle - 1}{\xi_{\circ,m}(r)},$$

$$\xi_{\circ,m}(r) = \langle \rho_m(\mathbf{x}) \rho_m(\mathbf{x} + \mathbf{r}) \rangle - 1.$$

At large separation, the sphere bias becomes independent of the separation r and in general can be predicted more accurately than the approximation for the full two-point clustering (2.227). It can be computed using large-deviation statistics with spherical collapse and a rescaling to the non-linear variance [184, 185, 8, 118]. Using the saddle-point approximation for two distant spheres of equal radius, the general expression (2.254a) simplifies to

$$b_{\circ,m}(\rho_m) = \frac{\delta(\rho_m) \sigma_L^2(R)}{\sigma_L^2(R \rho_m^{1/3}) \sigma_\mu^2}, \quad (3.3a)$$

and the normalisation (2.254b) becomes

$$\hat{b}_{\circ,m}(\rho_m) = \frac{b_{\circ,m}(\rho_m) - \langle b_{\circ,m}(\rho_m) \rangle}{\langle (\rho_m - 1) b_{\circ,m}(\rho_m) \rangle}. \quad (3.3b)$$

The validity of equation (3.3) has been established before for dark matter in large cosmological simulations [118, 159] and found to be surprisingly accurate even for separations $r \gtrsim 2R$, that are only slightly larger than twice the sphere radius.

As before, a theoretical prediction can be obtained using the linear variance computed from the input power spectrum of the simulation and the measured non-linear matter variance,

Table 3.2 Statistical properties of neutral hydrogen distribution in spheres of $R = 5 h^{-1} \text{Mpc}$ at redshifts $z = 1, 3, 5$. The measured non-linear variances σ of the log-density $\mu = \log \rho$ of neutral hydrogen (HI) in real and redshift-space, shown in the left two columns, can be compared to the matter variances in Table 3.1 to estimate the linear bias. The linear cross-correlation coefficients (3.12) between matter (m), neutral hydrogen in real-space (HI) and redshift-space (HIz) and mass-weighted halos (HM), shown in the right four columns, indicate a tight relation between the different fields.

z	$\sigma_{\mu, \text{HI}}$	$\sigma_{\mu, \text{HI}}^z$	$r_{\text{m, HI}}$	$r_{\text{m, HIz}}$	$r_{\text{m, HM}}$	$r_{\text{HM, HI}}$
1	1.228	1.311	0.969	0.918	0.982	0.945
3	0.986	1.081	0.963	0.944	0.987	0.989
5	0.798	0.878	0.958	0.943	0.983	0.990

stated in Table 3.1, as input. The sphere bias is measured from the simulation using a discretisation of equation (2.227). The average sphere correlation is determined by considering all pairs of spheres which are at separation r along one of the axis in the regular density grid extracted from the simulation box. The conditional mean is computed by selecting all pairs where one density is in a given bin of size $\Delta\rho = 0.05$ around the desired value ρ and computing the mean of the paired density ρ' . Error bars are determined by means of a jackknife estimator with 30 random subsamples. Due to the small box size of TNG100, it is difficult to get reliable measurements of the correlations in the large separation regime, such that we have to rely on a separation that just ensures non-overlapping spheres. Fig. 3.3 displays such a measurement of the density-dependent clustering signal for matter densities in real-space, which should be interpreted with caution (given that even the minimal separation of non-touching spheres is more than one tenth of the box size), but is completely consistent with the theoretical expectation and previous measurements in large N-body simulations. As expected, spheres of particularly large or small densities are more strongly clustered than average densities. When plotted as a function of the density, the clustering appears stronger at high redshifts, because the matter variance is smaller and hence the relative density contrast is larger than at low redshifts.

3.4 Statistics of tracer densities in spheres

Let us now turn to biased tracers of matter and describe how the previous results for matter densities in spheres can be mapped to tracer densities in spheres.

3.4.1 One-point PDF of tracer density

In general, one can express the respective one-point PDFs of matter and the tracer as marginals of their joint one-point PDF $\mathcal{P}(\rho_m, \rho_t)$

$$\mathcal{P}_t(\rho_t) = \int d\rho_m \mathcal{P}(\rho_m, \rho_t), \quad \mathcal{P}_m(\rho_m) = \int d\rho_t \mathcal{P}(\rho_m, \rho_t). \quad (3.4)$$

For simplicity, our bias model is formulated between matter and tracer densities (such as neutral hydrogen, halos or galaxies) for spheres of identical radii. While this is a simplistic approximation, as in general the relationship is non-local in both space and time [for a review see 68], we will show that it is sufficient for our purpose.

The tracer density PDF, \mathcal{P}_t , can be written as a convolution of the matter PDF, \mathcal{P}_m , and the conditional PDF of finding a certain tracer density given a matter density

$$\mathcal{P}_t(\rho_t) = \int d\rho_m \mathcal{P}_{\text{bias}}(\rho_t|\rho_m) \mathcal{P}_m(\rho_m), \quad (3.5)$$

where $\mathcal{P}_{\text{bias}}(\rho_t|\rho_m)$ is the conditional PDF (i.e. the probability of having a tracer density ρ_t given a matter density ρ_m). This conditional depends on the details of tracer formation and its associated parameters such as halo mass, smoothing scales, redshift and environment, but also includes scatter around any deterministic relation (stochasticity) which results from an incomplete understanding of the formation process. While in principle one could think that the full joint PDF is needed, one can separate this information into the marginals, the one-point PDFs which are of interest here, and correlations between matter and tracer densities that are independent of the marginal PDFs [186].

In the following, we will focus on the marginals and determine an accurate mean bias relation that allows for a one-to-one relation between the matter and tracer PDF. This is in the spirit of large-deviation statistics, which argues that the mean local gravitational evolution given by spherical collapse is adequate to predict the PDF of matter densities in spheres⁶. Equipped with a bias model for the mean relation $\rho_m(\rho_t)$, the tracer PDF \mathcal{P}_t is now obtained from the matter PDF \mathcal{P}_m by conservation of probability

$$\mathcal{P}_t(\rho_t) = \mathcal{P}_m(\rho_m(\rho_t)) |d\rho_m/d\rho_t|, \quad (3.6)$$

where it is required that $\rho_m(\rho_t)$ is a strictly monotonic function.

⁶The large-deviation principle states that the statistics is dominated by the path that minimises the ‘‘action’’ – or in our case the exponential decay of the PDF – in order to maximise the probability. This most likely path or dynamics can be decomposed into a gravitational part, given by the spherical collapse, and an astrophysical part, given by the mean bias relation.

3.4.2 Density-dependent clustering of tracers

Using a mean bias model, one can also relate the density-dependent clustering of matter in equation (3.3) to a tracer

$$b_{o,t}(\rho_t) = b_{o,m}(\rho_m(\rho_t)) \sqrt{\xi_{o,m}/\xi_{o,t}}, \quad (3.7)$$

where the ratio of correlation functions can be computed as a sphere-bias weighted mean of the bias relation

$$\sqrt{\xi_{o,m}/\xi_{o,t}} = \langle \rho_t(\rho_m) b_{o,m}(\rho_m) \rangle. \quad (3.8)$$

3.4.3 Parametrisation-independent bias functions

Previously, we have seen that the question of how to obtain an accurate model for the statistics of tracer densities in spheres boils down to successfully describing the effective mean bias relation between matter and the corresponding tracer densities in spheres.

The advantage of obtaining bias functions in a parametrisation-independent way is that they can be used as a guiding principle for finding suitable parametrisations with a small number of parameters that capture their functional form. This is particularly important if one is interested in the tails of the distribution where common polynomial bias models do not lead to satisfactory results. Since we want to map the matter PDF to the tracer PDF, let us rely on an ‘inverse’ bias model $\rho_m(\rho_t)$ writing the matter density as a function of the tracer density rather than the other way around.

Bias function from abundance matching

Following the ideas of [187, 154], a direct way to obtain a mean bias relation is to use the cumulative distribution functions (CDFs), defined as $\mathcal{C}(\rho) = \int_0^\rho d\rho' \mathcal{P}(\rho')$, and match their abundances

$$\mathcal{C}_m(\rho_m) = \mathcal{C}_t(\rho_t), \quad (3.9)$$

such that

$$\rho_m(\rho_t) = \mathcal{C}_m^{-1}(\mathcal{C}_t(\rho_t)), \quad \rho_t(\rho_m) = \mathcal{C}_t^{-1}(\mathcal{C}_m(\rho_m)). \quad (3.10)$$

Note that this bias function is built to relate the PDFs of tracer and matter densities by a one-to-one monotonic mapping, which does not assume a local relationship between matter and tracer densities. If there is a large correlation between the matter and its tracer field (for

a quantification see the cross-correlation coefficient defined below), this bias function also provides a good fit to a local scatter plot between matter and tracer densities.

Bias function from conditional mean

When assuming a local relation between tracer and matter densities, one can also infer mean bias relations from the conditional mean from the scatter plot (SP)

$$\rho_m^{\text{SP}}(\rho_t) := \langle \rho_m | \rho_t \rangle, \quad \rho_t^{\text{SP}}(\rho_m) := \langle \rho_t | \rho_m \rangle. \quad (3.11)$$

Note that, in contrast to the bias function from abundance matching, the composition of the inverse and forward conditional mean bias is not guaranteed to give the identity mapping $\rho_m^{\text{SP}}(\rho_t^{\text{SP}}(\rho_m)) \neq \rho_m$. In particular, this can be a signal for non-linear bias and a difference in the scatter when fixing matter or tracer density, respectively. We checked that those conditional mean inverse bias functions will be close to the inverse bias function inferred from the CDF method within 3%. We prefer the CDF method as it is guaranteed to provide a good description for mapping marginal PDFs with conservation of probability.

The (linear) cross-correlation coefficient r between matter and tracer densities in spheres is defined as

$$r = \frac{\langle \rho_m \rho_t \rangle - 1}{\sqrt{\langle \rho_m^2 \rangle_c \langle \rho_t^2 \rangle_c}}. \quad (3.12)$$

We show some of the correlation coefficients between matter, halo mass and neutral hydrogen in Table 3.2. For all redshifts, correlations between matter, neutral hydrogen and mass-weighted halos are all very high and above 95%, and neutral hydrogen is almost perfectly correlated with mass-weighted halos for the higher redshifts. Even when comparing matter in real-space to neutral hydrogen in redshift-space, thus absorbing redshift-space distortions in the bias, correlations are still well above 90%. This ensures that the bias function from abundance matching, determined purely from the marginal PDFs, will also be a good fit to the scatter plots.

3.4.4 Polynomial bias model in log-densities

Following [188, 159], we use a quadratic model for the (inverse) bias of log-densities $\mu = \log \rho$ in spheres (rather than for the density contrast) which reads

$$\mu_m = \sum_{n=0}^{n_{\text{max}}} b_n \mu_t^n, \quad n_{\text{max}} = 2. \quad (3.13)$$

A heuristic explanation for why a logarithmic transform helps is that it makes the underlying one-point PDFs of matter and biased tracers significantly more Gaussian [189, 190]. Hence, it provides a local remapping of non-linear densities that approximates initial (Lagrangian) densities for which local polynomial bias models are more adequate than for evolved (Eulerian) densities. As already emphasised in [188], the reason why equation (3.13) can be approximated by a linear bias model for the density fluctuations $\delta_t = \hat{b}_1 \delta_m$ on large scales is that the ranges of log-densities μ_t and μ_m become small and not because the bias relation itself becomes linear. This is particularly relevant when focusing on the tails of the distribution of densities and hence the regime where linear bias is insufficient.

Another approach for the bias function involves choosing a polynomial bias model for density contrast and matching the first few moments between matter and tracer distributions [158]. Since both the bias model and the matching technique are emphasising the PDF region around its peak, the approach is not suited in our situation, since we want to preserve the information in the tails.

3.5 Results

We extend recent results for the relation between neutral hydrogen and matter densities in spheres from [161], by determining simple, yet accurate, bias parametrisations for counts-in-cells statistics. We demonstrate their ability to predict neutral hydrogen counts-in-cells at a few percent accuracy using a small number of bias parameters. When combined with the cosmology-dependence of the underlying dark matter counts-in-cells, our findings lay the foundation for a joint fit of cosmological and bias parameters, allowing to infer cosmological information by marginalising over astrophysical uncertainties.

3.5.1 Joint density distributions and mean bias model

Joint dark matter and neutral hydrogen distribution

Fig. 3.4 presents scatter plots comparing densities in spheres for neutral hydrogen ρ_{HI} versus matter densities in real-space ρ_m for redshifts $z = 1, 3, 5$ and radius $R = 5 h^{-1} \text{Mpc}$.

We observe that the neutral hydrogen and matter distributions are closely related, as expected from Table 3.2. Moreover, the scatter around the mean bias (solid line) increases with cosmic time due to gravitational collapse and astrophysical effects. To make contact with observables, it is required to consider the effect of redshift-space distortions. The joint measurements of neutral hydrogen in redshift- and matter in real-space are shown with the shaded points. We show these because our theoretical prediction (2.223) is for

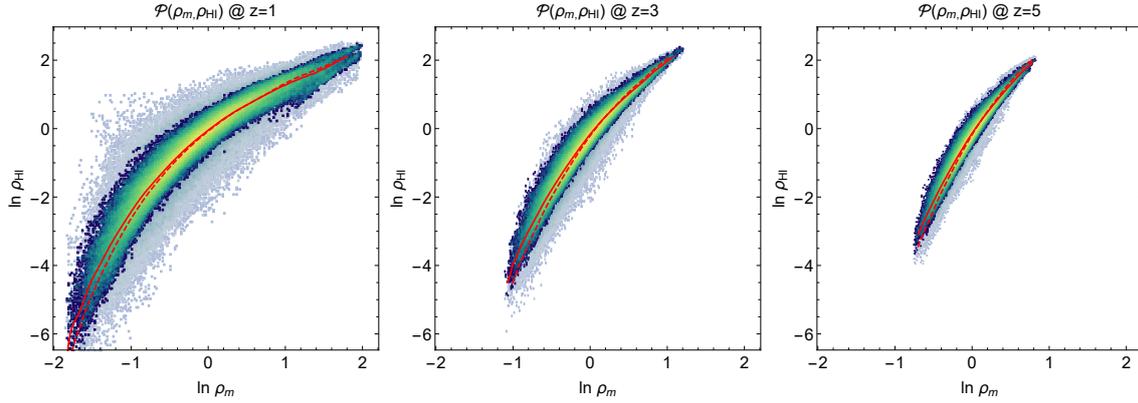


Fig. 3.4 HI vs. matter scatter plots in real- and redshift-space. Density scatter plots of the neutral hydrogen density ρ_{HI} in real-space (blue-green) and redshift-space (grey) versus the matter density in real-space ρ_m for radius $R = 5 h^{-1} \text{Mpc}$ at redshifts $z = 1, 3, 5$ (left to right). The colour (from light to dark) indicates the magnitude of the joint PDF in logarithmic scale and shows a concentration of probability around the mean bias. The figure also shows the parametrisation-independent bias obtained from the CDF in real-space (solid red line) and redshift-space (dashed red line). The main impact of redshift-space distortions is to increase the scatter while the mean bias relation is almost unchanged for average densities and mostly affected in the positive density tails.

matter in real-space and thus we incorporate the redshift-space mapping into the bias relation. Redshift-space distortions manifest themselves in an increased variance and are particularly strong at low redshifts where velocity dispersions within halos are typically larger. Moreover, as velocity dispersions become larger in overdense regions, the additional scatter increases with density, and coherent infall into overdensities increases the neutral hydrogen density in redshift-space (dashed lines) compared to real-space (solid lines). While most of the additional scatter when comparing neutral hydrogen densities in redshift-space to real-space matter densities comes from the non-linear mapping from real- to redshift-space, the scatter between redshift-space densities of neutral hydrogen and matter is larger than the scatter between the corresponding real-space densities, in particular at higher redshifts.

Joint neutral hydrogen and halo mass distribution

Furthermore, we present scatter plots between the mass-weighted halo density and neutral hydrogen in Fig. 3.5, which show a much more linear relation (in log-densities) than previously seen for the matter field. This reflects the fact that most of the neutral hydrogen mass is embedded into halos; nearly all mass at $z = 0$ and still 90% at $z = 5$ [161]. Since the amount of neutral hydrogen is sensitive to halo mass, the neutral hydrogen density in spheres is closely related to the mass-weighted halo density in spheres. Our scatter plots

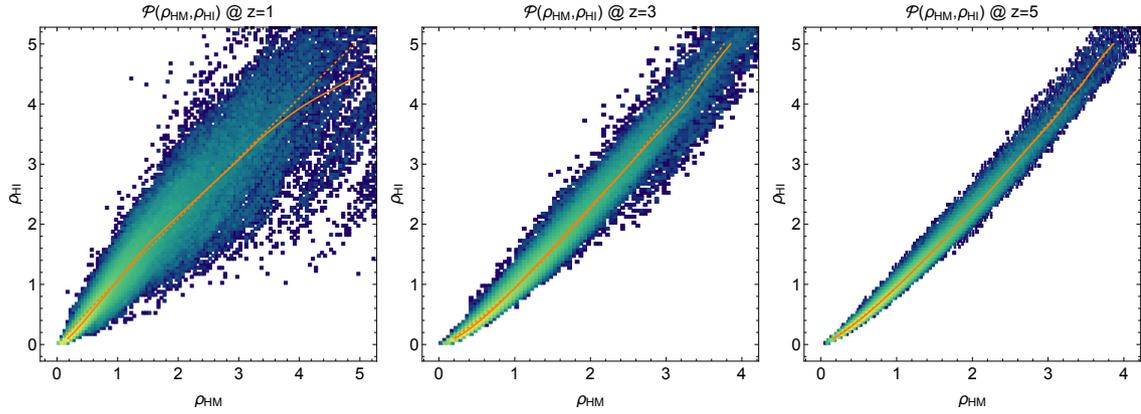


Fig. 3.5 HI vs. halo mass scatter plots in real-space. Density scatter plots of the real-space neutral hydrogen density ρ_{HI} versus the mass-weighted halo density ρ_{HM} for radius $R = 5 h^{-1} \text{Mpc}$ at redshifts $z = 1, 3, 5$ (left to right). The colour (from light to dark) indicates the magnitude of the joint PDF in logarithmic scale and shows a concentration of probability around the mean bias. The figure also shows the parametrisation-independent bias obtained from the CDF (solid orange line) and a linear fit in log-densities (dotted orange line). This plot shows that the neutral hydrogen closely traces the mass in halos with an almost linear relationship.

of average halo mass and neutral hydrogen in cells complement the halo HI mass functions from Fig. 4 of [161] showing the HI-halo relation on an object-by-object level. The relation of neutral hydrogen to mass-weighted halo densities is interesting for two reasons. First, mass-weighted halos are in turn closely related to luminosity-weighted galaxies. Hence, joint studies of intensity mapping and galaxy surveys [131, 132] could provide valuable information. In Fig. 3.6, we demonstrate the similarity of mass-weighted halo and neutral hydrogen PDFs by plotting them as a function of the log-density rescaled by the variance. We observe that both neutral hydrogen and halos are non-linearly biased with respect to the matter, but neutral hydrogen is close to a linearly biased version of halo mass. This motivates our quadratic bias parametrisation in log-densities from equation (3.13), which has been used before to model mass-weighted halos in [159]. Second, the close relation between halo mass and neutral hydrogen density in cells suggests that mocks obtained from populating bound dark matter structures with neutral hydrogen in a halo model approach [191, 192, 161, 193] are expected to give accurate results for counts-in-cells on those scales. This could make it possible to study counts-in-cells statistics of neutral hydrogen in larger volumes, taking advantage of cosmological simulations for dark matter.

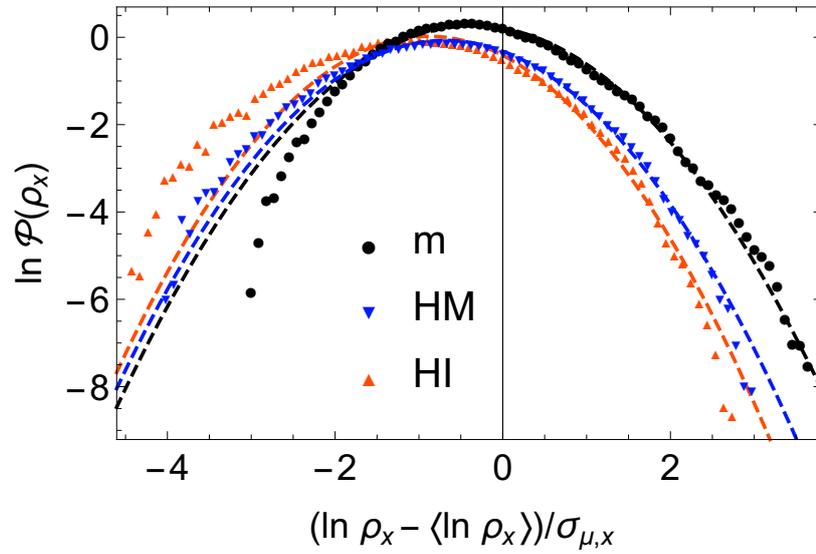


Fig. 3.6 Necessity of non-linear bias and beyond log-normal prediction. Comparison of measured density PDFs of matter (black), mass-weighted halos (blue) and neutral hydrogen (red) in spheres of radius $R = 5 h^{-1} \text{Mpc}$ at redshift $z = 3$ in real-space. The abscissa shows scaled log-densities such that for tracers that are linearly biased in log-densities curves would overlap. The lack of overlap between tracers and matter motivates the chosen quadratic bias model (3.13). Also shown are log-normal fits to the PDFs (dashed lines) which show significant deviations from the measurements for underdensities of matter and neutral hydrogen.

Mean bias parametrisations

Let us now focus on the mean CDF bias functions (3.10) which are depicted in the scatter plots in Figures 3.4 and 3.5. Moreover, we show in Fig. 3.7 a comparison of the CDF bias obtained from combining the measured neutral hydrogen PDF with either the measured matter PDF, or the fully-predictive matter PDF model with the HALOFIT variance both in real- and redshift-space. We find good agreement at sub-percent level close to the peak and deviations below 2% for a wide range of densities. This is encouraging, because it means that the fully-predictive matter PDF, combining (2.225) with the HALOFIT variance, can be used to extract a parametrisation-independent mean bias function from a measured neutral hydrogen PDF.

Next, we validate our bias parametrisation by comparing the quadratic fit in the log-density according to equation (3.13) and the parametrisation-independent CDF-bias (3.10) with measured HI PDF and predicted matter PDF and in Fig. 3.8. We fit the parametric bias model (3.13) to the parameter-independent bias function (3.10) using the chosen logarithmically spaced sampling points. We estimate the measurement errors at the sampling points using a jackknife estimator (i.e. we derive a CDF bias function for each subset and compute their scatter at the sampling points). Assuming the measurement errors to be normally distributed and performing an ordinary least squares regression, yields the bias parameters displayed in Table 3.3. As can be seen in Fig. 3.8, the quadratic bias model for the logarithmic densities agrees well with the measured function, both in real- and redshift-space. The residual plots show an accuracy of approximately 4% at $z = 1$ and 1% at $z = 3$ and 5 over a wide range of densities both in real (middle panel) and redshift-space (lower panel). When fitting the parametric model to the measured bias function (matter and HI from simulation), we reproduce the bias parameters from Table 3.3 at a few, sub-, and 10 percent level for b_0, b_1, b_2 respectively. We decided to employ the measurement errors from the purely measured scenario, as due to the high correlation between matter and neutral hydrogen, there are cancellations in the uncertainties which are not captured when using the theoretical matter CDF. When including the next higher order bias parameter b_3 , one finds that it is typically of order $b_2/10$, but can nevertheless slightly improve the approximation. However, we decided in favour of simplicity and truncated the expansion at second order.

Note that our bias parameters characterise the inverse relation (matter density as function of tracer density) and in particular our linear (inverse) bias b_1 will typically have values around $1/3 - 1/2$ signalling positive linear forward bias \tilde{b} around $2 - 3$ which is in line with previous studies of the bias from the two-point correlation at intermediate redshifts [161, 160].

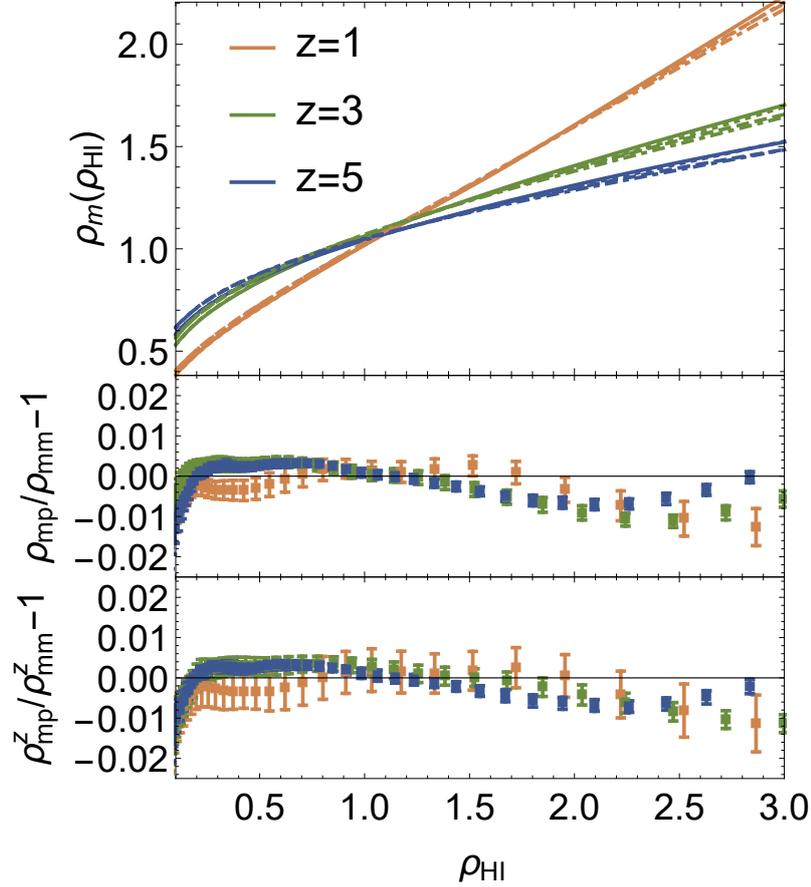


Fig. 3.7 Comparison of bias functions for fully-predictive and measured matter PDF. Upper panel: Bias functions for spheres of radius $R = 5 h^{-1} \text{Mpc}$ are displayed at redshifts $z = 1, 3, 5$ (orange to blue). The parametrisation-independent bias functions from the cumulative PDFs according to equation (3.10) using the measured HI PDF in real- and redshift-space with the measured real-space matter PDF are displayed in solid and dashed lines, respectively. Dotted and dot-dashed lines plot the bias function when using the fully-predictive matter PDF and the measured HI PDF in real- and redshift-space respectively. Note the axes are swapped compared to the scatter plots (Fig. 3.4). Middle and lower panel: Residuals between the parametrisation-independent bias functions for measured dark matter and measured HI (mm) and fully-predictive matter distribution together with measured HI (pm) in real-space (middle) and redshift-space (lower panel).

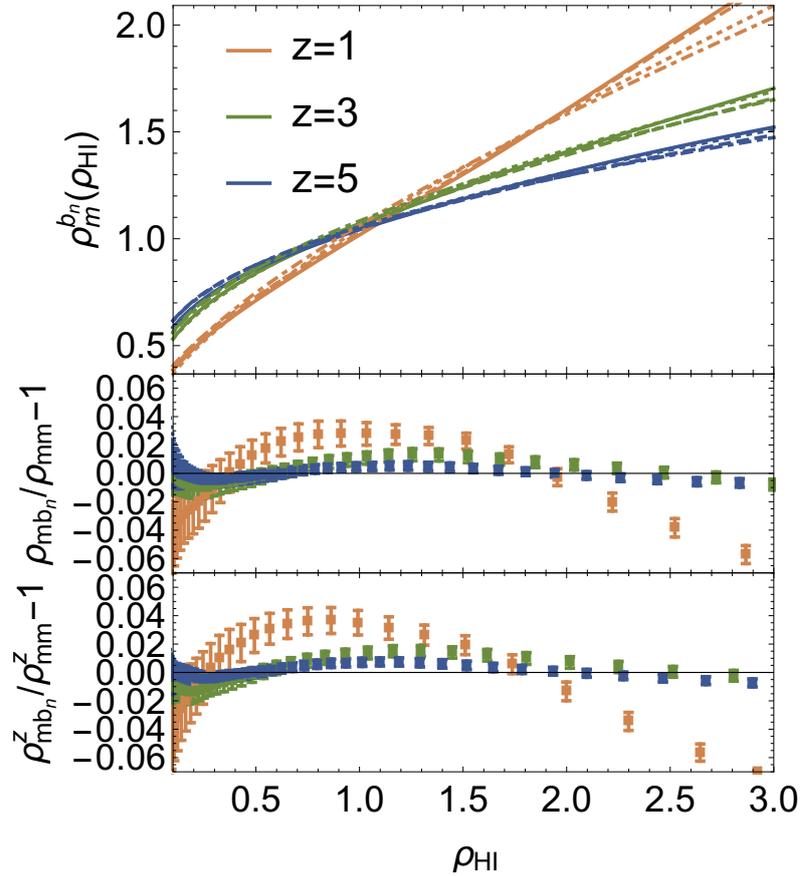


Fig. 3.8 Accuracy of quadratic parametrisation for the bias function. Upper panel: All bias functions plot matter densities as a function of the neutral hydrogen densities in spheres of radius $R = 5 h^{-1} \text{Mpc}$ at redshifts $z = 1, 3, 5$ (orange to blue). The quadratic bias model (3.13) with the fitted bias parameters from Table 3.3 in real-space (dotted lines) and redshift-space (dot-dashed lines) is compared to the measured bias function from Fig. 3.7 in real-space (solid lines) and redshift-space (dashed lines). Note that the axes are swapped compared to the scatter plots (Fig. 3.4). Middle and lower panel: Residuals between quadratic model (mb_n) and the measured bias function (mm) from the upper panel in real-space (middle) and redshift-space (lower panel).

Table 3.3 Collection of the bias parameters obtained from fitting the quadratic bias model from equation (3.13) to the predicted bias function (measured HI and predicted matter) obtained from the CDF according to equation (3.10) both in real (three left columns) and redshift-space (three right columns). The ratio between matter and HI log-variances in Tables 3.1 and 3.2 allows us to cross-check the linear bias parameter b_1 .

z	b_0	b_1	b_2	b_0^z	b_1^z	b_2^z
1	0.0451	0.5685	0.0514	0.0607	0.5387	0.0447
3	0.0693	0.3786	0.0309	0.0808	0.3486	0.0267
5	0.0463	0.3094	0.0255	0.0549	0.2846	0.0225

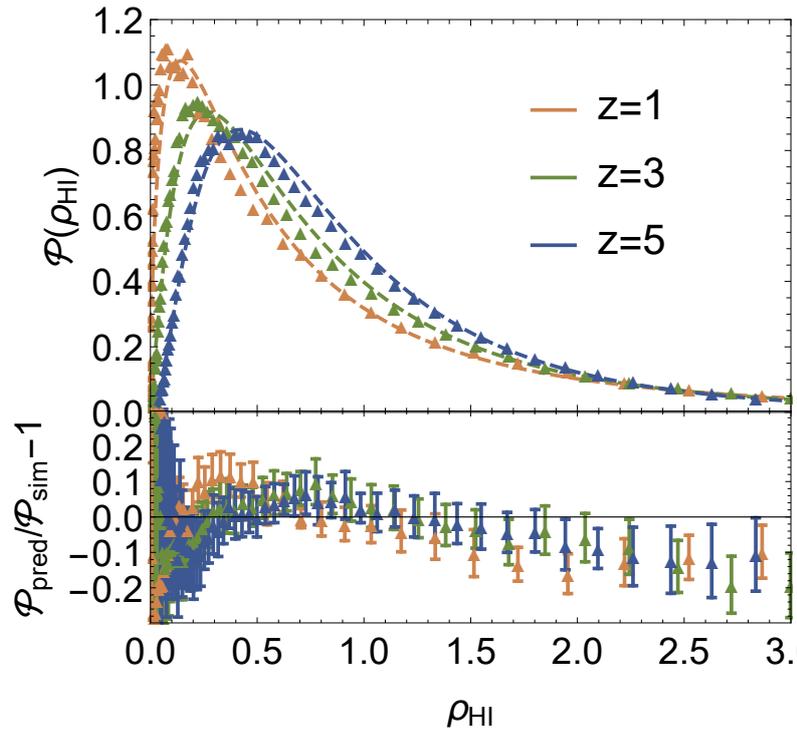


Fig. 3.9 Real-space HI PDF showing a few-percent level agreement between simulation and prediction. Upper panel: Measured PDF of HI densities in spheres of radius $R = 5 h^{-1} \text{Mpc}$ for redshifts $z = 1, 3, 5$ compared to the prediction from large-deviation statistics for matter combined with the quadratic log-bias model (dashed lines) for neutral hydrogen densities in real-space. Lower panel: Residuals between the theoretical predictions and the measured PDFs of HI densities in spheres in real-space.

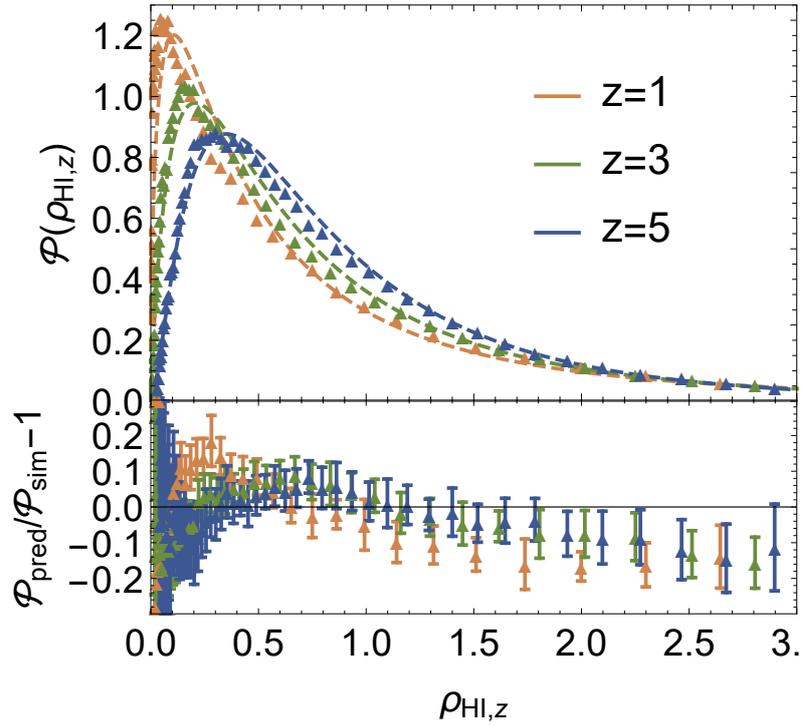


Fig. 3.10 Redshift-space HI PDF showing a few-percent level agreement between simulation and prediction. Upper panel: Measured PDF of HI densities in spheres of radius $R = 5 h^{-1}\text{Mpc}$ for redshifts $z = 1, 3, 5$ compared to the prediction from large-deviation statistics for matter combined with the quadratic log-bias model (dashed lines) for neutral hydrogen densities in redshift-space. Lower panel: Residuals between the theoretical predictions and the measured PDFs of HI densities in spheres in redshift-space.

3.5.2 One-point PDF of neutral hydrogen

Having established the accuracy of the bias model, let us now combine it with the one-point matter PDF to obtain a fully-predictive model for the one-point neutral hydrogen PDF. In this work, we fit the bias parameters at fixed cosmology to assess the accuracy at which this simple bias parametrisation combined with the predictive dark matter model can describe the neutral hydrogen PDF. In general, the framework allows for a joint fit of cosmology and bias, which could be used in a realistic application. In order to obtain an estimate for the cosmological parameters, one would then choose to marginalise over the bias parameters, given certain theoretical priors.

The results in real- and redshift-space are shown in comparison to the measurement from the IllustrisTNG simulation in Figures 3.9 and 3.10, respectively. In Fig. 3.9 we compare our fully-predictive theory for matter and the fitted bias model with the measured real-space neutral hydrogen PDF (upper panel). As the variance of the neutral hydrogen density field grows, the amplitude of the PDF tail increases, so the peak of the distribution moves towards lower densities, as they occupy more volume. The residuals (lower panel) show that the theory is able to describe the measurements over a large range in densities at a few percent level. In Fig. 3.10, we see that the redshift-space distortions can indeed be incorporated into the bias model, as the measured and predicted PDFs in redshift-space are still in good agreement (upper panel). As expected from the only slightly smaller correlation (cf. Table 3.2), we still see an agreement between theory and measurement at the few percent level, even though the residuals are slightly larger than for the real-space PDF (lower panel).

In Fig. 3.6, we also included a comparison of the PDFs of neutral hydrogen, mass-weighted halos and matter to the log-normal model [113] (dashed lines). Our results at $z = 3$ are representative, as the behaviour at other redshifts is very similar. Furthermore, the plot shows how the log-normal model fails in underdense regions, where the large-deviation statistics for matter is much more accurate. Additionally, when considering the central region of the PDF, the log-normal model typically leads to residuals of order 10% for dark matter [118, see Figure D1] and about 20% for mass-weighted halos [159, see Figure B1], both at low redshifts. At higher redshifts, we find residuals of 15-20% both for matter and tracers.

3.5.3 Density-dependent clustering of neutral hydrogen

The density-dependent clustering signal for biased tracers is interesting, because it offers to break the degeneracy between the non-linear matter variance and linear bias in the one-point tracer PDF, as demonstrated in [159] for the case of halos. Despite the aforementioned limitations of our small-box clustering measurements in IllustrisTNG, the density-dependence

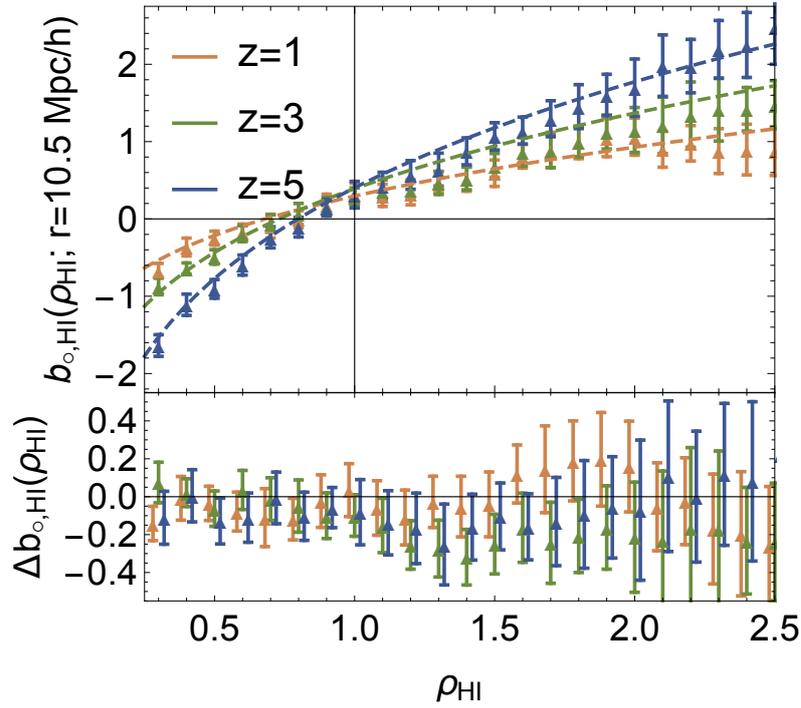


Fig. 3.11 Measured density-dependent HI clustering agrees with theoretical expectations. Upper panel: Density-dependent clustering encoded in the sphere bias at redshifts $z = 1, 3, 5$ (orange to blue) for neutral hydrogen in real-space as predicted from equation (3.7) (dashed lines) and measured (data points). Lower panel: Residuals between the theoretical prediction and simulation measurements. For better visibility, the abscissas for different redshifts are slightly shifted.

of neutral hydrogen clustering in real-space displayed in Fig. 3.11 is clearly a biased version of the density-dependent matter clustering shown in Fig. 3.3. The result is in line with the theoretical prediction (3.7) using the bias parameters found for the PDF and an approximation for the ratio of correlation functions $\sqrt{\xi_{o,m}/\xi_{o,t}} \simeq \exp(b_0)b_1$ based on a first order expansion of the log-bias model (3.13). The density-dependence of neutral hydrogen clustering is shallower than for dark matter, mainly due to linear bias, but also changes shape due to the non-linear bias term.

3.6 Conclusions and Outlook

Summary. Building on recent ideas from large-deviation statistics, an accurate theoretical model for counts-in-cells statistics of neutral hydrogen is described. The idea is to rely on analytical predictions for matter and relate them to tracers using a mean bias relation. When combining the analytical results for matter with a non-linear variance from HALOFIT,

one obtains a fully-predictive matter PDF that makes it possible to extract a non-parametric bias function from the neutral hydrogen PDF. Based on measurements in the hydrodynamic simulation IllustrisTNG, we determine the relation between matter and neutral hydrogen densities in spheres of $R = 5 h^{-1} \text{Mpc}$ from redshift $z = 5$ down to $z = 1$. The resulting non-parametric bias relation is well-described by a bias expansion up to second order in log-densities, which is in agreement with previous results for halos that host most of the neutral hydrogen.

The main results for the neutral hydrogen PDF are displayed in Figures 3.9 and 3.10 and demonstrate the few percent-level accuracy of the combined analytical model for matter and a mean bias fit both in real- and redshift-space. In addition, we detect a density-dependent clustering signal for neutral hydrogen (Fig. 3.11) that can, in principle, be used to break the degeneracy between the linear tracer bias and the non-linear variance and jointly constrain b_1 and σ_8 .

Fundamental physics. Future intensity mapping surveys will map gigantic volumes that are ideally suited for counts-in-cells statistics that probe the rare event tails and the growth of structure sensitive to dark energy [153]. Our formalism for HI counts-in-cells can be used to probe regions of particularly low and high matter density, which contain considerable information about fundamental physics such as primordial non-Gaussianity [156] and massive neutrinos [194]. 21cm offers a unique technique to observe the three-dimensional matter density field that allows to go beyond current galaxy surveys, where clustering properties of SDSS galaxy clusters are already used to approach constraints on neutrino mass [195]. We established a very close correlation between neutral hydrogen and mass-weighted halo counts-in-cells, which in turn is expected to translate to luminosity-weighted galaxies. Hence, our results could allow us to combine neutral hydrogen and galaxy counts-in-cells in order to harness synergies between intensity mapping and redshift galaxy surveys [131, 132].

Astrophysics. Another interesting direction could be to employ the accurate analytical, beyond log-normal model for one-point statistics of dark matter [7] to probe high-redshift astrophysics. Intensity mapping can be done with lines different from the 21cm spin-flip line of neutral hydrogen, which are sensitive to different astrophysical processes [196, 197] and can probe various environments such as hotter hydrogen gas ($\text{Ly}\alpha$), ionised regions (C II) or cool dense molecular gas (CO). The idea is to use the one-point statistics of intensity fluctuations which depend on both the spatial distribution of matter or halos, and also the corresponding luminosity functions. The luminosity functions contain interesting information about the detailed astrophysical conditions within the line emitters, such as star formation rates and metallicities, that can be constrained through the measured PDF of voxel intensity which complements information from the power spectrum [198]. In this context, [162]

introduced the probability distribution of voxel intensities and demonstrated its application to CO emission finding constraints on the luminosity function of the order of 10 percent. In this study, a log-normal matter distribution has been used in combination with a linear relation between halo mass and CO luminosity. Potentially, our more accurate first-principle matter PDF could tighten this constraint and enable a joint analysis of astrophysics and cosmology. Additionally, our non-linear bias model could be used to assess the robustness of the constraints regarding the assumed linear relation between halo mass and CO luminosity.

Observational limitations. We emphasise that we are considering single-dish like observations in this chapter, where neutral hydrogen fluctuations can be directly measured in configuration space. For interferometry observations, the directly observable quantity is the Fourier transform of the intensity flux. Thus, in that case, an approach bearing closer resemblance to observations will be to consider the PDF of mode amplitudes in Fourier-space. One of the main obstacles for intensity mapping observations is the fact that the amplitude of the galactic and extragalactic foregrounds can be several orders of magnitude higher than the one of the cosmological signal. Foreground cleaning is thus of pivotal importance and usually takes advantage of the rather smooth frequency spectra of foregrounds that disentangle them from the cosmic signal which maps the distribution of structures along the line-of-sight and hence has a significant amount of structure in frequency space. Foregrounds that are constant across the sky are not expected to constitute a serious problem for counts-in-cells statistics, as they just offset the overall mean density. As the non-parametric bias relation is insensitive to a multiplication by a constant factor and the bias parameter b_0 in the parametric bias relation is designed to capture this effect, the functional form of the bias relation is robust against such an offset. Foregrounds that are spatially varying on scales comparable to the size of the cells would add an extra foreground density fluctuation in every cell, roughly corresponding to extra scatter of the observed intensity around the true density. Importantly, the high correlation between the matter and HI fields is unaffected by foregrounds which simply add an additional, uncorrelated source of scatter. Our results demonstrate that neutral hydrogen counts-in-cells are rather robust against scatter in the underlying bias relation. Hence, one would expect that it is most important to model the effect of foregrounds on the mean relation, for which our formalism provides the basis. While it is beyond the scope of the present work, one could quantitatively assess the impact of foregrounds, as well as the window function and calibration on counts-in-cells statistics through mocks built for 21 cm intensity mapping experiments [199, 161].

Chapter 4

Computations for projected two- and three-point statistics¹

Summary

CMB-lensing, photometric galaxy surveys and, in general, cross-correlations between CMB-lensing and galaxy clustering are described by projected spectra. This motivates the study of projected power- and bispectra. However, projected spectra bring some technical difficulties with it: First, the number of projected bispectrum configurations grows cubic in the number of projection functions used. This means that for a galaxy survey with more than a few tomographic bins, it is intractable to directly compute the full (cross-) bispectrum signal vector. Second, the amount of resolved radial information is determined (roughly) by the inverse depth of the projection kernels and not the cut-off scale of perpendicular modes k_{\max}^{2D} . This causes problems when comparing 2D analyses with different projection depths. We illustrate the latter effect with a signal-to-noise analysis of the projected power spectrum. We explore theoretical uncertainties as a solution to this problem but find this route is only feasible for the power spectrum. For the power spectrum we show empirically that one can recover the full 3D Fisher information in tomographic surveys with sufficiently small projection depths (but fixed volume). For instance, using a projection depth of $l = 10 h^{-1} \text{Mpc}$ allows us to recover 99% of the 3D information. We introduce a solution to the first challenge by implementing a code that allows us to project the large number of projected bispectrum configuration efficiently. The main idea is to separate the projection integrals and use caching for reoccurring sub-results. We end this chapter by some remarks on the intrinsic challenges related to the large number of projected bispectra configurations.

¹This chapter based on [2].

Outline

This chapter is structured as follows: In section 4.1, we derive expressions for the projected power spectrum and bispectrum. These projection integrals are the 1D analogue to the 3D filtering discussed in section 2.4.2. In section 4.2, we study the statistical properties of the estimators for the projected spectra and illustrate the issue of finding a consistent perpendicular cut-off scale when comparing different projected spectra. Next, we review theoretical uncertainties and propose an efficient way to project them consistently in section 4.3. We turn our attention to the projected bispectrum and implement an efficient way to separate the required integrals in section 4.4. We briefly comment how one could further improve the algorithm in section 4.5.

4.1 Projected power- and bispectra

In this section, we derive in the flat sky approximation the effect of projections on 3D density fields. This allows us to express the 2D power spectrum and bispectrum in terms of the 3D theory predictions.

4.1.1 Projections

The density contrast can be projected along the LOS by means of a window function², W_X , where X characterizes the window. In real space, this reads

$$\delta_X(\mathbf{x}_\perp) = \int dx_\parallel W_X(x_\parallel) \delta(x_\parallel, \mathbf{x}_\perp) = \int \frac{d^3 \mathbf{p}}{(2\pi)^3} dx_\parallel e^{-i\mathbf{p}\cdot\mathbf{x}} W_X(x_\parallel) \delta(\mathbf{p}). \quad (4.1)$$

Hence we obtain for the Fourier transform

$$\begin{aligned} \delta_X(\mathbf{k}_\perp) &= \int \frac{dp_\parallel d^2 \mathbf{p}_\perp}{(2\pi)^3} dx_\parallel d^2 \mathbf{x}_\perp W_X(x_\parallel) \delta(\mathbf{p}) e^{-i\mathbf{p}\cdot\mathbf{x}} e^{i\mathbf{k}_\perp \cdot \mathbf{x}_\perp} \\ &= \int dx_\parallel \frac{dp_\parallel}{2\pi} d^2 \mathbf{p}_\perp \delta^D(\mathbf{p}_\perp - \mathbf{k}_\perp) e^{-ip_\parallel x_\parallel} W_X(x_\parallel) \delta(\mathbf{p}) \\ &= \int \frac{dp_\parallel}{2\pi} W_X(-p_\parallel) \delta(\mathbf{k}_\perp, p_\parallel). \end{aligned} \quad (4.2)$$

Throughout this dissertation, we separate the time dependency from the density contrast and include it into the window function; this is motivated by the perturbative expansion

²Throughout this thesis we use the terms window function, projection kernel and selection function interchangeably.

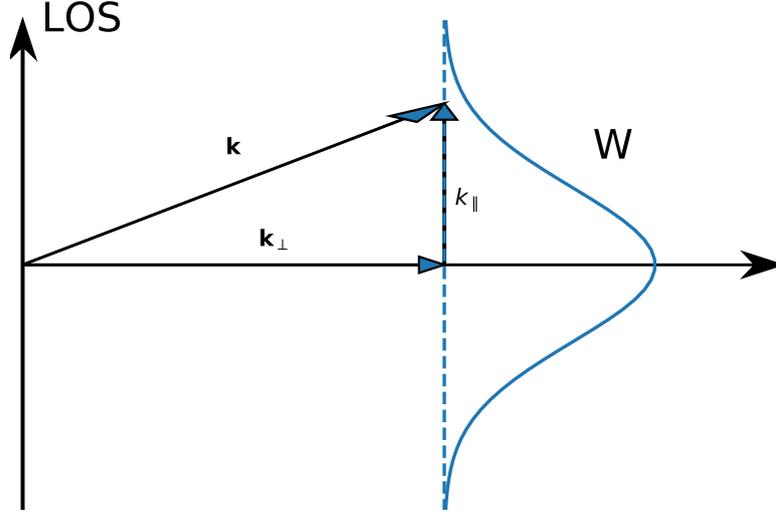


Fig. 4.1 The projected density field at each wave vector \mathbf{k}_\perp is a superposition of all wave vectors with perpendicular component \mathbf{k}_\perp and the weights set by a window function that depends on the LOS component k_\parallel of the wave vector.

of the density contrast (2.88). Fig. 4.1 illustrates the projection integral (4.2). For each perpendicular wave vector, \mathbf{k}_\perp , the projected field is a linear combination of all 3D modes with perpendicular component \mathbf{k}_\perp weighted by $W_X(-k_\parallel)$. Importantly, each 3D mode contributes to exactly one mode in the projected field. Thus, at linear order, the projected field is an isotropic and homogeneous Gaussian random field too. Those properties can be formally derived by swapping the order of the projections and averages when relating 2D to 3D correlators.

In this work, we model galaxy bins via 1D Top-Hat and Gaussian projection functions which read in Fourier space

$$W_{\text{TH}}(k_\parallel) = \text{sinc}(k_\parallel l/2) \quad (4.3a)$$

$$W_{\text{G}}(k_\parallel) = \exp[-(k_\parallel l/2)^2/2]. \quad (4.3b)$$

Here, l refers to the projection depth and we assume that the bins are thin enough so the time evolution can be assumed as constant. The CMB-lensing kernel, in contrast, is so broad that one needs to take the time evolution into account. This leads to a separate window function for each perturbative order. The n -th order CMB-lensing kernel is given by

$$W_{\kappa,n}(k_\parallel) = \frac{3}{2} \left(\frac{H_0}{c} \right)^2 \Omega_{\text{m},0} \int_0^{\chi_s} dx_\parallel e^{ik_\parallel x_\parallel} \frac{x_\parallel (\chi_s - x_\parallel)}{\chi_s} \frac{D^n(x_\parallel)}{a(x_\parallel)} \quad (4.4)$$

where D is the linear growth factor and χ_s the co-moving distance of the surface of last scattering. Assuming an EdS universe ($D = a$), allows us to analytically express the linear CMB-lensing window functions as

$$W_{\kappa,1}^{\text{EdS}}(k_{\parallel}) = \frac{3}{2} \left(\frac{H_0}{c} \right)^2 \Omega_{\text{m},0} \frac{\chi_s}{k_{\parallel}} j_1(k_{\parallel} \chi_s / 2) \exp[ik_{\parallel} \chi_s / 2]. \quad (4.5)$$

Here j_1 is a spherical Bessel function of the first kind. We used the fact that centring the window around χ and not the origin adds an additional phase factor

$$W_{\chi}(k_{\parallel}) = \int dx_{\parallel} \exp[ik_{\parallel} x_{\parallel}] W(x_{\parallel} - \chi) = \exp[i\chi k_{\parallel}] \int dx_{\parallel} \exp[ik_{\parallel} x_{\parallel}] W(x_{\parallel}) = \exp[ik_{\parallel} \chi] W_0(k_{\parallel}). \quad (4.6)$$

4.1.2 Projected power spectrum

The 3D power spectrum is defined as (2.53)

$$\langle \delta(\mathbf{k}) \delta(\mathbf{k}') \rangle = \delta^D(\mathbf{k} + \mathbf{k}') (2\pi)^3 P(k).$$

Changing the order of projection and average, then allows us to express the 2D power spectrum in terms of the 3D

$$\begin{aligned} \langle \delta_X(\mathbf{k}_{1,\perp}) \delta_Y(\mathbf{k}_{2,\perp}) \rangle &= \left\langle \int \frac{dk_{1,\parallel}}{2\pi} W_X(-k_{1,\parallel}) \delta(\mathbf{k}_{1,\perp}, k_{1,\parallel}) \int \frac{dk_{2,\parallel}}{2\pi} W_Y(-k_{2,\parallel}) \delta(\mathbf{k}_{2,\perp}, k_{2,\parallel}) \right\rangle \\ &= \int \frac{dk_{1,\parallel}}{2\pi} \frac{dk_{2,\parallel}}{2\pi} (2\pi)^3 P\left(\sqrt{\mathbf{k}_{1,\perp}^2 + k_{1,\parallel}^2}\right) \delta^D(\mathbf{k}_1 + \mathbf{k}_2) W_X(-k_{1,\parallel}) W_Y(-k_{2,\parallel}) \\ &= (2\pi)^2 \delta^D(\mathbf{k}_{1,\perp} + \mathbf{k}_{2,\perp}) P_{XY}^{2D}(\mathbf{k}_{1,\perp}), \end{aligned} \quad (4.7)$$

where we identified the 2D power spectrum as

$$P_{XY}^{2D}(\mathbf{k}_{\perp}) = \int_{-\infty}^{\infty} \frac{dk_{\parallel}}{2\pi} P\left(\sqrt{\mathbf{k}_{\perp}^2 + k_{\parallel}^2}\right) W_X(-k_{\parallel}) W_Y(k_{\parallel}). \quad (4.8)$$

We see that the projections source small scale information to larger scales. In case both window functions, centred around χ_X and χ_Y , are symmetric (4.7) can be simplified to

$$P_{XY}^{2D}(\mathbf{k}_{\perp}) = 2 \int_0^{\infty} \frac{dk_{\parallel}}{2\pi} P\left(\sqrt{\mathbf{k}_{\perp}^2 + k_{\parallel}^2}\right) \cos[(\chi_X - \chi_Y)k_{\parallel}] W_X(-k_{\parallel}) W_Y(k_{\parallel}). \quad (4.9)$$

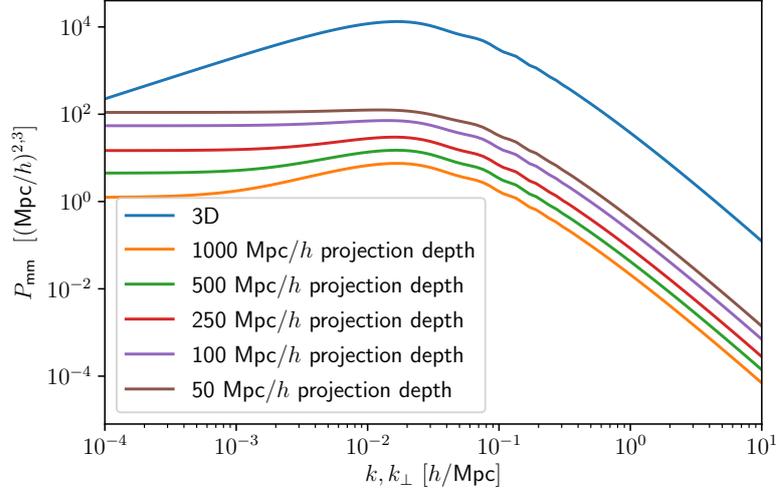


Fig. 4.2 Linear power spectrum (blue curve) and its 2D projections with a Gaussian kernel of varying depth (orange to brown). The projections change both the overall amplitude and the power spectrum's shape at low k . Instead of the k^{n_s} scaling at low k , the projected power spectra become constant.

The impact of the projection depth on the linear galaxy auto-power spectrum (i.e. $\chi_1 = \chi_2$) is illustrated in Fig. 4.2. When comparing projected power spectra (with Gaussian profile) to the linear power spectrum in 3D, projected power spectra differ by an overall factor that comes from the varying volume of the window functions in Fourier space. Since projections only source power from smaller to larger scales, the projections lead to an enhancement on large scales where the 3D power spectrum is increasing, i.e. for wave number below the wave number corresponding to matter-radiation equality and the peak of the 3D power spectrum. Moreover, the strength of the effect is increasing with decreasing projection depth i.e. wider projection kernels. In the large wave number limit, the integrand of the projection integral (4.8) becomes independent of the power spectrum and $P^{2D}(k_\perp) \propto P(k)$.

4.1.3 Projected bispectrum

Expressing the projected bispectrum

$$\langle \delta_{X_1}(\mathbf{k}_{1,\perp}) \delta_{X_2}(\mathbf{k}_{2,\perp}) \delta_{X_3}(\mathbf{k}_{3,\perp}) \rangle = (2\pi)^2 \delta^D(\mathbf{k}_{1,\perp} + \mathbf{k}_{2,\perp} + \mathbf{k}_{3,\perp}) B_{X_1 X_2 X_3}^{2D}(\mathbf{k}_{1,\perp}, \mathbf{k}_{2,\perp}, \mathbf{k}_{3,\perp}) \quad (4.10)$$

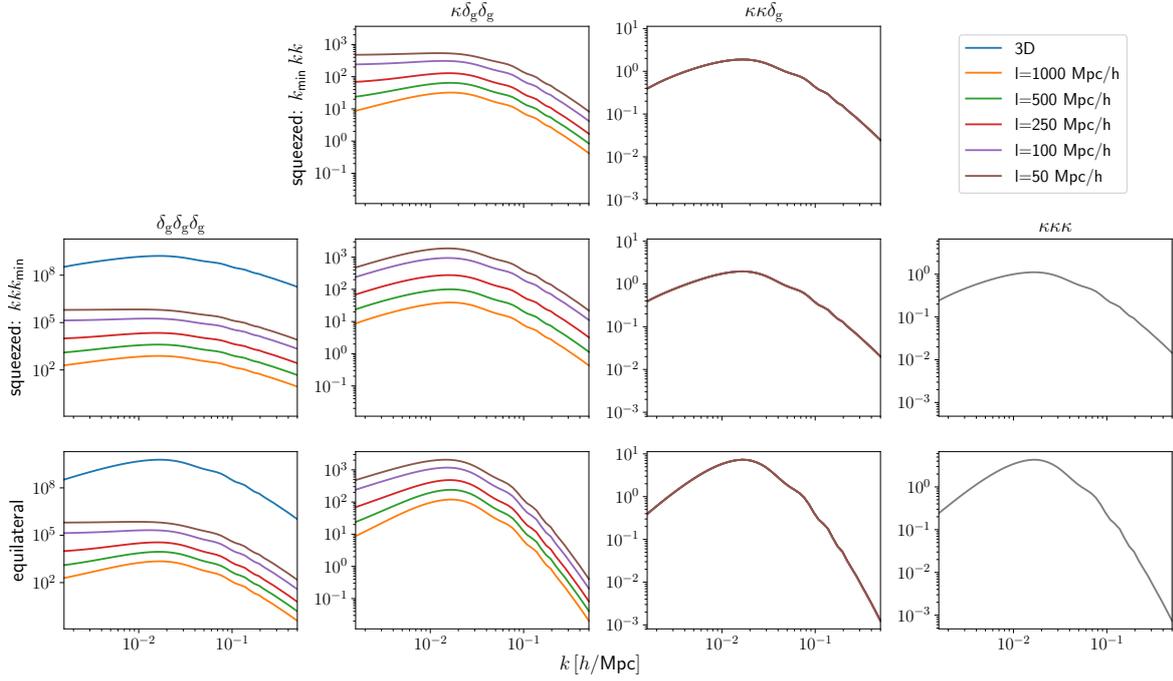


Fig. 4.3 The four columns show the four different (cross-)bispectra in the squeezed configuration (first two rows) and the equilateral configuration (bottom row). Where possible, we show the 3D bispectrum (blue) and projections with a Gaussian kernel of different depths (orange to brown). The lensing-lensing-clustering bispectra all lie onto each other.

in terms of the 3D fields by changing the order of the average and projection, yields

$$B_{X_1 X_2 X_3}^{2D}(\mathbf{k}_{1,\perp}, \mathbf{k}_{2,\perp}, \mathbf{k}_{3,\perp}) = \int \frac{dk_{1,\parallel}}{2\pi} \frac{dk_{2,\parallel}}{2\pi} dk_{3,\parallel} \delta^D \left(\sum_i k_{i,\parallel} \right) \quad (4.11)$$

$$W_{X_1}(k_{1,\parallel}) W_{X_2}(k_{2,\parallel}) W_{X_3}(k_{3,\parallel}) B \left(\sqrt{\mathbf{k}_{1,\perp}^2 + k_{1,\parallel}^2}, \sqrt{\mathbf{k}_{2,\perp}^2 + k_{2,\parallel}^2}, \sqrt{\mathbf{k}_{3,\perp}^2 + k_{3,\parallel}^2} \right).$$

Fig. 4.3 compares different projected bispectra for equilateral and squeezed configurations and a range of projection functions. κ refers to the CMB-lensing convergence and δ_g to Gaussian galaxy bins centred around $z = 0.57$ with projection depths ranging from $1000 h^{-1} \text{Mpc}$ to $50 h^{-1} \text{Mpc}$. We also show the 3D galaxy bispectrum and one sees that the projections have the strongest effect on largest scales.

When enumerating all possible projected bispectra, we stick to the convention that

- $k_1 \geq k_2 \geq k_3$
- (n_1, n_2, n_3) need to satisfy $n_i < n_j$ when $k_i = k_j$

where the n_i label the window functions involved.

4.2 Statistics

Having established theory predictions for the projected spectra, we turn to the task of estimating them. To this end, we turn the definitions of the power spectrum and bispectrum into estimators, $\hat{\mathcal{E}}$, by interpreting the averages as angular averages. This is possible thanks to the isotropy of the random field. We then characterize these estimators by calculating their mean and (co-)variance

$$C_{ij} = \langle \hat{\mathcal{E}}_i \hat{\mathcal{E}}_j \rangle - \langle \hat{\mathcal{E}}_i \rangle \langle \hat{\mathcal{E}}_j \rangle. \quad (4.12)$$

The following derivations of the covariance are carried out at leading order and assume that the underlying field is Gaussian. The 3D calculations in this subsection are closely oriented to section III of [200]. The 2D power spectrum, together with the signal-to-noise analysis, is our own work and the 2D bispectrum calculations can be found in [201].

4.2.1 Power spectrum

3D Power spectrum

Rearranging the power spectrum definition (2.53)

$$P(k) = \frac{1}{(2\pi)^3 \delta^D(\mathbf{k} + \mathbf{k}')} \langle \delta(\mathbf{k}) \delta(\mathbf{k}') \rangle \quad (4.13)$$

motivates the following estimator

$$\hat{P}(k) = \frac{k_f^3}{(2\pi)^3 V_P^{3D}(k)} \int_k d^3 \mathbf{p}_1 \int_k d^3 \mathbf{p}_2 \delta^D(\mathbf{p}_{1:2}) \delta(\mathbf{p}_1) \delta(\mathbf{p}_2). \quad (4.14)$$

Here we used the convention that $p_{i:j} = p_i + \dots + p_j$ and the volume, V_P^{3D} , of the k -bin, denoted by k , is

$$V_P^{3D}(k) = \int_k d^3 \mathbf{p} = \int_{k-\Delta k/2}^{k+\Delta k/2} p^2 dp \int d\Omega = 4\pi \Delta k \left(k^2 + \frac{\Delta k^2}{12} \right) \simeq 4\pi k^2 \Delta k. \quad (4.15)$$

The size of the survey sets a lower limit, called fundamental frequency k_f , on the radial width Δk of the k -bin. The fundamental frequency is determined via the size of the survey and for a cubic survey of volume V given by $k_f = 2\pi/V^{1/3}$. The Dirac delta in the power spectrum definition can be identified as a volume factor since

$$\delta^D(\mathbf{k}) = \int_V \frac{d^3 \mathbf{x}}{(2\pi)^3} \exp[i\mathbf{k} \cdot \mathbf{x}] \implies \delta^D(0) = \frac{V}{(2\pi)^3} = 1/k_f^3. \quad (4.16)$$

By construction, the estimator is unbiased:

$$\begin{aligned}
\langle \hat{P}(k) \rangle &= \frac{k_f^3}{(2\pi)^3 V_{\mathbf{P}}^{3D}(k)} \int_k d^3 \mathbf{p}_1 \int_k d^3 \mathbf{p}_2 \delta^D(\mathbf{p}_{12}) \langle \delta(\mathbf{p}_1) \delta(\mathbf{p}_2) \rangle \\
&= \frac{k_f^3}{V_{\mathbf{P}}^{3D}(k)} \int_k d^3 \mathbf{p}_1 \int_k d^3 \mathbf{p}_2 \delta^D(\mathbf{p}_{12})^2 P(p_1) \\
&= \frac{k_f^3}{V_{\mathbf{P}}^{3D}(k)} \int_k d^3 \mathbf{p} P(p) \delta^D(0) = P(k).
\end{aligned} \tag{4.17}$$

In this calculation, we assumed that the power spectrum is approximately constant over the k -bin. For large k -bins, one needs to improve the approximation of the integral. Doing so, leaves the estimator unbiased since one observes in a finite survey the same averaged signal.

In order to obtain the covariance (4.12) of this estimator, one needs to compute the interaction term

$$\begin{aligned}
\langle \hat{P}(k_i) \hat{P}(k_j) \rangle &= \frac{k_f^6}{(2\pi)^6 V_{\mathbf{P}}^{3D}(k_i) V_{\mathbf{P}}^{3D}(k_j)} \int_{k_i} d^3 \mathbf{p}_1 \int_{k_i} d^3 \mathbf{p}_2 \int_{k_j} d^3 \mathbf{p}_3 \int_{k_j} d^3 \mathbf{p}_4 \delta^D(\mathbf{p}_{12}) \delta^D(\mathbf{p}_{34}) \\
&\quad \times \langle \delta(\mathbf{p}_1) \delta(\mathbf{p}_2) \delta(\mathbf{p}_3) \delta(\mathbf{p}_4) \rangle \\
&= \frac{k_f^6}{(2\pi)^6 V_{\mathbf{P}}^{3D}(k_i) V_{\mathbf{P}}^{3D}(k_j)} \int_{k_i} d^3 \mathbf{p}_1 \int_{k_i} d^3 \mathbf{p}_2 \int_{k_j} d^3 \mathbf{p}_3 \int_{k_j} d^3 \mathbf{p}_4 \delta^D(\mathbf{p}_{12}) \delta^D(\mathbf{p}_{34}) \times \\
&\quad \times (\langle \delta(\mathbf{p}_1) \delta(\mathbf{p}_2) \rangle \langle \delta(\mathbf{p}_3) \delta(\mathbf{p}_4) \rangle + \langle \delta(\mathbf{p}_1) \delta(\mathbf{p}_3) \rangle \langle \delta(\mathbf{p}_2) \delta(\mathbf{p}_4) \rangle + \langle \delta(\mathbf{p}_1) \delta(\mathbf{p}_4) \rangle \langle \delta(\mathbf{p}_2) \delta(\mathbf{p}_3) \rangle) \\
&= I_1 + 2\delta_{ij}^K I_2.
\end{aligned} \tag{4.18}$$

Using Wick's theorem, we split the four-point correlator into three terms. I_1 refers to the integral of the first contraction and I_2 refers to those of the second and third. By means of relabelling, it can be seen that the third contraction is indeed the same as the second. The Dirac delta from the power spectrum definition ensures that the second and third contractions only contribute when $i = j$. The first integral is evaluated as

$$\begin{aligned}
I_1 &= \left(\frac{k_f^3}{V_{\mathbf{P}}^{3D}(k_i)} \int_{k_i} d^3 \mathbf{p}_1 \int_{k_i} d^3 \mathbf{p}_2 \delta^D(\mathbf{p}_{12})^2 P(p_1) \right) \times \\
&\quad \times \left(\frac{k_f^3}{V_{\mathbf{P}}^{3D}(k_j)} \int_{k_j} d^3 \mathbf{p}_3 \int_{k_j} d^3 \mathbf{p}_4 \delta^D(\mathbf{p}_{34})^2 P(p_3) \right) \\
&= \langle \hat{P}(k_i) \rangle \langle \hat{P}(k_j) \rangle.
\end{aligned} \tag{4.19}$$

This term cancels exactly the second term in the covariance expression (4.12). The second integral is given by

$$\begin{aligned}
I_2 &= \left(\frac{k_f^3}{(2\pi)^3 V_P^{3D}(k_i)} \right)^2 \prod_{j=1}^4 \left(\int_{k_i} d^3 \mathbf{p}_j \right) \delta^D(\mathbf{p}_{12}) \delta^D(\mathbf{p}_{34}) \langle \delta(\mathbf{p}_1) \delta(\mathbf{p}_3) \rangle \langle \delta(\mathbf{p}_2) \delta(\mathbf{p}_4) \rangle \\
&= \left(\frac{k_f^3}{V_P^{3D}(k_i)} \right)^2 \prod_{j=1}^4 \left(\int_{k_i} d^3 \mathbf{p}_j \right) \delta^D(\mathbf{p}_{12}) \delta^D(\mathbf{p}_{34}) \delta^D(\mathbf{p}_{13}) \delta^D(\mathbf{p}_{24}) P(p_1) P(p_2) \\
&= \left(\frac{k_f^3}{V_P^{3D}(k_i)} \right)^2 \int_{k_i} d^3 \mathbf{p}_1 \int_{k_i} d^3 \mathbf{p}_3 \delta^D(\mathbf{p}_{13}) \delta^D(\mathbf{p}_{13}) P^2(p_1) \\
&= \frac{k_f^3}{V_P^{3D}(k_i)} \frac{1}{V_P^{3D}(k_i)} \int_{k_i} d^3 p P^2(p) \simeq \frac{k_f^3}{V_P^{3D}(k_i)} P^2(k_i).
\end{aligned} \tag{4.20}$$

In total, the Gaussian power spectrum covariance is given by

$$\begin{aligned}
\text{Cov}_G^{3D} [P(k_i), P(k_j)] &= \langle \hat{P}(k_i) \hat{P}(k_j) \rangle - \langle \hat{P}(k_i) \rangle \langle \hat{P}(k_j) \rangle = (I_1 + 2\delta_{ij}^K I_2) - I_1 \\
&= \delta_{ij}^K \frac{2k_f^3}{V_P^{3D}(k_i)} \langle \hat{P}^2(k_i) \rangle \simeq \delta_{ij}^K \frac{2k_f^3}{V_P^{3D}(k_i)} P^2(k_i).
\end{aligned} \tag{4.21}$$

2D power spectrum

Analogously to the 3D power spectrum, the estimator for the 2D power spectrum is given by

$$\hat{P}_{XY}^{2D}(\mathbf{k}_\perp) = \frac{k_f^2}{(2\pi)^2 V_P^{2D}(\mathbf{k}_\perp)} \int_{k_\perp} d^2 \mathbf{p}_{1,\perp} \int_{k_\perp} d^2 \mathbf{p}_{2,\perp} \delta^D(\mathbf{p}_{1:2}) \delta_{X_1}(\mathbf{p}_{1,\perp}) \delta_{X_2}(\mathbf{p}_{2,\perp}). \tag{4.22}$$

The k -bin is defined as

$$V_P^{2D}(k_\perp) = \int_{k_\perp} d^2 \mathbf{p}_\perp = \int_{k_\perp - \Delta k/2}^{k_\perp + \Delta k/2} p_\perp dp_\perp \int_0^{2\pi} d\phi = 2\pi k_\perp \Delta k, \tag{4.23}$$

and for a quadratic survey of area A , the fundamental frequency is $k_f = (2\pi)/A^{1/2}$. Using a very similar calculation as in the 3D case, one finds that the estimator is unbiased and the covariance is given by

$$\begin{aligned}
\text{Cov}_G^{2D} [P_{X_1 X_2}^{2D}(k_{i,\perp}), P_{X_3 X_4}^{2D}(k_{j,\perp})] &= \\
&= \delta_{ij}^K \frac{k_f^2}{V_P^{2D}} \left[P_{X_1 X_3}^{2D}(k_{i,\perp}) P_{X_2 X_4}^{2D}(k_{i,\perp}) + P_{X_2 X_3}^{2D}(k_{i,\perp}) P_{X_1 X_4}^{2D}(k_{i,\perp}) \right].
\end{aligned} \tag{4.24}$$

Due to the window functions, the projection of the second and third contraction in (4.18) can be different which results in the two different cross-power spectra terms.

4.2.2 Power spectrum signal-to-noise analysis

Having derived the Gaussian covariance for the (projected) power spectrum, let us ignore shot-noise and compute the signal-to-noise (SN) ratios for the cases where analytical results exist.

3D power spectrum

Due to spatial homogeneity, different wave vectors are uncorrelated, so the SN of the linear power spectrum amounts to mode counting. Let us start the binning in k -space at the fundamental frequency k_f . Then there are $n = \left\lfloor \frac{k_{\max} - k_f}{\Delta k} \right\rfloor$ k -bins of size Δk up to the upper cut-off k_{\max} . For the SN ratio we find

$$\begin{aligned}
 \text{SN}_{3\text{D}} &= \sum_{ij} P(k_i) C^{-1}(k_i, k_j) P(k_j) \\
 &= \sum_i \frac{P(k_i)^2}{\frac{k_f^3}{4\pi(k_i^2 \Delta k + \Delta k^3/12)} 2P(k_i)^2} = \sum_i \frac{2\pi \Delta k}{k_f^3} \left(k_i^2 + \frac{\Delta k^2}{12} \right) \\
 &= \frac{2\pi \Delta k}{k_f^3} \left[\sum_{i=0}^{n-1} ((i+0.5) \cdot \Delta k + k_f)^2 + \frac{\Delta k^2}{12} \right] \\
 &= \frac{2\pi}{3} \left[\left(\frac{k_{\max}}{k_f} \right)^3 - 1 \right].
 \end{aligned} \tag{4.25}$$

Note that this is half the number of modes that fit into a spherical shell ranging from the lower to the upper cut-off scales of the analysis. The factor half comes from the fact that modes are only unique up to their sign.

Single projected power spectrum

In 2D, we can compute the (k_{\max}/k_f) scaling of the SN of a single auto-spectrum with a very similar calculation as for the 3D case. One finds

$$\text{SN}_{2\text{D}}(1\text{bin}) = \frac{\pi}{2} \left[\left(\frac{k_{\max}}{k_f} \right)^2 - 1 \right]. \tag{4.26}$$

As in 3D, this is the number of modes that fit into a hollow circle with limits given by the lower and upper analysis cut-offs. Cross-power spectra P_{XY} have a different covariance structure

$$C_{XY,XY}(k_i, k_j) \propto \delta_{ij}^K (P_{XY}^2(k_i) + P_{XX}(k_i)P_{YY}(k_i)) \quad (4.27)$$

that does not allow for analytic results.

Multiple projected power spectra

As different projected power spectra are correlated, their individual SNs do not simply add up. Here, we will see that including the appropriate cross-power spectra P_{XY} into the analysis of the corresponding auto-power spectra effectively removes the correlation between the auto-spectra. We demonstrate this explicitly for the case of two auto-spectra and the general case follows by induction.

Let us start with two projection windows X and Y . Due to spatial homogeneity, we know that different scales are uncorrelated, so the result from the previous subsection applies, and we only have to investigate the effect of the two bins for one particular k . The signal vector is then given by $S_P = \begin{bmatrix} P_{XX} & P_{YY} & P_{XY} \end{bmatrix}$ and the covariance is

$$C_P = \begin{bmatrix} 2P_{XX}^2 & 2P_{XY}^2 & 2P_{XX}P_{XY} \\ 2P_{XY}^2 & 2P_{YY}^2 & 2P_{YY}P_{XY} \\ 2P_{XX}P_{XY} & 2P_{YY}P_{XY} & P_{XX}P_{YY} + P_{XY}^2 \end{bmatrix}. \quad (4.28)$$

Here we dropped the mode counting prefactor $\frac{k_f^2}{2\pi k_i \Delta k}$. One sees that the off-diagonal terms prohibits a simple adding of the SN's from the different power spectra. However, the cross-power spectrum, P_{XY} , effectively decorrelates the two auto-spectra and one finds

$$\text{SN}_{2D}(2 \text{ bin}) = S_P C_P^{-1} S_P = 1 \equiv 2 \text{SN}_{2D}(1 \text{ bin}). \quad (4.29)$$

Combining this with result from the previous section we find by induction

$$\text{SN}_{2D}(\text{full analysis}) = \# \text{ auto-spectra} \cdot \frac{\pi}{2} \left[\left(\frac{k_{\max}}{k_f} \right)^2 - 1 \right]. \quad (4.30)$$

This result illustrates an important problem when working with projected (power) spectra: There is an implicit second cut-off scale for radial modes (roughly) given by the inverse projection depth. This is why even for a fixed survey volume and perpendicular k_{\max}^{2D} , the total SN increases linearly with the number of bins. In terms of the projected (power) spectra

(4.7), the additional radial modes enter through the broader projection kernels (in Fourier space) that correspond to thinner bins.

One can show that adding the cross-correlations, P_{XY}^{2D} to an auto-power spectrum SN analysis (e.g. P_{XX}^{2D}), does not improve the SN compared to the auto-power spectrum SN alone. Thus, assuming a Gaussian covariance, cross-correlations only contribute to the SN when both auto-power spectra are included.

Adding shot-noise would increase the relevance of cross-correlations since they do not suffer from shot-noise. In particular, cross-correlations would contribute to the SN for all spectra combinations and not only by effectively de-correlating auto-power spectra. Depending on the shot-noise amplitude, cross-correlations can be even more important for the overall SN than the auto-spectra.

4.2.3 Bispectrum

3D Bispectrum covariance

From the bispectrum's definition (2.58), we derive

$$\hat{B}(k_1, k_2, k_3) = \frac{k_f^3}{(2\pi)^3 V_B} \int_{k_1} d^3 \mathbf{p}_1 \int_{k_2} d^3 \mathbf{p}_2 \int_{k_3} d^3 \mathbf{p}_3 \delta^D(\mathbf{p}_{123}) \delta(\mathbf{p}_1) \delta(\mathbf{p}_2) \delta(\mathbf{p}_3) \quad (4.31)$$

as our bispectrum estimator. The volume of the k -bin, V_B , is given by

$$\begin{aligned} V_B^{3D} &= \int_{k_1} d^3 \mathbf{p}_1 \int_{k_2} d^3 \mathbf{p}_2 \int_{k_3} d^3 \mathbf{p}_3 \delta^D(\mathbf{p}_{123}) \\ &= \int \frac{d^3 \mathbf{x}}{(2\pi)^3} \prod_{i=1}^3 \left(\int_{k_i - \Delta k/2}^{k_i + \Delta k/2} p_i^2 dp_i \int_0^{2\pi} d\phi_i \int_0^\pi \sin \theta_i d\theta_i \exp[i(x p_i \cos \theta_i)] \right) \\ &= \int \frac{4\pi x^2 dx}{(2\pi)^3} \prod_{i=1}^3 \left(-2\pi \cdot 2 \int_{k_i - \Delta k/2}^{k_i + \Delta k/2} p_i^2 dp_i \frac{\sin(x p_i)}{x p_i} \right) \\ &= -32\pi \int_0^\infty x^2 dx \prod_{i=1}^3 \left(\int_{k_i - \Delta k/2}^{k_i + \Delta k/2} p_i^2 dp_i \sqrt{\frac{\pi}{2x p_i}} J_{1/2}(x p_i) \right) \\ &= 8\pi^2 \prod_{i=1}^3 \left(\int_{k_i - \Delta k/2}^{k_i + \Delta k/2} p_i dp_i \right) \\ &= 8\pi^2 k_1 k_2 k_3 (\Delta k)^3. \end{aligned} \quad (4.32)$$

In the derivation, we first wrote the Dirac delta in its Fourier representation and then identified $j_0 = \sin x/x$. The spherical Bessel functions in turn can be represented in terms of Bessel

functions which allows us to analytically solve the integral over x [202, Eq. (1) on page 411]. The integral over the Bessel functions is only non-zero if the vectors form a valid triangle.

By construction, the bispectrum estimator is unbiased

$$\begin{aligned} \langle \hat{B}(k_1, k_2, k_3) \rangle &= \frac{k_f^3}{(2\pi)^3 V_B} \int_{k_1} d^3 \mathbf{p}_1 \int_{k_2} d^3 \mathbf{p}_2 \int_{k_3} d^3 \mathbf{p}_3 \delta^D(\mathbf{p}_{123}) \langle \delta(\mathbf{p}_1) \delta(\mathbf{p}_2) \delta(\mathbf{p}_3) \rangle \\ &= \frac{k_f^3}{V_B} \int_{k_1} d^3 \mathbf{p}_1 \int_{k_2} d^3 \mathbf{p}_2 \int_{k_3} d^3 \mathbf{p}_3 \delta^D(\mathbf{p}_{123})^2 B(p_1, p_2, p_3) \\ &\simeq B(k_1, k_2, k_3). \end{aligned} \quad (4.33)$$

Here, we used, as in the previous subsection, the fact that one of the Dirac deltas corresponds to a volume factor and thus cancels the k_f^3 term.

The calculation of the interaction term in the covariance is more involved than for the power spectrum since it involves a six-point function that has, according to Wick's theorem, fifteen possible contractions. However, assuming that the two triangles are not degenerate, we only have to contract each density with a density of the other bispectrum/triangle because of the delta functions that come with the resulting power spectra. This means the two triangles have to be congruent. The covariance is proportional to some integral, that we are about to compute, times all possible non-zero contractions. We will denote the triplet of integers labeling a wave vector configuration by a capital character $I = (i_1, i_2, i_3)$. We can then extend the Kronecker delta to triangles via

$$\delta_{IJ}^K = \delta_{i_1 j_1}^K \delta_{i_2 j_2}^K \delta_{i_3 j_3}^K. \quad (4.34)$$

From the argument above, we can write

$$C_{IJ}^B = I_B \sum_{L \in \sigma(J)} \delta_{IL}^K = \delta_{IJ}^K s_B I_B \quad (4.35)$$

where we sum over the 6 permutations of J . One can then summarize the number of non-zero contractions with the symmetry factor s_B that is 1, 2 and 6 for scalene, isosceles and equilateral triangles, respectively. We now calculate I_B assuming without loss of generality that the contraction

$$\text{contraction} : \langle \delta(\mathbf{p}_1) \delta(\mathbf{p}_4) \rangle \langle \delta(\mathbf{p}_2) \delta(\mathbf{p}_5) \rangle \langle \delta(\mathbf{p}_3) \delta(\mathbf{p}_6) \rangle \quad (4.36)$$

is non-zero. This yields

$$\begin{aligned}
I_B &= \left(\frac{k_f^3}{(2\pi)^3 V_{B_i}} \right)^2 \int_{k_1} d^3 \mathbf{p}_1 \int_{k_2} d^3 \mathbf{p}_2 \int_{k_3} d^3 \mathbf{p}_3 \int_{k_1} d^3 \mathbf{p}_4 \int_{k_2} d^3 \mathbf{p}_5 \int_{k_3} d^3 \mathbf{p}_6 \delta^D(\mathbf{p}_{123}) \delta^D(\mathbf{p}_{456}) \times \\
&\quad \times \langle \delta(\mathbf{p}_1) \delta(\mathbf{p}_4) \rangle \langle \delta(\mathbf{p}_2) \delta(\mathbf{p}_5) \rangle \langle \delta(\mathbf{p}_3) \delta(\mathbf{p}_6) \rangle \\
&= \left(\frac{k_f^3}{(2\pi)^3 V_{B_i}} \right)^2 \int_{k_1} d^3 \mathbf{p}_1 \int_{k_2} d^3 \mathbf{p}_2 \int_{k_3} d^3 \mathbf{p}_3 \int_{k_1} d^3 \mathbf{p}_4 \int_{k_2} d^3 \mathbf{p}_5 \int_{k_3} d^3 \mathbf{p}_6 \delta^D(\mathbf{p}_{123}) \delta^D(\mathbf{p}_{456}) \times \\
&\quad \times (2\pi)^9 P(p_1) P(p_2) P(p_3) \delta^D(\mathbf{p}_{1:4}) \delta^D(\mathbf{p}_{2:5}) \delta^D(\mathbf{p}_{3:6}) \\
&= (2\pi)^3 \left(\frac{k_f^3}{V_{B_i}} \right)^2 \int_{k_1} d^3 \mathbf{p}_1 \int_{k_2} d^3 \mathbf{p}_2 \int_{k_3} d^3 \mathbf{p}_3 \delta^D(\mathbf{p}_{123})^2 P(p_1) P(p_2) P(p_3) \\
&\simeq (2\pi)^3 \frac{k_f^3}{V_{B_i}} P(k_1) P(k_2) P(k_3).
\end{aligned} \tag{4.37}$$

Hence, the Gaussian bispectrum covariance is given by

$$\begin{aligned}
\text{Cov}_G^{3D} [B^{3D}(k_{i_1}, k_{i_2}, k_{i_3}), B^{3D}(k_{j_1}, k_{j_2}, k_{j_3})] &= \delta_{IJ}^K s_B (2\pi)^3 \frac{k_f^3}{V_{B_i}^{3D}} P(k_1) P(k_2) P(k_3) \\
&= \delta_{IJ}^K \frac{s_B \pi k_f^3}{k_1 k_2 k_3 (\Delta k)^3} P(k_1) P(k_2) P(k_3).
\end{aligned} \tag{4.38}$$

As for the power spectrum, the calculations and expressions assume thin k -bins but can easily be generalized to broader k -bins. In fact, it is standard practice to reduce the number of bispectrum configurations via a coarse binning.

2D Bispectrum

In this section, we work with 2D vectors and discard the perpendicular subscript for clarity. Following the previous sections, an estimator for the 2D bispectrum is given by

$$\hat{B}_{X_1 X_2 X_3}(k_1, k_2, k_3) = \frac{k_f^2}{(2\pi)^2 V_B} \int_{k_1} d^2 \mathbf{p}_1 \int_{k_2} d^2 \mathbf{p}_2 \int_{k_3} d^2 \mathbf{p}_3 \delta(\mathbf{p}_{123}) \delta_{X_1}(\mathbf{p}_1) \delta_{X_2}(\mathbf{p}_2) \delta_{X_3}(\mathbf{p}_3) \tag{4.39}$$

where the fundamental frequency is given by $k_f^2 = (2\pi)^2/A = k_f$. The volume of the (k_1, k_2, k_3) -bin is given by

$$\begin{aligned}
V_{\mathbf{B}}^{2D} &= \int_{k_1} d^2\mathbf{p}_1 \int_{k_2} d^2\mathbf{p}_2 \int_{k_3} d^3\mathbf{p}_3 \delta^D(\mathbf{p}_{123}) = \prod_{i=1}^3 \left(\int_{k_i-\Delta k/2}^{k_i+\Delta k/2} p_i dp_i \int_0^{2\pi} d\phi_{p_i} \right) \delta^D(\mathbf{p}_{123}) \\
&= \prod_{i=1}^3 \left(\int_{k_i-\Delta k/2}^{k_i+\Delta k/2} p_i dp_i \int_0^{2\pi} d\phi_{p_i} \right) \int \frac{d^2\theta}{(2\pi)^2} e^{i(\mathbf{p}_1+\mathbf{p}_2+\mathbf{p}_3)\theta} \\
&= \prod_{i=1}^3 \left(\int_{k_i-\Delta k/2}^{k_i+\Delta k/2} p_i dp_i \right) \int \frac{d^2\theta}{(2\pi)^2} (2\pi)^3 J_0(p_1\theta) J_0(p_2\theta) J_0(p_3\theta) \\
&= \prod_{i=1}^3 \left(\int_{k_i-\Delta k/2}^{k_i+\Delta k/2} p_i dp_i \right) 2\pi \Lambda(p_1, p_2, p_3) \simeq 2\pi k_1 k_2 k_3 (\Delta k)^3 \Lambda(k_1, k_2, k_3)
\end{aligned} \tag{4.40}$$

where we used the integral representation of the Bessel function

$$J_0(x) = \int_0^{2\pi} \frac{d\phi}{2\pi} e^{ix \cos \phi} \tag{4.41}$$

and utilized [203, formula no. 6.578.9] for the angular integration which yields

$$\Lambda(k_1, k_2, k_3) = \begin{cases} \left[\frac{1}{4} \sqrt{2k_1^2 k_2^2 + 2k_1^2 k_3^2 + 2k_2^2 k_3^2 - k_1^4 - k_2^4 - k_3^4} \right]^{-1} & \text{if } |k_1 - k_2| < k_3 < k_1 + k_2 \\ 0 & \text{otherwise.} \end{cases} \tag{4.42}$$

By construction, the estimator is unbiased. Moreover, a similar calculation as in 3D shows

$$\text{Cov}_G^{2D} [B_{\mathbf{X}}^{2D}(k_{i_1}, k_{i_2}, k_{i_3}), B_{\mathbf{Y}}^{2D}(k_{j_1}, k_{j_2}, k_{j_3})] = \frac{(2\pi)k_f^2}{k_{i_1} k_{i_2} k_{i_3} (\Delta k)^3} \frac{1}{\Lambda(k_{i_1}, k_{i_2}, k_{i_3})} \text{PPP} \tag{4.43}$$

where

$$\text{PPP} = \sum_{L \in \sigma(J)} \delta_{IL}^K P_{X_{i_1} Y_{l_1}}^{2D}(k_{i_1}) P_{X_{i_2} Y_{l_2}}^{2D}(k_{i_2}) P_{X_{i_3} Y_{l_3}}^{2D}(k_{i_3}). \tag{4.44}$$

Note that the Kronecker delta in (4.44) enforces statistical homogeneity by requiring that the two triangles are congruent. This makes the covariance diagonal. As in the 3D case, k -bins are often chosen so wide that the thin-bin approximation used here does not apply.

Power-bispectrum covariance

The Gaussian covariance of power- and bispectrum is zero since it corresponds to a connected five-point function. For completeness, we give here the leading order of the power-bispectrum

covariance. Using the definitions from above yields at leading order for the covariance in 3D

$$\text{Cov}^{3\text{D}} [P(k_i), B(k_{j_1}, k_{j_2}, k_{j_3})] = 2(\delta_{i j_1}^K + \delta_{i j_2}^K + \delta_{i j_3}^K) \frac{k_f^3}{V_P^{3\text{D}}(k_i)} P(k_i) B(k_{j_1}, k_{j_2}, k_{j_3}). \quad (4.45)$$

The Kronecker delta ensures that the magnitude of the power spectrum vector corresponds to at least one leg of the bispectrum. For the covariance of the projected spectra one finds

$$\begin{aligned} & \text{Cov}^{2\text{D}} [P_{X_1 X_2}^{2\text{D}}(k_i), B_{Y_1 Y_2 Y_3}^{2\text{D}}(k_{j_1}, k_{j_2}, k_{j_3})] \\ &= \frac{k_f^2}{V_P^{2\text{D}}(k_i)} \sum_{L \in \sigma(J)} \delta_{i l_1}^K \left(P_{X_1 Y_{l_1}}(k_i) B_{X_2, Y_{l_2}, Y_{l_3}}(k_i, k_{l_2}, k_{l_3}) + P_{X_2 Y_{l_1}}(k_i) B_{X_1, Y_{l_2}, Y_{l_3}}(k_i, k_{l_2}, k_{l_3}) \right). \end{aligned} \quad (4.46)$$

4.3 Theoretical uncertainties

The (perturbative) theory predictions of the power spectrum and bispectrum gradually lose their accuracy with increasing wave vectors. Typically, one deals with that by means of a hard cut-off, k_{max} , in the analysis (see eg. section 4.2.2). The implicit assumption behind this is that the predictions are perfectly accurate up to the cut-off and (completely) inaccurate for smaller scales. A further problem arises when working with projected spectra. In section 4.2.2, we showed that perpendicular cut-offs do not directly control the amount of small scale radial information that is included in the analysis. Theoretical uncertainties offer a solution to both problems. The idea is to formally include all modes into the analysis and to weight modes according to the accuracy of their predictions.

4.3.1 Theoretical uncertainties in three dimensions

A first model for theoretical uncertainties was developed in [204]. The core idea is to add theoretical uncertainties, C_{th} , to the statistical uncertainties, C , that were discussed in the previous section

$$C \rightarrow C + C_{\text{th}}. \quad (4.47)$$

The theoretical uncertainties grow with increasing wave vectors and thus gradually decrease the weight of small scale modes until the error bars saturate. The hard cut-off in contrast assigns the most weight (due to cosmic variance) to the modes closest to the cut-off which are also the most uncertain modes of the analysis. Fig. 4.4 illustrates the effect of theoretical uncertainties on the forecasted errors (see section 5.3.1) in 3D. We compare the k_{max} dependence of the marginalized error bars without theoretical uncertainties and with theoretical

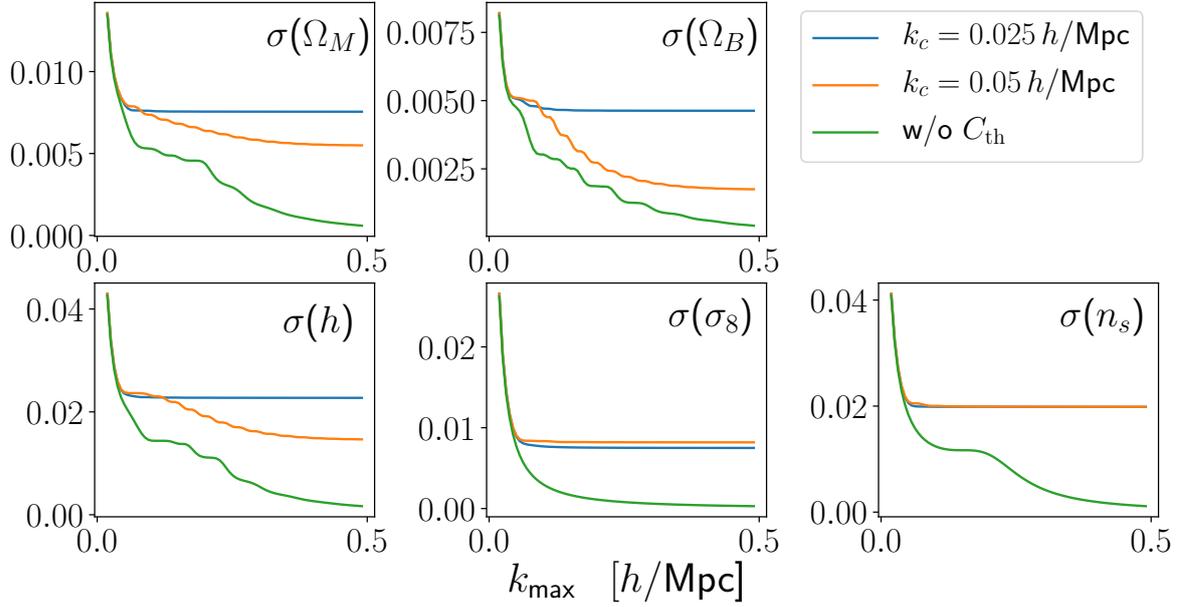


Fig. 4.4 Unmarginalized 1σ error bars without theoretical uncertainties (green line) and with theoretical uncertainties (orange and blue) as a function of the upper cut-off k_{\max} . The error bars of the cosmological parameters are decreasing with k_{\max} without theoretical uncertainties. Using theoretical uncertainties, the error bars saturate. For a smaller correlation length in the correlation function, the saturation happens at a larger scale (blue curve) than with the larger correlation length (orange). The results come from a power spectrum analysis of a cubic survey of side length $1000 h^{-1}\text{Mpc}$ at redshift $z=0.57$.

uncertainties of two different correlation lengths. For a smaller correlation length, the error bars saturate earlier.

Theoretical uncertainties, e , have mean zero and their covariance can be decomposed into two components: An amplitude, E , and a correlation function, ρ ,

$$C_{\text{th}}^{3\text{D}}(k_1, k_2) = \langle e^{3\text{D}}(k_1) e^{3\text{D}}(k_2) \rangle_c = E(k_1) \rho(k_1, k_2) E(k_2). \quad (4.48)$$

The scale of the fluctuations, E , is set by the next-order solution of the perturbative expansion. Since we are working at linear order, we adopt the following 1-loop matter power spectrum envelope introduced by [204]

$$E(k) = P(k, z) \left(\frac{D(z)}{D(0)} \right)^2 \left(\frac{k}{0.31 h \text{Mpc}^{-1}} \right)^{1.8}, \quad (4.49)$$

together with a Gaussian correlation function

$$\rho_{\text{P}}^{3\text{D}}(k_1, k_2) = \exp \left[-\frac{(k_1 - k_2)^2}{2k_c^2} \right]. \quad (4.50)$$

The (inverse) correlation length k_c is a free parameter that should reflect the scale on which variations in the spectra occur. In our setting, the relevant variations are the Baryonic Acoustic Oscillations that correspond in Fourier space to oscillations of length $k_{\text{BAO}} \simeq 0.05 h^{-1} \text{Mpc}$. We follow the choice of [204] and use an (inverse) correlation length of $k_c = k_{\text{BAO}}/2$ as this is the characteristic scale on which the power spectrum varies. For the 1-loop bispectrum envelope [204] chose

$$E(k_1, k_2, k_3) = B_{\text{mmm}}(k_1, k_2, k_3, z) \left(\frac{D(z)}{D(0)} \right)^2 \left(\frac{\bar{k}}{0.31 h \text{Mpc}^{-1}} \right)^{1.8}, \quad \bar{k} = \frac{1}{3}(k_1 + k_2 + k_3) \quad (4.51)$$

and assumed a Gaussian correlation function

$$\rho_{\text{B}}^{3\text{D}}(k_I, k_J) = \prod_{\alpha=1}^3 \exp \left[-(k_{i_\alpha} - k_{j_\alpha})^2 / 2k_c^2 \right], \quad \text{where } k_{i_1} \geq k_{i_2} \geq k_{i_3}. \quad (4.52)$$

We, again, adopt these choices and use $k_c = k_{\text{BAO}}/2$ for the correlation length.

Formal derivation

The addition of the theoretical uncertainties to the cosmic variance in (4.47) was motivated as follows [204]: At leading order, the likelihood of the power spectrum is Gaussian. Modelling the theoretical uncertainties as a mean-zero Gaussian process then yields the following joint likelihood function³

$$\begin{aligned} \mathcal{L}(\mathbf{t}, \mathbf{e}|\mathbf{d}) &= \frac{1}{(2\pi)^n \sqrt{\det C \det C_{\text{th}}}} \exp \left[-\frac{1}{2}(\mathbf{d} - \mathbf{t} - \mathbf{e})^T C^{-1} (\mathbf{d} - \mathbf{t} - \mathbf{e}) \right] \exp \left[-\frac{1}{2}\mathbf{e}^T C_{\text{th}}^{-1} \mathbf{e} \right] \\ &= \frac{1}{(2\pi)^n \sqrt{\det C \det C_{\text{th}}}} \exp \left[-\frac{1}{2}(\mathbf{d} - \mathbf{t})^T C^{-1} (\mathbf{d} - \mathbf{t}) \right] \exp \left[-\frac{1}{2}\mathbf{e}^T C_{\text{th}}^{-1} \mathbf{e} + \mathbf{e}^T C^{-1} (\mathbf{d} - \mathbf{t}) \right] \end{aligned} \quad (4.53)$$

where the data vector is \mathbf{d} , the theory predictions are \mathbf{t} and the theoretical uncertainties \mathbf{e} . The idea is to marginalize over the theoretical uncertainties to incorporate them into the error

³Strictly speaking, only the first term is the likelihood function and the second provides a prior for the theoretical uncertainties. The product is proportional to the posterior $p(\mathbf{d}, \mathbf{e}|\mathbf{t})$, so one can marginalize over the theoretical uncertainties \mathbf{e} .

budget. The marginalization is done using the Gaussian integral for symmetric matrix A

$$\int d\mathbf{e} \exp \left[-\frac{1}{2} \mathbf{e}^T A \mathbf{e} + \mathbf{e}^T B \right] = \sqrt{\frac{(2\pi)^n}{\det A}} \exp \left[\frac{1}{2} B A^{-1} B \right]. \quad (4.54)$$

Using

$$A = C_{\text{th}}^{-1}, \quad B = C_{\text{th}}^{-1}(\mathbf{d} - \mathbf{t}), \quad (4.55)$$

yields

$$\mathcal{L}(\mathbf{t}|\mathbf{d}) = \int d\mathbf{e} \mathcal{L}(\mathbf{t}, \mathbf{e}|\mathbf{d}) = \frac{1}{\sqrt{(2\pi)^n \det [C + C_{\text{th}}]}} \exp \left[-\frac{1}{2} (\mathbf{d} - \mathbf{t})^T (C + C_{\text{th}})^{-1} (\mathbf{d} - \mathbf{t}) \right]. \quad (4.56)$$

Here, we used

$$\begin{aligned} & [C^{-1} - C^{-1}(C^{-1} + C_{\text{th}}^{-1})^{-1}C^{-1}] (C + C_{\text{th}}) \\ &= 1 + C^{-1}C_{\text{th}} - C^{-1}(C^{-1} + C_{\text{th}}^{-1})^{-1} - C^{-1}(C^{-1} + C_{\text{th}}^{-1})^{-1}C^{-1}C_{\text{th}} \\ &= 1 + C^{-1}C_{\text{th}} - C^{-1}C_{\text{th}} = 1 \\ &\implies [C^{-1} - C^{-1}(C^{-1} + C_{\text{th}}^{-1})^{-1}C^{-1}] = (C + C_{\text{th}})^{-1}. \end{aligned} \quad (4.57)$$

The marginal likelihood (4.56) provides the formal justification for the modification of the covariance in (4.47).

4.3.2 Projected theoretical uncertainties

In analogy to the projected power spectrum (4.8) and bispectrum (4.11), we project the theoretical uncertainties of the power spectrum via

$$e_{XY}^{2D}(k_{\perp}) = \int \frac{dk_{\parallel}}{2\pi} W_X(-k_{\parallel}) W_Y(k_{\parallel}) e \left(\sqrt{k_{\parallel}^2 + \mathbf{k}_{\perp}^2} \right) \quad (4.58)$$

and of the bispectrum via

$$\begin{aligned} & e_{X_1 X_2 X_3}^{2D}(k_{1,\perp}, k_{2,\perp}, k_{3,\perp}) \\ &= \int \frac{dk_{1,\parallel}}{2\pi} \frac{dk_{2,\parallel}}{2\pi} dk_{3,\parallel} \delta^D \left(\sum_i k_{i,\parallel} \right) W_{X_1}(k_{1,\parallel}) W_{X_2}(k_{2,\parallel}) W_{X_3}(k_{3,\parallel}) e(k_1, k_2, k_3). \end{aligned} \quad (4.59)$$

A consistent model for the projected systematic uncertainties in 2D, C_{th}^{2D} , is achieved by swapping the average with the projection integrals in the covariance definition. For the power

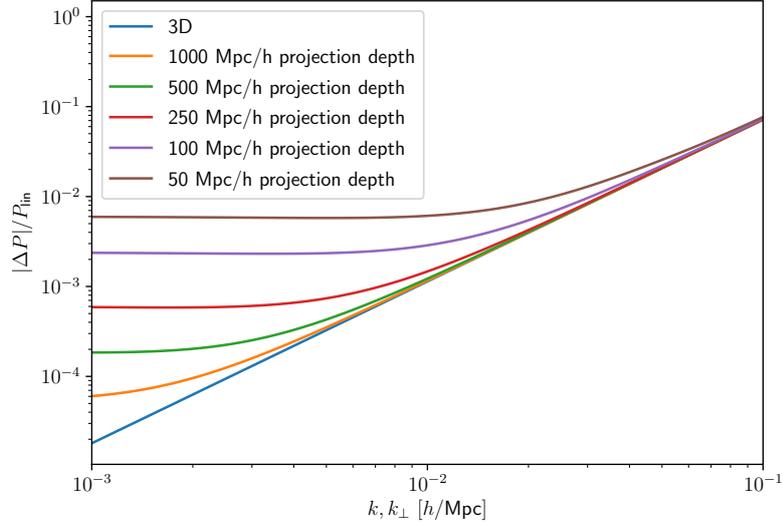


Fig. 4.5 Relative importance of the one-loop envelope to the linear power spectrum in 3D (blue curve) and for various projected power spectra (orange to brown). The ratio in 3D follows a power law, whereas the ratios of the projected power spectra become constant at low k . At large k , the ratios of the projected power spectra approach the 3D value from above.

spectrum, we obtain

$$\begin{aligned}
 C_{\text{th}}^{2\text{D}}(\mathbf{k}_{1,\perp}, \mathbf{k}_{2,\perp}) &= \langle e_{X_1 X_2}^{2\text{D}}(\mathbf{k}_{1,\perp}) e_{Y_1 Y_2}^{2\text{D}}(\mathbf{k}_{2,\perp}) \rangle \\
 &= \int \frac{dk_{1,\parallel}}{2\pi} \frac{dk_{2,\parallel}}{2\pi} W_{X_1}(k_{1,\parallel}) W_{X_2}(-k_{1,\parallel}) W_{Y_1}(k_{2,\parallel}) W_{Y_2}(-k_{2,\parallel}) C_{\text{th}}^{3\text{D}}(k_1, k_2).
 \end{aligned} \tag{4.60}$$

For generic (power spectrum) correlation functions, the resulting 2D integral does not have a closed form solution. Even worse, theoretical uncertainties yield non-sparse covariances where each independent entry corresponds to a 2D integral. This is numerically not feasible. However, in case of a Gaussian correlation functions, the integral can be separated using

$$\begin{aligned}
 \exp\left[-\frac{(k_1 - k_2)^2}{2r^2}\right] &= \frac{\sqrt{2}}{\sqrt{\pi r}} \int_{-\infty}^{\infty} dc \exp\left[-\frac{(k_1 - c)^2}{r^2}\right] \exp\left[-\frac{(k_2 - c)^2}{r^2}\right] \\
 &\simeq \frac{\sqrt{2}}{\sqrt{\pi r}} \Delta c \sum_{i=1}^N \exp\left[\frac{(k_1 - (c_{\min} + i\Delta c))^2}{r^2}\right] \exp\left[\frac{(k_2 - (c_{\min} + i\Delta c))^2}{r^2}\right],
 \end{aligned} \tag{4.61}$$

which makes the problem tractable since it enables caching. The idea is to store the result of each 1D integral labelled by (k_{\perp}, X, i) - where X characterizes the window and i characterizes the term in the above expansion - in a hash table. The computation of the covariance matrix then simply amounts to combining those cached results. In our settings, we achieve

subpercent accuracy for all configurations using

$$c_{\min} = -3.2, \quad c_{\max} = 3.2, \quad \Delta c = \frac{c_{\max} - c_{\min}}{N}, \quad N = 1500 \quad (4.62)$$

for the discretization.

The above approach to the bispectrum's theoretical uncertainties becomes computationally intractable when projecting. Due to the implicit wave vector ordering in the correlation function (4.52), the direct projection is a four-dimensional integral that cannot be separated. As we were not able to sufficiently speed up the computations, we do not use theoretical uncertainties for the projected bispectrum.

4.3.3 Empirical validation

We test the equivalence between three- and two-dimensional matter power spectrum analysis (without RSDs) empirically by performing both analyses and comparing the forecasted error bars. We control the sourcing of small-scale information to large scales with theoretical uncertainties. To this end, we use a cubic survey of side length $1000 h^{-1} \text{Mpc}$. The corresponding fundamental frequency is roughly four times smaller than the correlation length of the theoretical uncertainties which ensures that the theoretical uncertainties are approximately constant over k -bins with width of the fundamental frequency. The projections are done with Top-Hat window functions of depths $1000 h^{-1} \text{Mpc} / n$, where n is the number of bins. Our findings are summarized in Fig. 4.6. We show the ratio between the error bars in the two-dimensional setting and the values the three-dimensional analysis as a function of the 2D cut-off scale, k_{\max}^{2D} . For a given number of bins, the 2D error bars saturate, due to the projected theoretical uncertainties. In addition, those values converge to the 3D values as the number of bins increases.

4.4 Efficient bispectrum projection integrals

The number of (cross-)bispectrum configurations scales cubic both in the cut-off⁴ and the number of projection functions. Since each configuration corresponds to a 2D integral, it is not feasible to obtain the projected theory predictions for moderately large cut-offs and moderately many projection kernels. In this section, we are introducing strategies that allow us to compute those projections efficiently.

⁴The exact number of triangles as a function of the cut-off is described the OEIS sequence A002623 [205].

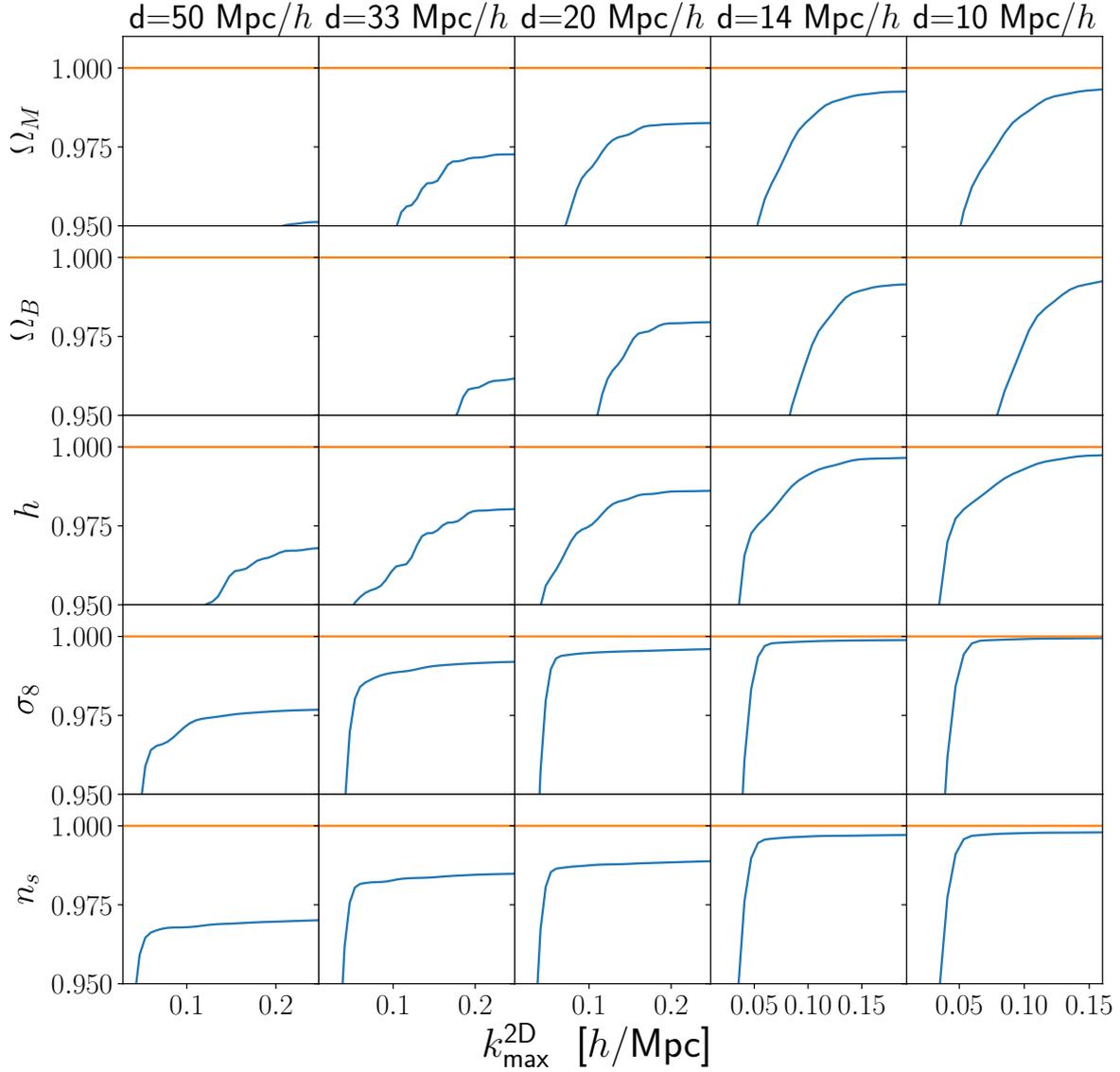


Fig. 4.6 The ratio of (unmarginalized) error bars in 3D and 2D, $\sigma_x^{3D}/\sigma_x^{2D}$, is shown as a function of k_{\max}^{2D} . We use the saturated value for the 3D uncertainty and refer to Fig. 4.4 where we studied the saturation in detail. For each parameter (row) and projection depth (column), the 2D error bar saturates due to theoretical uncertainties. As the number of tomographic bins increases (from left to right), this saturated value approaches the 3D value.

4.4.1 Separating the bispectrum

The core idea to speed up the computation of the projected bispectrum is separating the bispectrum projection integral (4.11) by using the Fourier representation of the Dirac delta

$$\begin{aligned}
& B_{X_1 X_2 X_3}^{2D}(k_{1,\perp}, k_{2,\perp}, k_{3,\perp}) \\
&= \int \frac{dk_{1,\parallel}}{2\pi} \frac{dk_{2,\parallel}}{2\pi} dk_{3,\parallel} \delta^D\left(\sum_i k_{i,\parallel}\right) W_{X_1}(k_{1,\parallel}) W_{X_2}(k_{2,\parallel}) W_{X_3}(k_{3,\parallel}) B(k_1, k_2, k_3) \quad (4.63) \\
&= \int_{-\infty}^{\infty} dx_{\parallel} \int \prod_{i=1}^3 \left[\frac{dk_{i,\parallel}}{2\pi} W_{X_i}(k_{i,\parallel}) \exp(-ik_{i,\parallel}x_{\parallel}) \right] B(k_1, k_2, k_3).
\end{aligned}$$

In this work, we only work with separable bispectra,⁵ so it suffices to discuss how the resulting four one-dimensional integrals can be solved. Let us start with the three inner integrals that are Fourier transformations of the window function times some contribution from the bispectrum. These terms can be parametrized as

$$I_W(x_{\parallel}, \mathbf{k}_{\perp}; \alpha, \beta, \gamma, \delta, \varepsilon) = \int_{-\infty}^{\infty} \frac{dk_{\parallel}}{2\pi} e^{-ik_{\parallel}x_{\parallel}} W(k_{\parallel}) k^{\alpha} P(k)^{\beta} \left(\frac{k_{\parallel}}{k}\right)^{\gamma} \exp\left[-\delta k_{\parallel}^2 \sigma_v^2 / 2\right] \mathcal{M}^{\varepsilon}(k) \quad (4.64)$$

where $k = \sqrt{k_{\parallel}^2 + \mathbf{k}_{\perp}^2}$. Let us trace back the origin and possible values of the five parameters from the theory predictions we derived in chapter 2:

- The matter bispectrum (2.105) generates terms with $\alpha \in \{-2, 0, 2, 4\}$, $\beta \in \{0, 1\}$ and $\gamma = \delta = \varepsilon = 0$. The α 's come from the cosine, $\mu(k_1, k_2) = \frac{k_3^2 - k_1^2 - k_2^2}{2k_1 k_2}$ in the gravitational kernel F_2 . The term $\alpha = -2$ only occurs together with $\beta = 1$. The terms with $\beta = 0$ and $\alpha > 0$ are numerically challenging for slowly decaying window function since the integrand formally diverges. This is in particular an issue for the lensing kernel. In the next section, we present analytic results for those integrals.
- Linear RSDs generate terms with $\gamma = 2$ that come from the cosine of the wave vector with the LOS in the redshift-space kernels, Z . The multiplicative prefactor that captures non-linear redshift-space distortions correspond to $\delta = 1$.
- The PNG templates (see section 2.2.4) introduce terms with Poisson factor (2.67), \mathcal{M} , where $\varepsilon = 1 - 2\beta$. In those integrals, $\beta \in \{0, 1/3, 2/3, 1\}$. The scale-dependent bias terms (see section 2.4.5) corresponds to terms with $\alpha \in \{0, 1, 2\}$ and $\beta = 1 \Rightarrow \varepsilon = -1$.

⁵While the perturbative solutions are separable, the requirement restricts the possible non-linear RSD models (see section 2.3.3)

For each k_{\perp} in the analysis, we need to compute (4.64) as a function of x_{\parallel} . In most cases, this can be achieved by means of a Fast Fourier Transform (FFT). The result can then be cached since FFTs are expensive compared to looking up values in a hash table. In the next section, we will derive analytic formulae for the cases in which the FFT method cannot be used. The outer integration is then done by combining the pre-computed results.

4.4.2 Analytic results for the CMB-lensing function

In this subsection, we derive analytic results for integrals of the form

$$I_{\kappa,n}(x_{\parallel}, \mathbf{k}_{\perp}; \alpha) = \int \frac{dk_{\parallel}}{2\pi} e^{-ik_{\parallel}x_{\parallel}} k^{\alpha} W_{\kappa,n}(k_{\parallel}) \quad (4.65)$$

where $W_{\kappa,n}$ is the lensing window function (4.4). Please note, that the case $\alpha = 0$ corresponds to the inverse Fourier transformation, so one recovers the lensing kernel in real space. The other terms that occur have $\alpha = 2, 4$. Thus, we can use $k^2 = k_{\parallel}^2 + k_{\perp}^2$, $k^4 = k_{\parallel}^4 + 2k_{\parallel}^2 k_{\perp}^2 + k_{\perp}^4$, to restrict ourselves to integrals of the form

$$\tilde{I}_{\kappa,n}(x_{\parallel}, \mathbf{k}_{\perp}; \alpha) = \int \frac{dk_{\parallel}}{2\pi} e^{-ik_{\parallel}x_{\parallel}} k_{\parallel}^{\alpha} W_{\kappa,n}(k_{\parallel}). \quad (4.66)$$

As a first step to analytically solve (4.66), we approximate the time evolution in the lensing kernels (4.4) with polynomials. For the application in this thesis, we do this by fitting a sixth-order polynomial centred around $\chi_s/2$ to the n -th order growth factor, D^n using the least square method and weights $1/D^n$. The sixth order approximation leads to a relative error of less than 10^{-4} for the first and second order. We refer to Fig. 5.2 to see the behavior of linear and quadratic lensing functions we are interested in. The approach is fully general and higher orders of the lensing kernel will simply have different coefficient in the polynomial approximation - and possibly require higher order polynomials too.

Before starting with the derivations, we quickly review spherical Bessel functions whose properties are a useful tool in the computations.

Spherical Bessel functions

The spherical Bessel functions can be defined as

$$j_n(z) = (-1)^n z^n \left(\frac{1}{z} \frac{d}{dz} \right)^n \frac{\sin(z)}{z} \quad (4.67)$$

and evaluating this expression gives the first five as

$$\begin{aligned}
j_0(z) &= \frac{\sin(z)}{z} \\
j_1(z) &= \frac{\sin(z)}{z^2} - \frac{\cos(z)}{z} \\
j_2(z) &= \left(\frac{3}{z^2} - 1\right) \frac{\sin(z)}{z} - 3 \frac{\cos(z)}{z^2} \\
j_3(z) &= \left(1 - \frac{15}{z^2}\right) \frac{\cos(z)}{z} + \left(\frac{15}{z^3} - \frac{6}{z}\right) \frac{\sin(z)}{z} \\
j_4(z) &= \left(\frac{10}{z} - \frac{105}{z^3}\right) \frac{\cos(z)}{z} + \left(1 - \frac{45}{z^2} + \frac{105}{z^4}\right) \frac{\sin(z)}{z}.
\end{aligned} \tag{4.68}$$

Spherical Bessel function scale as $1/z$ in the high z limit, so they all suffer from slow decay. For numerical accuracy, it is useful to utilize the low z approximation

$$j_l(z) \simeq \frac{z^l}{(2l+1)!!} \left(1 - \frac{z^2}{6+4l}\right) \tag{4.69}$$

for $z \ll 1$.

CMB-lensing window in Fourier space

The approximated CMB-lensing function (4.4) consists of a polynomial that approximates the growth factor and the lensing kernel. For each monomial in this approximation, we find the following analytical expressions in Fourier space

$$\begin{aligned}
&\int_{-\chi_s/2}^{\chi_s/2} \frac{dx_{\parallel}}{\chi_s} x_{\parallel}^n (\chi_s/2 - x_{\parallel})(\chi_s/2 + x_{\parallel}) \exp[ix_{\parallel}k_{\parallel}] \\
&= \frac{1}{2} \left(\frac{\chi_s}{2}\right)^{n+2} \int_{-1}^1 dy (1-y^2)y^n \exp[iyq] \\
&= \begin{cases} 2 \left(\frac{\chi_s}{2}\right)^2 \frac{j_1(q)}{q} & n = 0 \\ 2i \left(\frac{\chi_s}{2}\right)^3 \frac{j_2(q)}{q} & n = 1 \\ 2 \left(\frac{\chi_s}{2}\right)^4 \left(\frac{j_2(q)}{q^2} - \frac{j_3(q)}{q}\right) & n = 2 \\ 2i \left(\frac{\chi_s}{2}\right)^5 \left(3 \frac{j_3(q)}{q^2} - \frac{j_4(q)}{q}\right) & n = 3 \\ 2 \left(\frac{\chi_s}{2}\right)^6 \left(\frac{j_5(q)}{q} - 6 \frac{j_4(q)}{q^2} + 3 \frac{j_3(q)}{q^3}\right) & n = 4 \\ 2i \left(\frac{\chi_s}{2}\right)^7 \left(-\frac{j_6(q)}{q} - 10 \frac{j_5(q)}{q^2} + 15 \frac{j_4(q)}{q^3}\right) & n = 5 \end{cases} \tag{4.70}
\end{aligned}$$

where we substituted $q = \chi_s k_{\parallel}/2$ and $y = 2x_{\parallel}/\chi_s$. The CMB-lensing window function in Fourier space (4.4) is then the weighted sum of those terms.

Integral identities for spherical Bessel functions

Having obtained an analytical expression for the lensing window function in Fourier space, we now discuss the projection integral of the bispectrum (4.66). Using the the Bessel function's integral representation in terms of a Legendre polynomial P_l

$$j_l(k_{\parallel}r) = \frac{(-i)^l}{2} \int_{-1}^1 d\mu P_l(\mu) e^{ik_{\parallel}r\mu} \quad (4.71)$$

allows us to compute the Fourier transforms of spherical Bessel functions with polynomial coefficients for $n \geq 0$ as follows

$$\begin{aligned} \int dk_{\parallel} k_{\parallel}^n j_l(k_{\parallel}r) \exp[-ik_{\parallel}x_{\parallel}] &= \int_{-1}^1 d\mu P_l(\mu) \int dk_{\parallel} k_{\parallel}^n \exp[ik_{\parallel}(r\mu - x_{\parallel})] \frac{(-i)^l}{2} \\ &= \int_{-1}^1 d\mu P_l(\mu) \int dk_{\parallel} \frac{1}{(ir)^n} \frac{\partial^n}{\partial \mu^n} \exp[ik_{\parallel}(r\mu - x_{\parallel})] \frac{(-i)^l}{2} \\ &= \int_{-1}^1 d\mu P_l(\mu) \int d\tilde{k}_{\parallel} \frac{1}{(ir)^n r} \frac{\partial^n}{\partial \mu^n} \exp[i\tilde{k}_{\parallel}(\mu - x_{\parallel}/r)] \frac{(-i)^l}{2} \\ &= \int_{-1}^1 d\mu P_l(\mu) \frac{1}{(ir)^n r} \frac{\partial^n}{\partial \mu^n} \int d\tilde{k}_{\parallel} \exp[i\tilde{k}_{\parallel}(\mu - x_{\parallel}/r)] \frac{(-i)^l}{2} \\ &= \int_{-1}^1 d\mu P_l(\mu) \frac{1}{(ir)^n r} \frac{\partial^n}{\partial \mu^n} (2\pi \delta^D(\mu - x_{\parallel}/r)) \frac{(-i)^l}{2} \\ &= \int_{-1}^1 d\mu P_l(\mu) \frac{1}{(ir)^n r} \pi (-i)^l \frac{\partial^n}{\partial \mu^n} \delta^D(\mu - x_{\parallel}/r) \\ &= \int_{-1}^1 d\mu \frac{\pi (-i)^{l+n}}{r^{n+1}} \delta^D(\mu - x_{\parallel}/r) (-1)^n \frac{\partial^n}{\partial \mu^n} P_l(\mu) \\ &= \begin{cases} \frac{\pi}{r^{n+1}} (-1)^n (-i)^{n+l} \frac{\partial^n}{\partial \mu^n} P_l(\mu) \Big|_{\mu=\frac{x_{\parallel}}{r}} & \text{if } |x_{\parallel}/r| < 1 \\ 0 & \text{otherwise.} \end{cases} \end{aligned} \quad (4.72)$$

This approach sheds a light on another approach we could have taken instead of using FFTs to solve (4.64): Approximate the integrands with some base functions that can be integrated analytically. However, since the integration routine spends most of the time in the outer x integration of (4.63) and the FFTs are sufficiently accurate, this approach would only increase the complexity.

4.4.3 Analytic results for Gaussian selection functions

Using Gaussian selection functions leads to analytical solutions too. Integrals of the form

$$I_G(x, \mathbf{k}_\perp; \alpha \in \{0, 2, 4\}, \beta = 0, \gamma = 0, \delta = 0, \varepsilon = 0) = \int_{-\infty}^{\infty} dk_\parallel \frac{e^{-ik_\parallel x_\parallel}}{2\pi} W_G(k_\parallel) k^\alpha, \quad (4.73)$$

can be integrated to

$$I_G(x, \mathbf{k}_\perp; \alpha) = \frac{e^{-\tilde{x}^2/2l^2}}{\sqrt{2\pi}l^{2\alpha+1}} \cdot \begin{cases} 1 & \alpha = 0 \\ (\mathbf{k}_\perp^2 l^4 + l^2 - \tilde{x}^2) & \alpha = 2 \\ \mathbf{k}_\perp^4 l^8 + 2\mathbf{k}_\perp^2 l^6 + l^4(3 - 2\mathbf{k}_\perp^2 \tilde{x}^2) - 6l^2 \tilde{x}^2 + \tilde{x}^4 & \alpha = 4 \end{cases} \quad (4.74)$$

where $\tilde{x} = x_\parallel - \chi$ and we used the Gaussian profile

$$W_G(k_\parallel) = \exp[-(k_\parallel l)^2/2] \exp[ik_\parallel \chi] \quad (4.75)$$

which corresponds in real space to a bin with depth l . Those analytical solutions provide a way to check the accuracy of the FFTs. Moreover, one could also include the Gaussian FoG damping by substituting $\tilde{l}^2 = l^2 + \sigma_v^2$. This illustrates an interesting point: Projections with deep kernels ($l \gg \sigma_v$) are robust against inaccurate non-linear RSD models and photo-z errors as described in section 4.5 of [79]. We will explore this idea in chapter 5.

4.4.4 Symmetries of a survey with identical and thin bins

For moderately many bins, the bulk of the computations in a projected power spectrum and bispectrum forecasting scenario is spent in the outer integral of (4.63). In the outer integral, we combine the cached results of the one-dimensional integrals to the desired bispectrum signal. In a setting with identical and thin bins that are merely centred around different positions, many bispectra are in fact identical up to a constant (coming from the time evolution). We have introduced the same idea for the power spectrum in (4.9) which reduced the scaling for the number of different projected (cross-)power spectra from quadratic to linear in the number of bins, n . For the galaxy bispectrum, we can reduce the scaling from n^3 to $3 \cdot n \cdot (n - 1)/2 + 1$.

To see this, let us label the n (identical) galaxy bins from 0 to $n - 1$ with increasing co-moving distance and enumerate all combinations with different separations between the bins. For the first bin, there are no restrictions and so each of the n possible positions needs to be taken into account. Using symmetry, we only need to consider cases where the second bin

it further away than the first; let us call its index i . The last bin has to have a relative position to the first that is between $-(n-1)+i$ and $n-1$ to make sure all bins are within the survey. Lastly, in the case where $i=0$ (first bin identical with the second), we can use symmetry and only compute cases where the third bin is further away than the first two. Overall

$$n + \sum_{i=1}^{n-1} (n+n-1-i) = \frac{3}{2}n(n-1) + 1. \quad (4.76)$$

4.5 Outlook: Signal compression

In this chapter, we have discussed among other things strategies to speed up projection integrals. The separation of the projection integrals for the theoretical uncertainties of the projected power spectrum and of the projection integrals for the projected bispectrum made both problems tractable. However, even with the achieved improvements, the large number of cross-bispectra configurations still brings several challenges with it:

- Each configuration requires a one-dimensional integral which makes it currently impossible to compute the theory predictions more than a few times for different cosmological parameters. This in turn prohibits error forecasts beyond Fisher forecasting and makes it impossible to fit the theory to data without further approximations. Amplitude parameters such as the bias parameters, f_{NL} and σ_8 , are an exception since caching of the full predictions could allow us to perform the mentioned tasks.
- The large theory vectors are not only challenging to compute, but also make solving the linear systems - which are necessary for SN analyses, Fisher forecasts and fits - very expensive for non-Gaussian covariances. In general, solving a linear system is of complexity $O(m^3) = O(n^9 T^3)$ where m is the length of the theory vector that itself scales cubic with the number of tomographic bins, n , and triangle configurations, T . Note that in the case of Gaussian covariances, one can use spatial homogeneity and group configurations with congruent triangles together in order to achieve a block-diagonal covariance matrix that reduces the computational complexity to $O(n^9 T)$.
- It is not feasible to obtain non-Gaussian, empirical covariances for the projected bispectrum because of the vast amount of N-body simulations required to constrain all the independent parameters in the matrix.

One way to tackle these issues together is the use of signal compression techniques that reduce the numerical complexity of the problem while still preserving most of the cosmological information. The complexities arising from the large number of 3D bispectrum configurations

are well known and work has been done to find good statistics that capture non-Gaussian information while still being numerically tractable [55, 206–209].

Chapter 5

Projected two- and three-point statistics: Forecasts and mitigation of non-linear RSDs¹

Summary

The combination of two- and three-point clustering statistics of galaxies and the underlying matter distribution has the potential to alleviate degeneracies between cosmological parameters and nuisance parameters and can lead to tighter constraints on parameters describing the composition of the Universe and the dynamics of inflation. Here we investigate the relation between biases in the estimated parameters and inaccurate modelling of non-linear redshift-space distortions for the power spectrum and bispectrum of projected galaxy density fields and lensing convergence. Non-linear redshift-space distortions are one of the leading systematic uncertainties in galaxy clustering. Projections along the line-of-sight suppress radial modes and thus allow a trade-off between biases due to non-linear redshift-space distortions and statistical uncertainties. We investigate this bias-error trade-off for a CMASS-like survey with a varying number of redshift bins. Improved modelling of the non-linear redshift-space distortions allows the recovery of more radial information when controlling for biases. Not modelling non-linear redshift-space distortions inflates error bars for almost all parameters by 20%. The information loss for the amplitude of local non-Gaussianities is smaller, since it is best constrained from large scales.

¹This chapter is based on [2].

Outline

This chapter is structured as follows: In section 5.1, we motivate the chapter. In section 5.2.1, we summarise the relevant theory predictions from chapter 2. We then specify the implementation details of the ideas from chapter 4 in section 5.2.2. In section 5.3, we discuss the inference techniques used in our analysis. The results are then presented in section 5.4 and we conclude in section 5.5.

5.1 Introduction

Our understanding of the Universe has been shaped by observations of the CMB and LSS over the last thirty years [20–22]. Future CMB-lensing experiments are expected to refine this picture [37, 38]. Upcoming galaxy surveys that trace the matter distribution of the LSS such as LSST [26], SPHEREx [27], Euclid [28] and DESI [29] will contribute complementary information about the late time evolution of the Universe. Moreover, those surveys will achieve exquisitely small statistical errors due to the vast volumes they cover and high number density of tracers they resolve. Given their three-dimensional origin, upcoming LSS datasets are predicted to eventually contain more information about cosmological parameters than the CMB.

The early Universe’s density distribution was very close to a Gaussian random field [54] which is fully described by the two-point correlation function or its Fourier transform, the power spectrum. The subsequent non-linear evolution changed the matter distribution which manifests itself in a modification of the power spectrum on small scales and non-vanishing higher order correlation functions. The bispectrum, which is the Fourier transform of the three-point-correlation function, is known to contain most of the non-linear information on mildly non-linear scales. In addition, it allows us to break degeneracies between bias and amplitude parameters [210, 200].

CMB-lensing captures the integrated effect of matter onto CMB photons along their path from the surface of last scattering through the LSS to us. Accordingly, lensing spectra are described by projected spectra of the LSS. But there are also very valid reasons to study galaxy clustering statistics in projection. Firstly, tomographic surveys infer the redshift bins of objects and not their precise positions. Moreover, CMB-lensing - galaxy-clustering cross-correlations require a two-dimensional clustering analysis. Lastly, projections offer a way to suppress non-linear RSDs.

RSDs are generated by galaxies’ peculiar velocities parallel to the LOS. The resulting Doppler redshift is degenerate with the redshift from the Hubble flow which is used to determine the radial positions. On large scales, the effect is well described perturbatively,

but on smaller scales one has to resort to empirical models. Work has been done in this direction [71, 82, 78, 211] but the fundamental issue of potentially biased estimates caused by inaccurate modelling remains.

In this chapter, we are quantifying the parameter shifts due to inaccurate RSD modelling. In particular, we study how those biases depend on the chosen projection depth and RSD model used. Given the exquisitely small statistical errors of upcoming surveys, it is worthwhile to make estimators more robust in order to confidently leverage the small statistical uncertainties. [212] recently investigated the parameter shifts arising from an incomplete or incorrect account of bias parameters and selection effects.

In addition, we forecast error bars to investigate the relation between statistical and systematic uncertainties. The constraining power of the galaxy bispectrum for future galaxy surveys has been studied in [213, 112, 212] and power spectrum forecasts for CMB-lensing - galaxy-clustering cross-correlations were performed in [214].

While we employ the flat sky approximation here, there is a growing literature that studies angular (cross-)correlation functions [215–225]. The step from flat to curved sky is conceptually straightforward in our framework using the FFTlog-algorithm [226], but we leave this for future work.

5.2 Theory predictions

5.2.1 Statistics in 3D

Putting all the results from chapter 2 together, we can write down all matter-galaxy cross-power spectra and cross-bispectra. We are interested in those cross-correlations, since the projection of the matter fields (with the appropriate kernel) yields the theory prediction for the lensing convergence and projections of the galaxy field yield the predictions for the projected galaxy spectra. Since the matter fields (and the lensing convergence) do not suffer from RSDs, we use the F_2 kernels (2.98) for their perturbation expansion. The galaxy fields in contrast, are affected by biasing, RSDs and scale-dependent bias from PNGs and thus the Z_i kernels (2.166) are needed in the expansion. The local, equilateral and orthogonal PNG templates are discussed in section 2.2.4 and the Finger-of-God damping model is introduced in section 2.3.3. The leading order power spectra are given by

$$P_{\text{gg}}(k, \mu) = D_{\text{FoG}}^{\text{P}}(k_{\parallel}) \left[Z_1(k)^2 P(k) + \frac{1}{\bar{n}} \right] , \quad (5.1)$$

$$P_{\text{gm}}(k, \mu) = \sqrt{D_{\text{FoG}}^{\text{P}}(k_{\parallel})} Z_1(k) P(k) , \quad (5.2)$$

$$P_{\text{mm}}(k) = P(k) \quad , \quad (5.3)$$

where μ is the cosine of the angle of the wave vector with the LOS. The bispectra are given by

$$B_{\text{ggg}} = D_{\text{FoG}}^{\text{B}}(k_{1,\parallel}, k_{2,\parallel}, k_{3,\parallel}) \times \left[2Z_1(k_1)Z_1(k_1)Z_2(k_1, k_2)P(k_1)P(k_2) + 2 \text{perm.} \right. \\ \left. + Z_1(k_1)Z_1(k_2)Z_1(k_3)B_{\text{prim}}(k_1, k_2, k_3) + \frac{b_1}{\bar{n}} (Z_1(k_1)P(k_1) + 2 \text{perm.}) + \frac{1}{\bar{n}^2} \right], \quad (5.4)$$

$$B_{\text{ggm}} = D_{\text{FoG}}^{\text{B}}(k_{1,\parallel}, k_{2,\parallel}) \left[2Z_1(k_1)Z_2(k_1, k_3)P(k_1)P(k_3) \right. \\ \left. + 2Z_1(k_2)Z_2(k_2, k_3)P(k_2)P(k_3) + 2Z_1(k_1)Z_1(k_2)F_2(k_1, k_2)P(k_1)P(k_2) \right. \\ \left. + Z_1(k_1)Z_1(k_2)B_{\text{prim}}(k_1, k_2, k_3) + \frac{1}{\bar{n}} \left(b_1 + \frac{b_{\Psi}(k_3)}{M(k_3)} \right) P(k_3) \right], \quad (5.5)$$

$$B_{\text{gmm}} = D_{\text{FoG}}^{\text{B}}(k_{1,\parallel}) \left[2F_2(k_1, k_2)Z_1(k_1)P(k_1)P(k_2) + 2F_2Z_1(k_1)P(k_1)P(k_3) \right. \\ \left. + 2Z_2P(k_2)P(k_3) + Z_1(k_1)B_{\text{prim}}(k_1, k_2, k_3) \right], \quad (5.6)$$

$$B_{\text{mmm}} = \left[2F_2P(k_1)P(k_2) + 2 \text{perm.} \right] + B_{\text{prim}}(k_1, k_2, k_3). \quad (5.7)$$

For the sake of compactness, we omitted the arguments of the bispectra. The matter bispectrum depends only on the magnitude of the three wave vectors. RSDs break the statistical isotropy and introduce an explicit dependence on the projection of the wave vector of the galaxy fields on the LOS.

5.2.2 Projected statistics and survey specifications

In chapter 4, we derived the expressions for the projected power spectrum (4.8) and projected bispectrum (4.63) in the flat sky approximation. Both integrals allow us to compute the projected spectra from the 3D theory predictions. The projection kernel for the lensing convergence is determined from GR (4.4) and it remains to specify the tomographic galaxy bins that constitute the galaxy survey.

In this work, we consider a setting with n identical tomographic galaxy bins with projection depths, l , that are centred around

$$\chi_i = \chi_c + [i - (n + 1)/2] l \quad 1 \leq i \leq n \quad (5.8)$$

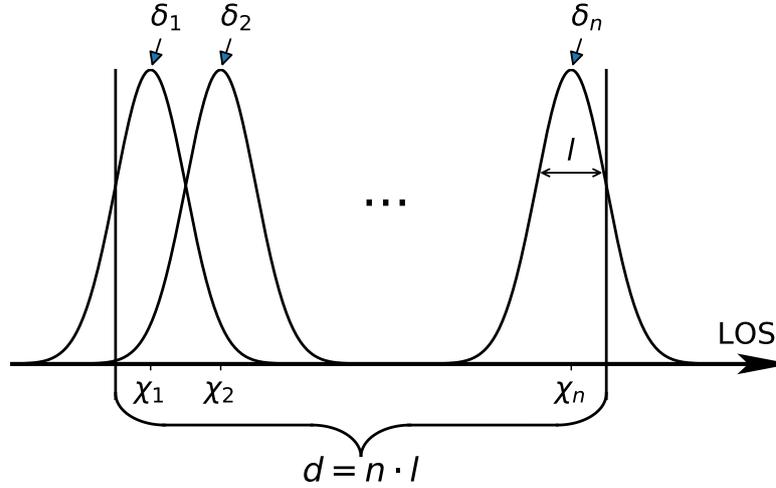


Fig. 5.1 We use a galaxy survey of fixed length d consisting of n identical tomographic bins along the LOS. The centre of each bin is given by (5.8). Beside the Gaussian windows that are displayed here, we also use Top-Hat windows since they cover the survey volume evenly.

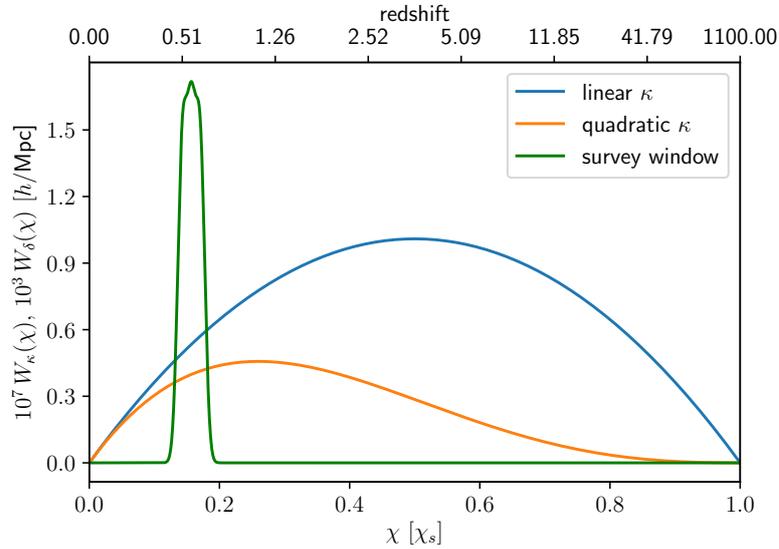


Fig. 5.2 Overview of the windows used in this chapter. The CMB-lensing kernel (blue) peaks halfway between the observer and the surface of last scattering at χ_s . The orange curve shows the lensing kernel multiplied with the growth factor. We show in green the sum of three galaxy bins with width and distance between the their centres of $200 h^{-1} \text{Mpc}$ centred around $z=0.57$. The combined galaxy window shown here is constructed to resemble the CMASS galaxy sample.

where χ_c is the co-moving distance to the centre of the survey. Since we have a CMASS-like survey in mind, we choose $\chi_c = \chi(z = 0.57)$. The full survey specifications are summarised in Table 5.1b. For simplicity, we ignore the time evolution within the tomographic (galaxy) bins. Fig. 5.1 gives an overview of the galaxy survey in mind. In addition to galaxy clustering, we consider CMB-lensing and illustrate the overall setting along the LOS in Fig. 5.2.

While we do not work with photometric redshift errors in this work, they can be directly included in our formalism as discussed in section 4.5 of [79]. For instance, assuming a Gaussian error distribution of the photometric redshift errors, one can capture these errors by changing the velocity dispersion as $\sigma^2 = \sigma_v^2 + \sigma_z^2$. Here σ_z^2 is the uncertainty in the co-moving position of the tracers. Other functional forms for the photo- z errors could also be included but this would introduce similar complications as discussed for non-Gaussian selection functions in chapter 4.

For our forecasts and bias estimations, we vary the number of tomographic bins such that the depth of each bin scales as $l = d/n$, where d is the survey depth along the LOS. We use a k -space binning of k_f (PS), $4k_f$ (BS) and take three base points per fundamental frequency to obtain the averaged signal over k -bins. The power spectrum and bispectrum covariances are given in section 4.2. We use the Planck best-fit cosmology [20] as our fiducial cosmology: $\Omega_B = 0.0494$, $\Omega_M = 0.3144$, $h = 0.6732$, $\sigma_8(z = 0.57) = 0.6029$, $n_s = 0.966$.

5.3 Inference methods

In this section, we first review Fisher forecasting and then outline our approach for computing parameter shifts due to inaccurate theoretical predictions.

5.3.1 Fisher forecasting

The Cramer-Rao bound provides a lower bound on the statistical error for any unbiased estimator in terms of the inverse of the Fisher information (matrix)

$$F_{ij} = \langle (\log \mathcal{L})_{,ij} \rangle, \quad (5.9)$$

where i, j label the parameters of interest, \mathcal{L} is the likelihood function and all quantities are evaluated at the maximum-likelihood point. Assuming a Gaussian likelihood for the power

spectrum and bispectrum, the Fisher information can be calculated as [227]

$$\begin{aligned} F_{ij} &= \frac{1}{2} \text{Tr} [C^{-1}C_{,j}C^{-1}C_{,i} + C^{-1}(\mu_{,i}\mu_{,j}^T + \mu_{,j}\mu_{,i}^T)] \\ &\simeq \mu_{,i}^T C^{-1} \mu_{,j}. \end{aligned} \quad (5.10)$$

The theory vector μ contains the spectra of interest and C is the covariance. The derivatives in 3D can be computed via finite differences [228]. Using the product rule and (4.8, 4.63) allows us to compute the derivatives in 2D. The (un)marginalised error forecasts are then given by

$$\sigma_i^2 = \begin{cases} 1/F_{ii} & \text{unmarginalised} \\ (F^{-1})_{ii} & \text{marginalised.} \end{cases} \quad (5.11)$$

5.3.2 Parameter estimation

In a scenario where we are interested in the parameter shifts due to inaccurate (theoretical) modelling, we fit some theoretical model μ_θ to the underlying ground truth μ_{true} . Assuming a Gaussian distribution, this is done by choosing the parameters θ that maximise the following log-likelihood,

$$-\chi^2 = -\frac{1}{2}(\mu_{\text{true}} - \mu_\theta)^T C^{-1}(\mu_{\text{true}} - \mu_\theta) + \text{const.} \quad (5.12)$$

Since we investigate small biases, we can linearise the theoretical model around the best fit parameters θ_* as

$$-\chi^2 = -\frac{1}{2}v^T C^{-1}v, \quad (5.13)$$

where

$$v = \mu_{\text{true}} - \mu_{\theta_*} - (\theta_* - \theta) \cdot \left. \frac{\partial \mu_\theta}{\partial \theta} \right|_{\theta_*}. \quad (5.14)$$

Since we have access to the ground truth parameters, we use those as best fit parameters. The likelihood of the linearised model has an explicit minimum

$$\theta = \theta_* + F^{-1}\mathbf{b}, \quad (5.15)$$

where F is the Fisher information and \mathbf{b} is given by

$$\mathbf{b} = (\mu_{\text{true}} - \mu_{\theta_*}) C^{-1} \left. \frac{\partial \mu_\theta}{\partial \theta} \right|_{\theta_*}, \quad (5.16)$$

where for each component of \mathbf{b} , one partial derivative is taken. (5.15) allows us to compute the biases from using an inaccurate model.

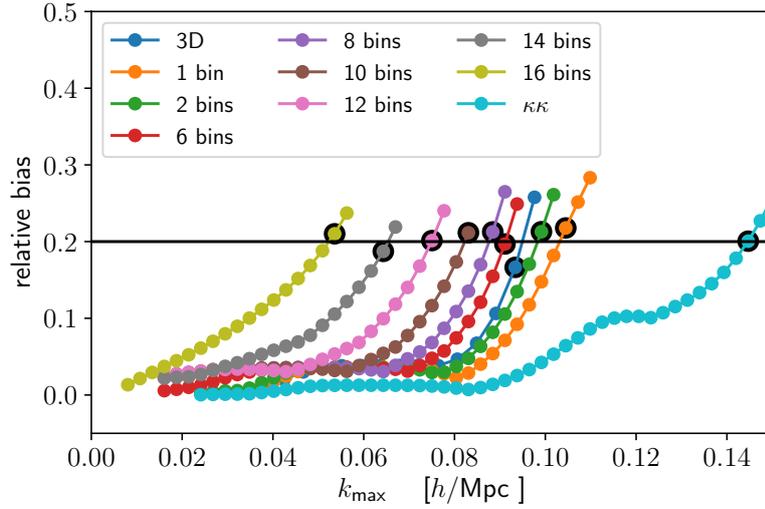


Fig. 5.3 Relation between k_{\max} and maximal relative bias for a varying number of redshift bins for a CMASS-like survey (see Table 5.1a). The points with the black border are closest to 0.2 for each configuration and thus used as cut-offs in this chapter. Their numerical values can be found in Table 5.1b. The 3D curve was determined from a cubic survey of side length $l = (A \cdot d)^{1/3}$ with the same redshift and shot-noise as the 2D setting.

5.4 Results

In this section, we first establish a consistent analysis cut-off, k_{\max}^{2D} , as a function of the number of redshift bins that allows us to compare 2D surveys with different projection depths without the need of theoretical uncertainties. We then use this cut-off to analyse the error-bias trade off as a function of the FoG model (2.118, 2.122). We end the section with more optimistic forecasts.

Throughout this section, we use the finding that the cross-covariances between power spectra and bispectra are negligible at large scales [229, 230, 213]. When not stated otherwise, we work with a CMASS-like survey as described in Fig. 5.2 and Table 5.1a.

5.4.1 Choosing the cut-off scale for projected spectra

As explained in section 4.3.2, it is unfeasible to directly implement theoretical uncertainties for the projected bispectrum. Thus, we need another approach to control for theoretical uncertainties in the matter predictions. In this work, we control these systematics by choosing a cut-off scale, k_{\max} , that ensures that all parameter shifts due to inaccurate matter modelling are below 20% of the corresponding error bars. We estimate the parameter shifts by fitting a linear matter power spectrum to the HALOFIT [4] predictions as described in section 5.3.2; both models are without RSDs. Fig. 5.3 illustrates the monotonic relation between cut-off

Table 5.1 (a): Characterisation of the CMASS-like survey we use in this chapter. The velocity dispersion parameters were obtained from fits against N -body simulations. (b): Cut-offs for different projection depths that ensure all relative biases in the matter predictions are below 20% for a CMASS-like survey described on the right side. The 3D cut-off comes from a cubic survey with the same volume. The values correspond to the black points in Fig. 5.3.

parameter	(a)	(b)	
	value	Type	k_{\max} [$h^{-1}\text{Mpc}$]
depth	$590 h^{-1}\text{Mpc}$	3D*	0.093
Area	$2345 \text{Mpc}^2 h^{-2}$	1 bin	0.10
z_{eff}	0.57	2 bins	0.099
\bar{n}	$2 \cdot 10^{-4} \text{Mpc}^3 h^{-3}$	4 bins	0.094
b_1	2.31	6 bins	0.091
b_{s^2}	$-\frac{4}{7}(b_1 - 1)$	8 bins	0.088
b_2	0.77	10 bins	0.083
σ_P	$4 h^{-1}\text{Mpc}$	12 bins	0.075
σ_B	$5.5 h^{-1}\text{Mpc}$	16 bins	0.054
		lensing	0.14

scale and maximal relative biases for different projection depths. The cut-off is chosen to be the value where the maximal relative bias in the cosmological parameters is closest to 20%. Those values are marked in black in the figure and the numerical values are reported in Table 5.1b. We want to point out that the precise values of those cut-offs are specific for the survey specified in Table 5.1a.

Since the amount of imprecise, small-scale information that gets sourced to larger scales increases with decreasing projection depth, we see that the cut-off decreases as the number of tomographic bins increases. The effective cut-off in 3D, where no sourcing happens, lies in between those extremes because of the different k dependence of the 2D and 3D covariance functions. The lensing cut-off is significantly larger for two reasons. Firstly, the kernel is very narrow in Fourier space and secondly, it peaks at early times, where non-linearities are small.

In Fig. 5.4 we compare the forecasted error bars from the galaxy power spectrum and bispectrum using the chosen cut-offs in two and three dimensions. As the number of tomographic bins increases, one gains information by resolving more of the modes parallel to the LOS from cross-correlations between the tomographic bins but loses information at the same time from the overall decreasing cut-off scale. For both the power spectrum (dashed line) and the bispectrum (dotted), we see an increase in information until ~ 10 bins when the latter effects overtake and the information decreases again. This approach allows us to recover more than 80% of bias/amplitude parameters and more than 90% of cosmological

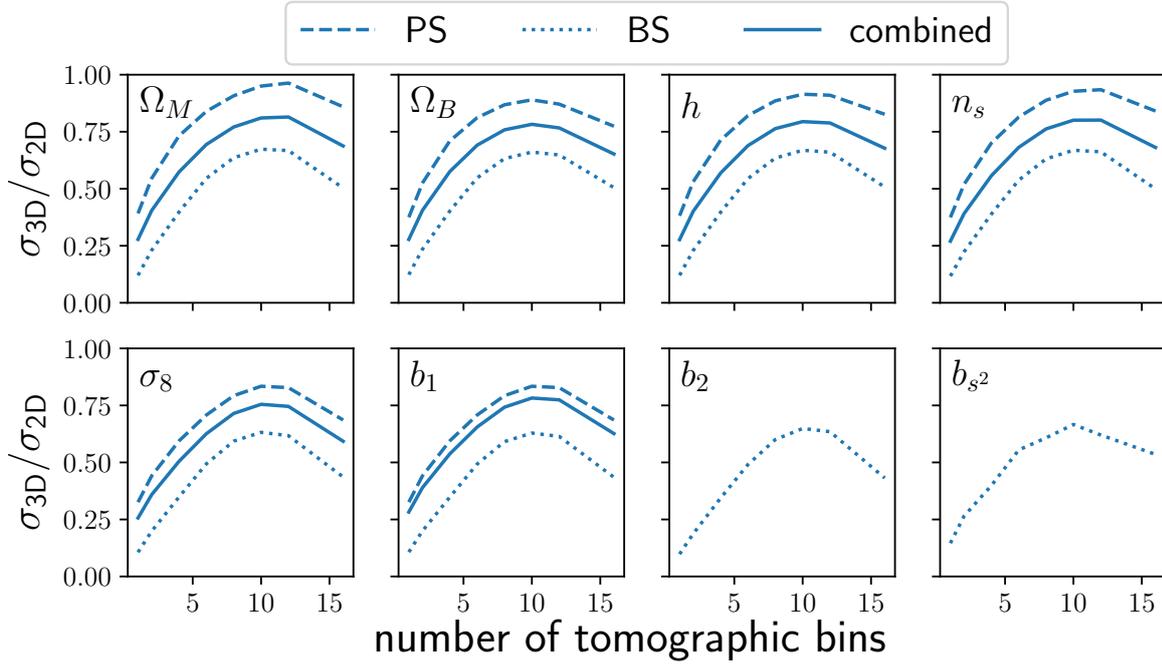


Fig. 5.4 The ratio of unmarginalised error bars in two and three dimensions is shown for the galaxy power spectrum (dashed), galaxy bispectrum (dotted) and combined (solid lines). With an increasing number of tomographic bins, the 2D error bars first decrease due to the increase in resolved radial information. Around 10 bins, the decrease in the cut-off k_{\max}^{2D} takes over and the error increases again.

parameters in a power spectrum analysis. In a pure bispectrum analysis, two thirds of the Fisher information can be recovered compared to a 3D analysis.

5.4.2 Signal-to-noise analysis

There are three effects that determine the signal-to-noise (SN) scaling with respect to the number of tomographic bins: 1) As one increases the number of redshift bins, the projection depth decreases and more radial signal is resolved. 2) If at least one galaxy field is involved, k_{\max}^{2D} decreases with smaller projection depth, which in turn decreases the SN too. 3) The galaxy selection function changes when we vary the number of Gaussian profiles. Whereas the first two effects are relevant on all scales, the latter effect's size decreases with the number redshift bins, and is negligible from 4 bins onward. This justifies the use of Gaussian bins. Fig. 5.5 displays the SN for all individual power spectra and bispectra in our fiducial cosmology including RSDs. Due to our conservative cut-offs, the galaxy clustering and lensing auto-power spectra have the most SN. The galaxy bispectrum's SN is strongly increasing with the number of bins and nearly reaches the lensing power spectrum SN at its

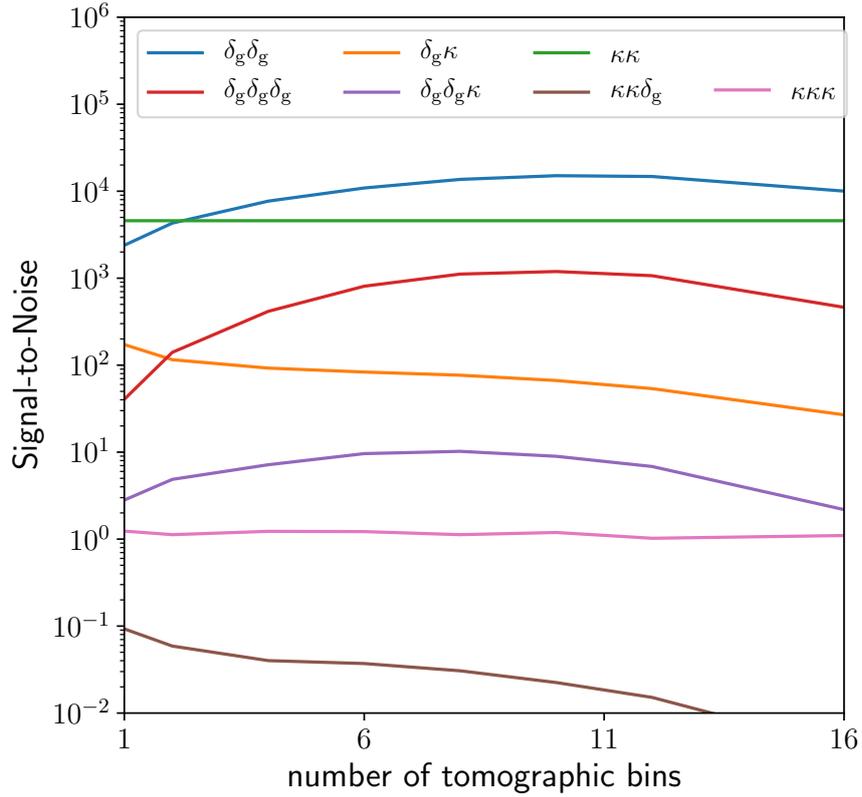


Fig. 5.5 Signal-to-noise scaling for different power spectra and bispectra as a function of the number of tomographic bins. The dependence has three components: 1) With more tomographic bins, more radial information is resolved. 2) The cut-off, k_{\max}^{2D} , decreases with the number of tomographic bins which in turn decreases the SN. 3) The galaxy selection function is changing when varying the number of Gaussian profiles. This effect is negligible from four bins onward.

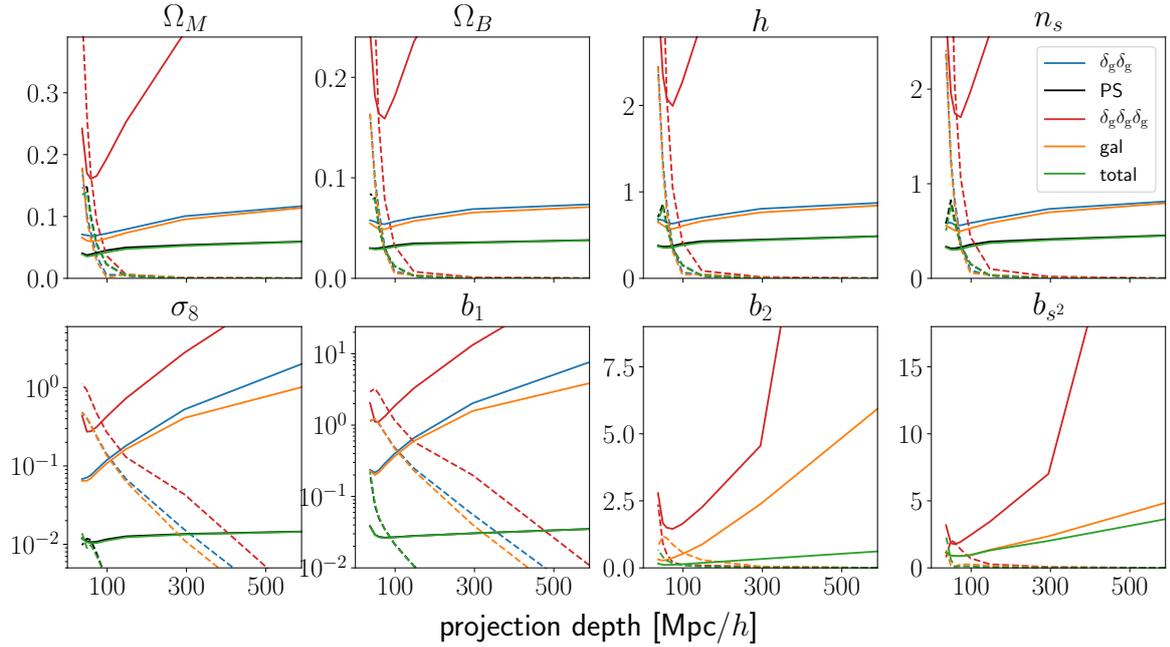


Fig. 5.6 Marginalised 1σ error bars on cosmological parameters as a function of the projection depth for a selection of cross-spectra are displayed in solid lines. The Fisher forecasts assume FoG Model 2, but are actually independent of the chosen FoG model. The corresponding dashed lines represent five times the bias due to inaccurate theoretical modelling for the FoG model that misspecifies the velocity dispersion by 50%. All forecasted errors are minimal at $\sim 60 h^{-1}\text{Mpc}$.

maximum at 10 bins. The cross-spectra tend to be much smaller than those auto-correlations since the overlap between the galaxy survey and the lensing kernel is small (see Fig. 5.2). The lensing bispectrum SN shows small fluctuations since the binning of its (constant) cut-off is changing to ensure it is consistent with the changing galaxy clustering binning.

In section 4.2.2, we derived analytical results for the SN scaling of most projected (cross-) power spectra using their Gaussian covariance. The scaling (4.30) is independent of the chosen RSD model. For the bispectrum, no analytical results exist but their SN scalings are still approximately independent of the RSD model.

5.4.3 Error - bias tradeoff

In this subsection, we investigate the error-bias trade-off in two scenarios. First, we assume a ΛCDM cosmology and forecast the parameter uncertainties and shifts due to inaccurate FoG modelling. Next, we perform a similar analysis for the bias and PNG parameters for a fixed cosmology. In both cases, we assume a ground-truth Gaussian FoG model with velocity dispersions specified in Table 5.1a. We then study inaccurate models by changing

the velocity dispersion parameters by -100% (i.e. no FoG modelling), -50% , -10% , 10% and 50% . Since we find that the biases are approximately independent of the sign of the shift in the velocity dispersions, we restrict ourselves to the first three cases and refer to them as model 1-3. The statistical uncertainties that we report together with biases are always the ones obtained from the inaccurate FoG model. In practice however, the uncertainties are nearly independent of the chosen FoG model.

Moreover, we only show the results of the following five spectra combinations that we consider most interesting: Galaxy power spectrum ($\delta_g \delta_g$), all power spectra combined (PS), galaxy bispectrum ($\delta_g \delta_g \delta_g$), galaxy power spectrum and bispectrum combined (gal) and all power spectra plus bispectra combined (total).

Cosmological forecasts

We fit the three above-mentioned FoG models to the fiducial model and report the error bars and relative biases for all parameters. Fig. 5.6 shows the scaling of the relative biases and error bars for the model where the velocity dispersion is 50% smaller than its fiducial value (Model 2). The five combinations can roughly be grouped into three sets. The galaxy bispectrum has the least constraining power of the considered spectra and the derived error bars decrease very strongly with the number of tomographic bins until they reach a minimum around 10 bins (projection depth $l \simeq 60 h^{-1} \text{Mpc}$). For smaller projection depths, the error bars increase again due to the decreasing k_{max}^{2D} . The galaxy power spectrum and the combination of galaxy power spectrum and bispectrum scales more weakly with the number of resolved bins and have their minimum around 10 bins too. Finally, adding information from galaxy lensing to either the galaxy power spectrum alone or both power spectrum and bispectrum allows us to lower error bars once more. This is partly because CMB-lensing allows us to break the b_1 - σ_8 degeneracy. It does not help to constrain b_{s2} better, as the trace-free part of the tidal tensor is rather uncorrelated to the trace of the tidal tensor that corresponds to the amplitude/bias parameters, which CMB-lensing can constrain well. The relative biases remain close to zero for all parameters, with the exception of σ_8 and b_1 , until a projection depth of $l \sim 100 h^{-1} \text{Mpc}$, when they start growing quickly. Changing the FoG model, has a marginal impact on the error bars but shifts the curves of the relative biases left (right) if the FoG model becomes more (less) accurate.

In Table 5.2, we summarise our findings for the optimal error bars (left side) together with the relative biases (right side) conditioned on being smaller than 20%. We observe that the error bar difference across the spectra is significantly larger than the difference within the spectra across FoG models.

Table 5.2 Optimal forecasted relative errors for three FoG models considered. For each spectra, we give the number of tomographic bins such that the maximal relative bias is below 20% and the statistical uncertainties are minimal. On the left side, we report the marginalised relative 1σ uncertainties. On the right side, we show the corresponding relative biases.

	relative 1σ uncertainties [%]					relative biases [%]				
	$\delta_g \delta_g$	PS	$\delta_g \delta_g \delta_g$	gal	total	$\delta_g \delta_g$	PS	$\delta_g \delta_g \delta_g$	gal	total
Model 1: $\sigma_v = 0 \cdot \sigma_{v, \text{ground truth}}$										
bins	4	4	6	4	4					
Ω_M	25	16	60	23	15	1.9	2.2	-5.4	2.3	2.6
Ω_B	120	70	360	110	68	1.4	1.8	-5	1.5	2.0
h	100	64	330	99	62	1.7	1.7	-5.2	1.9	2.0
n_s	65	40	200	61	39	-1.5	-1.9	6.2	-1.7	-2.2
σ_8	30	2.1	65	27	2.0	10	-1.4	18	10	-1.6
b_1	29	1.2	74	26	1.2	-10	10	-18	-10	10
b_2			210	110	24			-1.9	-9.1	6.7
b_{s^2}			320	180	170			8.1	-2.9	-0.8
Model 2: $\sigma_v = 0.5 \cdot \sigma_{v, \text{ground truth}}$										
bins	4	6	6	4	6					
Ω_M	25	14	61	23	13	1.5	10	-3.8	1.7	11
Ω_B	120	66	360	110	63	1.0	7.4	-3.6	1.1	8.0
h	100	60	340	99	57	1.3	7.5	-3.7	1.4	8.2
n_s	65	37	210	61	35	-1.1	-8.5	4.5	-1.3	-9.2
σ_8	30	1.9	70	27	1.8	7.5	-6.2	13	7.7	-6.5
b_1	29	1.2	80	26	1.2	-7.5	16	-13	-7.7	16
b_2			210	110	18			-1.3	-6.8	15
b_{s^2}			330	180	130			5.9	-2.2	-2.4
Model 3: $\sigma_v = 0.9 \cdot \sigma_{v, \text{ground truth}}$										
bins	8	10	10*	8	10					
Ω_M	22	12	50	19	11	-2.6	13	-5.1	-2.5	13
Ω_B	110	60	330	97	58	-3	8.7	-4.6	-2.9	9.0
h	93	54	310	85	52	-3	8.5	-4.9	-2.9	8.6
n_s	58	33	180	52	31	3.5	-9.8	5.7	3.4	-9.9
σ_8	15	1.8	41	14	1.7	13	-5.4	14	15	-4.9
b_1	12	1.2	44	11	1.2	-14	9.3	-13	-15	9.6
b_2			200	48	14			-2.1	-12	12
b_{s^2}			240	120	120			5.0	-1.4	0.5

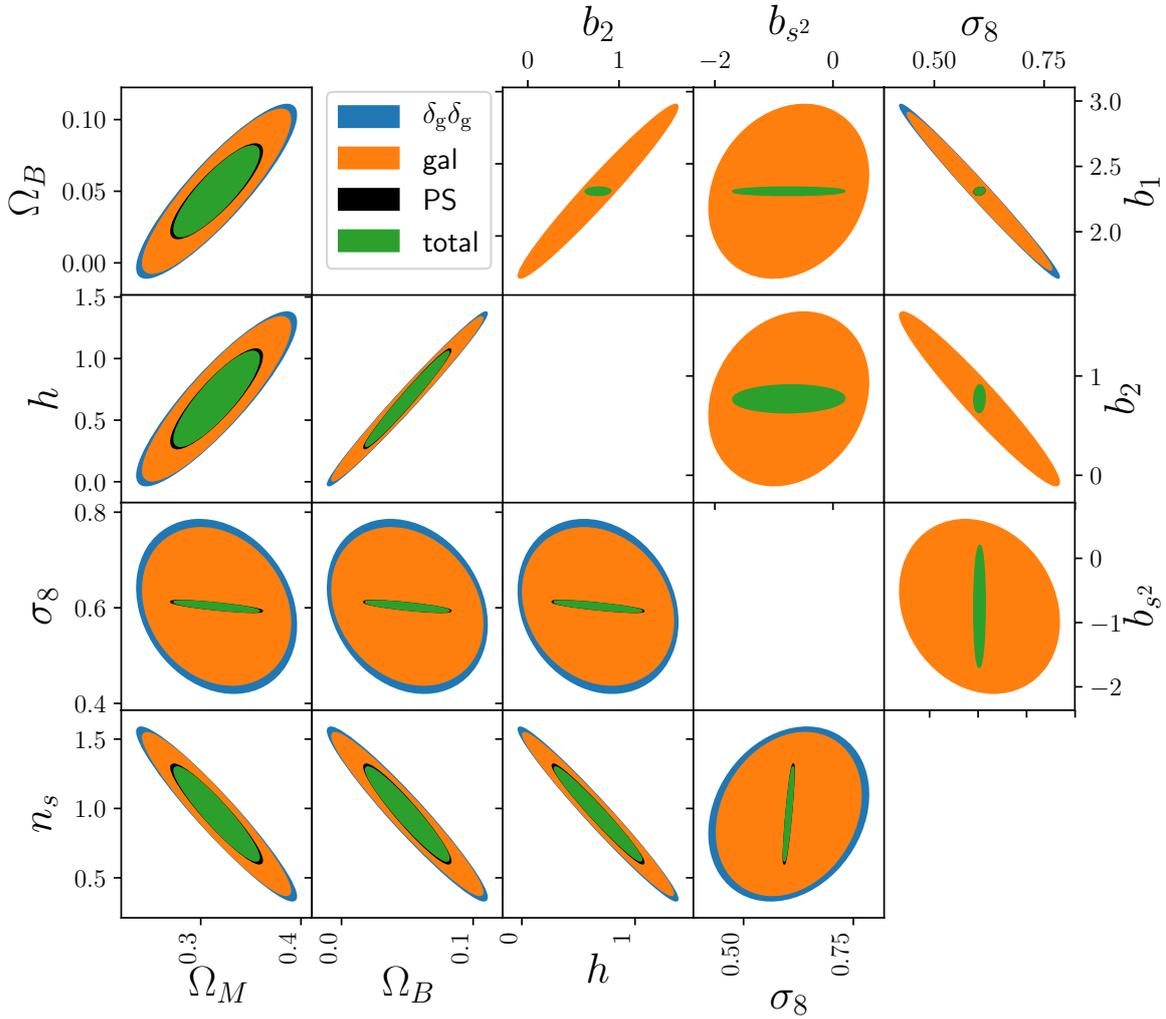


Fig. 5.7 The 1σ contour plots for four different spectra using FoG Model 2 (velocity dispersion underestimated by 50%). The bins are chosen to minimise the errors controlling for systematic uncertainties. The optimal number of bins are reported in Table 5.2.

The decrease in Fisher information for projection depths $\leq 60 h^{-1}\text{Mpc}$ implies that even with a velocity dispersion that is off by 10%, one is able to fully recover the available information in two dimensions. In contrast, a misspecification of 50% leads to a $\sim 10\%$ increase in the error for cosmological parameters and more for bias parameters. The worst case scenario of not modelling the FoG effect at all, inflates the error bars by $\sim 20\%$.

The relative biases tend to be strongest in parameters that affect the amplitude strongly and tend to be positive, since the models considered underestimate the FoG damping. The 2D contour plots of the best case Fisher information matrix of Model 2 (see Table 5.2) is displayed in Fig. 5.7. One sees that all spectra have approximately the same covariance structure and that CMB-lensing helps to break the σ_8 - b_1 degeneracy. The positive correlation

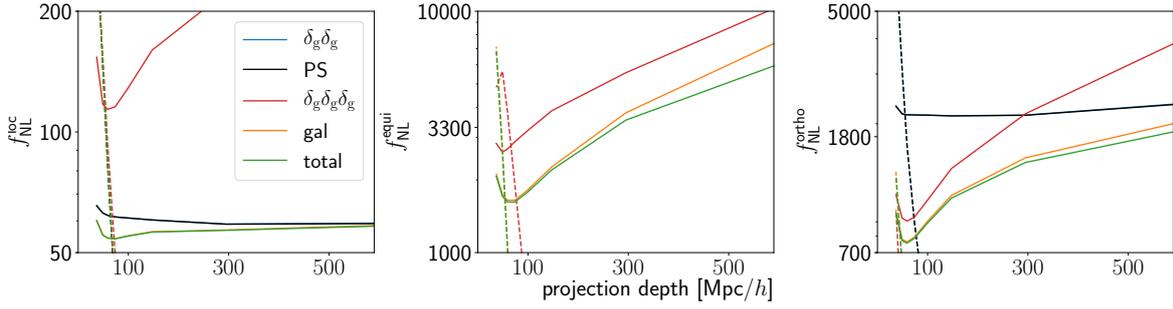


Fig. 5.8 The marginalised 1σ error bars on PNGs as a function of the projection depth for a selection of cross-spectra are displayed in solid lines. The Fisher forecasts assume FoG Model 2, but are almost independent of the chosen FoG model. The corresponding dashed lines represent five times the bias due to inaccurate theoretical modelling for the FoG model that misspecifies the velocity dispersion by 50%. Each subplot corresponds to a different forecasts and we marginalised over the bias parameters.

between b_1 and b_2 is explained as follows: The three bias/amplitude parameters σ_8 , b_1 and b_2 are all pairwise anticorrelated. However, the positive definiteness of the covariance matrix pushes the weakest among them, $b_2 - b_1$, to a positive value in the joint analysis.

PNG forecasts

Assuming a known cosmology, we now forecast bias and f_{NL} parameters for the local, equilateral and orthogonal shape. We perform a separate forecast for each template and FoG model and illustrate the dependence on the number of tomographic bins in Fig. 5.8.

The scale-dependent bias in the galaxy power spectrum yields the best constraints for the local shape. Since the survey is roughly four times as wide as deep, the constraints from the power spectrum have no dependency on the projection depth. The error from the bispectrum analysis, in contrast, shows a strong dependence on the projection depths. For few bins, the error is very large but decreases quickly with decreasing projection depths. As with the cosmological parameters, the error bars become worse for more than 10 bins due to the decreasing cut-off. Since the equilateral shape does not lead to scale-dependent bias, it can only be constrained from the bispectrum and the power spectra only contribute towards reducing the uncertainty in the bias parameters. The constraints for the orthogonal shape from the power spectrum and bispectrum are of similar order, since the scale-dependent bias only scales as $1/k$.

The relative biases are close to zero for large projection depths and only start playing a role around $100 h^{-1}\text{Mpc}$. We report the optimal forecasts with relative biases below 20% in Table 5.3. We observe that the error bar differences across spectra are larger than across FoG models given a spectra. For the local shape, the latter differences are basically zero, since

only the largest scales are relevant. The constraints of the equilateral and orthogonal shape in contrast improve by $\sim 10\%$ when modelling the FoG effect precisely.

Table 5.3 Optimal forecasted errors for a bias- f_{NL} analysis with fixed cosmological parameters. We present separate forecasts for the three FoG models considered. For each spectra, we give the number of tomographic bins such that the maximal relative bias is below 20% and the statistical uncertainties are minimal. On the left side, we report the marginalised 1σ uncertainties. On the right side, we show the corresponding relative biases.

	1σ uncertainties					relative biases [%]				
	$\delta_{\text{g}}\delta_{\text{g}}$	PS	$\delta_{\text{g}}\delta_{\text{g}}\delta_{\text{g}}$	gal	total	$\delta_{\text{g}}\delta_{\text{g}}$	PS	$\delta_{\text{g}}\delta_{\text{g}}\delta_{\text{g}}$	gal	total
Model 1: $\sigma_{\text{v}} = 0 \cdot \sigma_{\text{v, ground truth}}$										
bins	2	2	6	2	2					
$f_{\text{NL}}^{\text{local}}$	60	60	120	56	56	-1.4	-1.4	-1.1	-1.7	-1.7
b_1	0.020	0.020	0.29	0.020	0.020	16	16	-0.47	17	17
b_2			1.1	0.18	0.18			11	7.2	7.2
b_{s^2}			1.3	1.3	1.3			0.81	-0.79	-0.79
bins			6	2	2					
$f_{\text{NL}}^{\text{equi}}$			2900	2300	2200			-18	-0.51	-0.47
b_1			0.51	0.016	0.016			-15	20	20
b_2			2.2	0.36	0.35			20	3.9	4.0
b_{s^2}			1.6	1.4	1.4			9.1	-0.92	-0.90
bins	6	6	10	4	4					
$f_{\text{NL}}^{\text{ortho}}$	2200	2200	920	900	880	11	11	12	4.5	4.4
b_1	0.16	0.16	0.42	0.068	0.067	19	19	17	15	15
b_2			1.4	0.21	0.21			-4.4	8.9	9.0
b_{s^2}			1.4	1.1	1.1			3.7	-0.88	-0.95

138 Projected two- and three-point statistics: Forecasts and mitigation of non-linear RSDs

	1 σ uncertainties					relative biases [%]				
	$\delta_g \delta_g$	PS	$\delta_g \delta_g \delta_g$	gal	total	$\delta_g \delta_g$	PS	$\delta_g \delta_g \delta_g$	gal	total
Model 2: $\sigma_v = 0.5 \cdot \sigma_{v, \text{ground truth}}$										
bins	2	2	8	2	2					
$f_{\text{NL}}^{\text{local}}$	60	60	110	56	56	-1.1	-1.1	-1.6	-1.3	-1.3
b_1	0.020	0.020	0.30	0.020	0.020	12	12	0.064	12	12
b_2			1.1	0.18	0.18			11	5.4	5.4
b_{s^2}			1.3	1.3	1.3			4.1	-0.60	-0.59
bins			6	2	2					
$f_{\text{NL}}^{\text{equi}}$			2900	2300	2200			-13	-0.38	-0.35
b_1			0.51	0.016	0.016			-11	15	15
b_2			2.2	0.36	0.35			15	2.9	3.0
b_{s^2}			1.6	1.4	1.4			6.8	-0.69	-0.67
bins	6	6	10	4	4					
$f_{\text{NL}}^{\text{ortho}}$	2200	2200	930	900	890	8.3	8.3	8.6	3.3	3.3
b_1	0.16	0.16	0.42	0.068	0.067	15	15	13	11	11
b_2			1.4	0.21	0.21			-3.1	6.6	6.7
b_{s^2}			1.4	1.1	1.1			2.8	-0.66	-0.71
Model 3: $\sigma_v = 0.9 \cdot \sigma_{v, \text{ground truth}}$										
bins	6	6	10	6	6					
$f_{\text{NL}}^{\text{local}}$	61	61	120	54	54	-3.2	-3.2	-4.1	-3.8	-3.8
b_1	0.017	0.017	0.31	0.016	0.016	17	17	-0.035	18	18
b_2			1.1	0.12	0.12			2.9	8.0	8.0
b_{s^2}			1.3	0.88	0.87			1.1	-1.1	-1.1
bins			10	4	4					
$f_{\text{NL}}^{\text{equi}}$			2700	1800	1800			-7.3	-0.38	-0.36
b_1			0.50	0.013	0.013			-5.9	10	10
b_2			2.1	0.28	0.28			7.5	2.1	2.1
b_{s^2}			1.5	1.0	1.0			3.9	-0.76	-0.75
bins	10	10	10	10	10					
$f_{\text{NL}}^{\text{ortho}}$	2200	2200	910	790	780	7.4	7.4	1.1	4.2	4.2
b_1	0.16	0.16	0.37	0.059	0.058	11	11	0.52	14	14
b_2			1.3	0.19	0.18			1.4	4.3	4.4
b_{s^2}			1.3	1.2	1.2			0.69	3.5	3.5

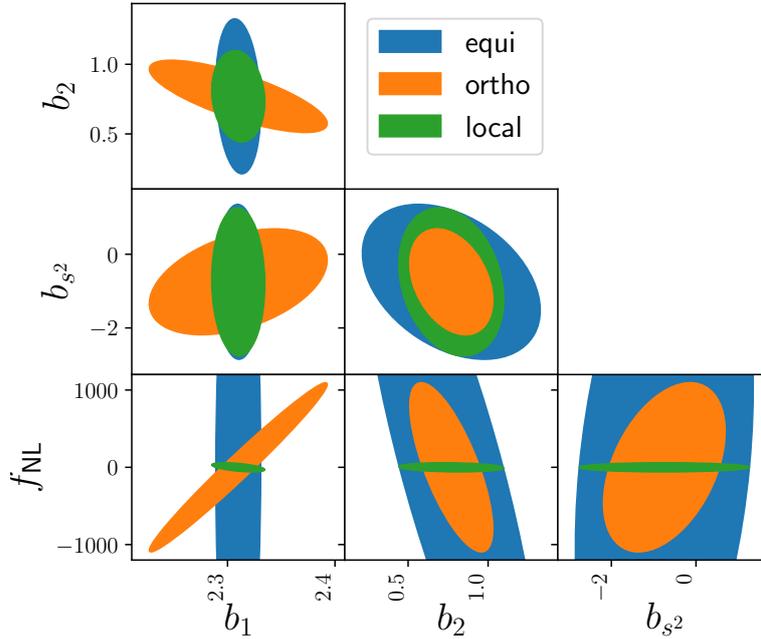


Fig. 5.9 1σ contour plots for an f_{NL} -bias forecast combining all power spectra and bispectra and using FoG Model 2. The bins are chosen to minimise the errors controlling for systematic uncertainties. Each colour corresponds to an independent forecast.

We show the correlation between bias and f_{NL} parameters in Fig. 5.9. One sees that the weaker scale dependence of the scale-dependent bias of the orthogonal PNGs, compared to local PNGs, leads to a considerable correlation between b_1 and f_{NL} . For the equilateral PNGs, the bias is not scale dependent on large scales and cannot be disentangled from the ordinary linear bias. Thus, we do not use it to constrain equilateral PNGs which is why the correlation between f_{NL} and b_1 is low.

5.4.4 Optimistic forecasts

We believe an optimistic but still realistic scenario is given by an upper cut-off of $k_{\text{max}}^{2\text{D}} = 0.15 h\text{Mpc}^{-1}$ and 6 redshift bins. This is in line with the choice of [112, 213]. [212] were able to work with larger cut-offs for the power spectrum by using separate values for the power spectrum and bispectrum and computing parameter shifts (5.15) relative to the next perturbative order instead of the fully non-linear HALOFIT prediction. In Table 5.4, we summarise the error bars for the optimistic scenario in our fiducial cosmology with RSDs. Since the cut-off for CMB-lensing has previously been close to $0.15 h\text{Mpc}^{-1}$, there are no significant improvements there. In contrast, the error bars from the galaxy power spectrum shrink by a factor of 2 and for the bispectrum by a factor of 3 compared to the best case

Table 5.4 Forecasts for the relative 1σ uncertainties of cosmological parameters in a CMASS-like survey with $k_{\max} = 0.15 h^{-1}\text{Mpc}$ and 6 tomographic bins without (left side) and with (right) CMB prior.

	relative 1σ uncertainties without CMB prior [%]					relative 1σ uncertainties with CMB prior [%]				
	$\delta_g \delta_g$	PS	$\delta_g \delta_g \delta_g$	gal	total	$\delta_g \delta_g$	PS	$\delta_g \delta_g \delta_g$	gal	total
Ω_M	13	11	22	10	9.1	2.1	1.9	2.5	1.9	1.8
Ω_B	50	43	91	41	37	1.1	1.0	1.3	1.0	0.95
h	47	40	85	38	34	0.76	0.70	0.90	0.70	0.66
n_s	31	26	57	24	22	0.42	0.40	0.46	0.41	0.39
σ_8	15	1.3	31	11	1.2	0.88	0.57	0.88	0.88	0.57
b_1	15	0.8	36	11	0.8	0.96	0.68	4.7	0.96	0.67
b_2			80	51	10			63	10	9.9
b_{s^2}			110	56	51			75	50	50

forecasts with the conservative cut-off from Table 5.1b. The combined error bars from an analysis with all spectra shrink by 30%. Adding a CMB prior leads to dramatic improvement for all parameters that can be constrained from CMB observations. This is expected because only the next generation galaxy surveys will contain a comparable information content to current CMB surveys. The CMB prior is based on Appendix A of [228]. Let us stress that CMASS is a galaxy sample from a moderate-sized survey volume and thus constraints from this sample should not be expected to be competitive with Planck constraints. Upcoming surveys will map regions that are 50-100 times larger than CMASS. This is expected to translate into seven- to ten-times tighter parameter constraints.

Table 5.5 contains our PNG forecasts in the optimistic scenario. Since PNGs are best determined on large scales, adding small scale information decreases the constraints only by $\sim 20\%$. Since the large volume of future surveys translates into a smaller k_f , constraints on local and orthogonal type non-Gaussianity are expected to improve by more than the volume related factor of seven to ten mentioned above. This is due to the scale-dependent bias, which dominates on large scales and leads to additional survey volume dependence.

5.4.5 A simple signal compression approach

Following the ideas of section 4.5, we implement a simple signal compression method. Since most bispectrum configurations correlate fields across different galaxy-bins, we reduce the number of galaxy cross-spectra by introducing a maximal correlation length for all spectra that contain two or more galaxy fields ($\langle \delta_g \delta_g \rangle$, $\langle \delta_g \delta_g \kappa \rangle$ and $\langle \delta_g \delta_g \delta_g \rangle$). I.e. we only include configurations in the analysis where the maximal distance between bins is below some

Table 5.5 Forecasts for the 1σ uncertainties of f_{NL} and bias parameters in a CMASS-like survey with $k_{\text{max}} = 0.15 h^{-1}\text{Mpc}$ and 6 tomographic bins.

	$\delta_g \delta_g$	PS	$\delta_g \delta_g \delta_g$	gal	total
$f_{\text{NL}}^{\text{local}}$	57	57	76	45	44
b_1	0.012	0.012	0.11	0.011	0.011
b_2			0.50	0.075	0.075
$b_{s,2}$			0.56	0.38	0.38
$f_{\text{NL}}^{\text{equi}}$			2300	1200	1200
b_1			0.20	0.0088	0.0088
b_2			1.1	0.13	0.13
$b_{s,2}$			0.68	0.40	0.40
$f_{\text{NL}}^{\text{ortho}}$	1750	1750	680	630	620
b_1	0.13	0.13	0.12	0.049	0.048
b_2			0.52	0.15	0.15
$b_{s,2}$			0.56	0.39	0.39

maximal correlation length. Restricting the maximal correlation length reduces the power spectrum scaling (in the number of bins, n) from n^2 to $n \cdot \text{mcl}$ and for the bispectrum n^3 to $n \cdot \text{mcl}^2$ where, mcl is the maximum correlation length in units of the projection depth.

Physically speaking, this approach is sensible, since the correlation between bins is decreasing with increasing distance while the noise terms become constant. This is because the covariance of two cross-spectra contains components proportional to auto-spectra (cf. 4.24, 4.43). Hence, those configurations carry little SN and Fisher information.

The maximal correlation length needed to recover at least 99% of the Fisher information of the full analysis is summarised in Table 5.6. We find that the maximum correlation length is independent of the FoG model used. The bias-error scaling as a function of the maximal correlation length for the PS and total spectra combinations is shown in Fig. 5.10. The same scalings of the $\delta_g \delta_g$ and gal spectra combinations are reported in Fig. 5.11. In both figures, we show the forecasted errors in bold lines and five times the relative biases in dashed lines. The rightmost values in each subplot correspond to the full analysis and as one lowers the maximal correlation length, one observes that the errors and biases first remain constant indicating the intuition that distant cross-bin spectra do not carry much Fisher information. At around a maximal correlation length of $\sim 200 h^{-1}\text{Mpc}$, the errors start increasing. The relative biases decrease too, which is partly driven by the increasing statistical uncertainties. We find that generally, this signal compression approach does not allow us to improve the bias-error trade-offs previously discussed in section 5.4.3. Restricting the maximal correlation length by means of this 99% threshold speeds up the computations

Table 5.6 The maximal correlation lengths needed to recover 99% of the Fisher information for all cosmological parameters compared to the full analysis. The results are independent of the FoG model. All entries are in $h^{-1}\text{Mpc}$.

	$\delta_g \delta_g$	PS	$\delta_g \delta_g \delta_g$	gal	total
n=2	590	590	590	590	590
n=4	295	295	295	295	295
n=6	295	295	295	295	295
n=8	221	221	221	221	221
n=10	177	177	177	177	177
n=12	147	147	196	196	196
n=16	110	110	184	184	184

for 16 bins by a factor of more than 10. For a smaller total number of bins, the speed-ups are smaller.

5.5 Conclusion

In this chapter, we quantified the statistical power of two- and three-point correlators of projected density fields in constraining cosmological parameters and primordial non-Gaussianity. We investigated the trade-off between statistical errors and biases induced by imperfect modelling of non-linear redshift-space distortions, in particular the Finger-of-God effect. We developed an efficient implementation of the projection integrals required to predict projected power spectra and bispectra.

We controlled theoretical uncertainties by computing cut-offs that depend on the projection depth. Those cut-offs are chosen such that the maximal relative biases due to inaccuracies in the matter predictions are less than 20%. This approach allows us to recover more than 80% of the information in bias/amplitude parameters and more than 90% of the information in cosmological parameters in a power spectrum analysis. In a bispectrum analysis, this approach allows us to recover 70% of the 3D Fisher information.

Next, we studied the relation between FoG modelling and relative biases in the forecasted parameters for a CMASS-like survey. We found that the resulting biases are independent of whether one over- or underestimates the FoG damping. We found that not modelling the FoG effect inflates error bars by 20% when controlling for biases. A model that underestimates the velocity dispersion by 50% leads to an increase of 10% of the errors and a model whose velocity dispersion differs by 10% is able to recover the full information while maintaining all relative biases smaller than 20%.

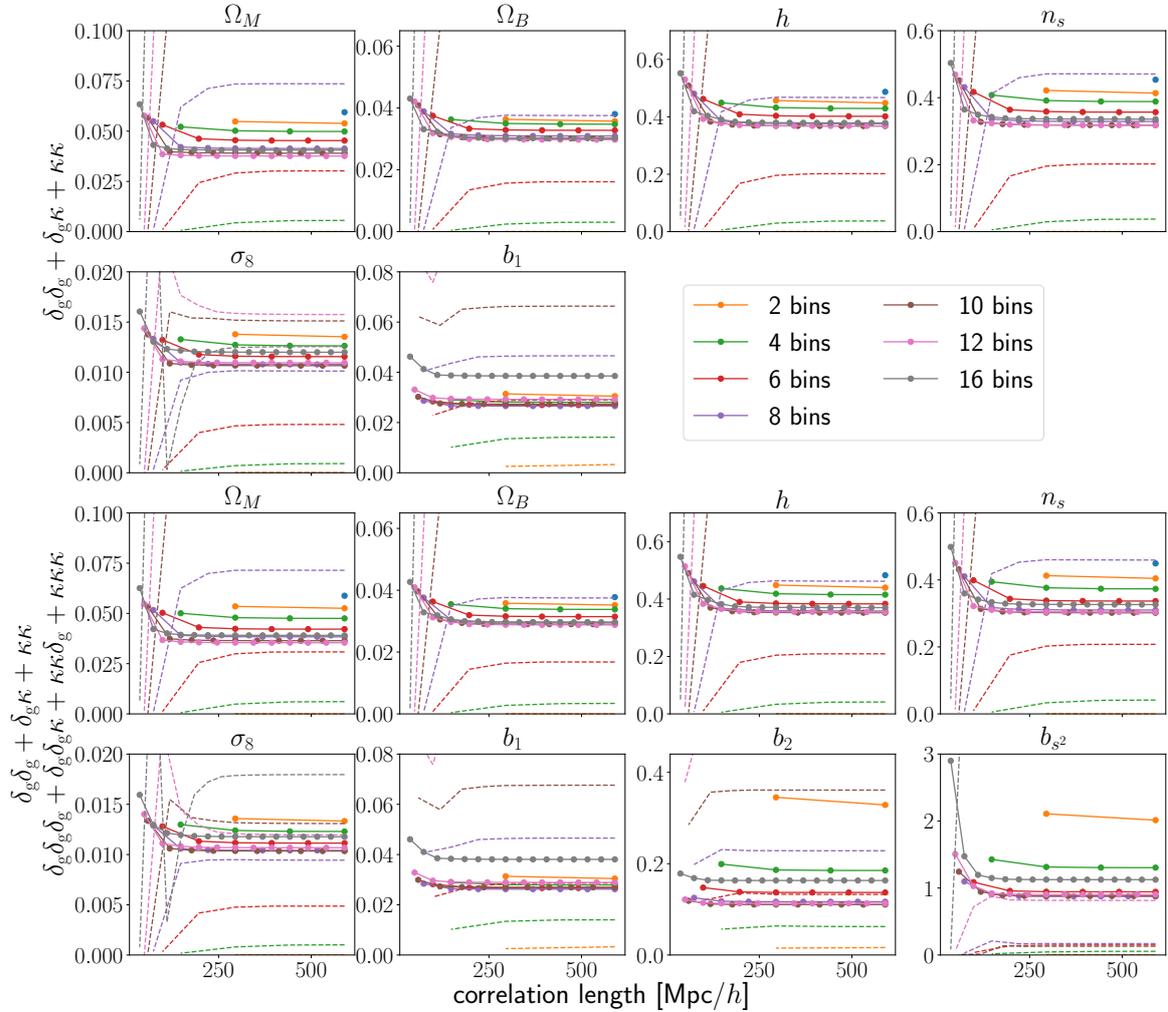


Fig. 5.10 Error bars (solid lines) and five times the biases (dashed lines) for joint clustering-lensing spectra as a function of the maximal correlation length using FoG Model 2. The rightmost values in each subplot correspond to the full analysis and the leftmost values to an analysis with only the auto-spectra. The intersection of dashed and solid lines marks the point where the relative bias becomes 20%. The top two rows correspond to a power spectrum analysis and the bottom two rows show an analysis with all power spectra and bispectra. Each subplot shows the scaling for a different parameter.

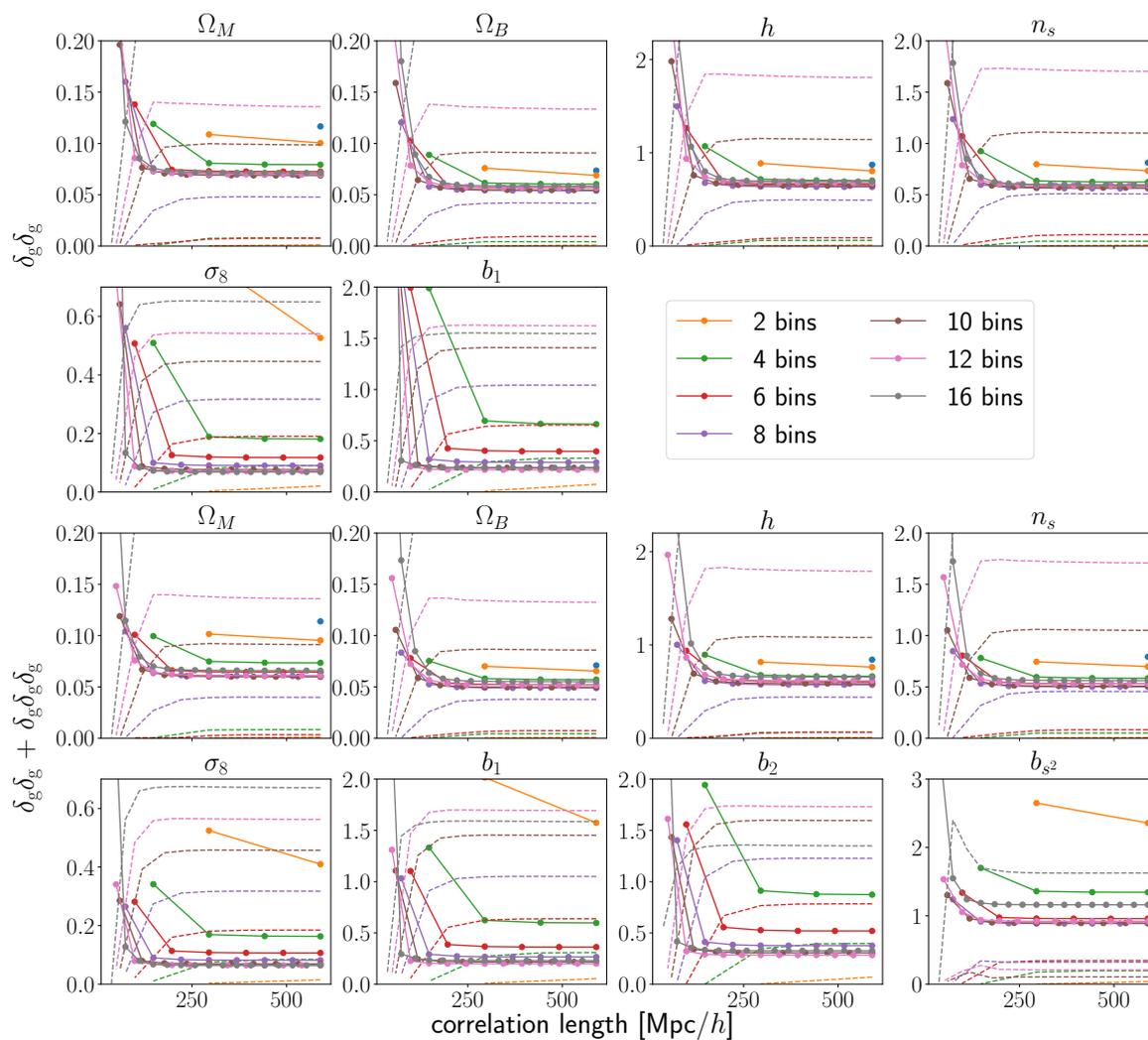


Fig. 5.11 Similar as Fig. 5.10 but for the $\delta_g \delta_g$ (top two rows) and gal combination (bottom two rows) of spectra.

We performed a similar analysis for PNGs of the local, equilateral and orthogonal shape. Here, the necessity to model the FoG effect depends crucially on whether or not one best constrains the PNGs from the scale-dependent bias or from the template. Whereas in the former case, one can recover most of the information without modelling non-linear RSDs, one can improve the error bars by 10% and 20% for the template dominated orthogonal and equilateral shape respectively with an accurate FoG model.

In agreement with [213] we find that the power spectrum in redshift-space does not suffer as severely from the bias-amplitude degeneracies as in real-space. Accordingly, the potential of the bispectrum to further reduce those degeneracies is diminished.

In a more optimistic scenario where one can control systematic uncertainties up to $k_{\max} = 0.15 h^{-1} \text{Mpc}$, one can expect further improvements of more than 100% for spectra that contain galaxy clustering information. This translates into 30% improvements of combined (PS, total) error bars. Let us stress that future surveys would further improve constraints by 700-1000% due to the much larger volumes enabling LSS constraints as tight as those provided by Planck. With these surveys, the projected power spectrum and bispectrum provide a conservative yet powerful analysis toolkit.

Lastly, we studied the impact of dropping cross-bin galaxy correlations. We find that this is not a tool to improve the bias-error trade-off. However, it is possible to reduce the number of clustering bispectrum configurations considerably without losing much Fisher information. In practice, we were able to speed up the most numerically-demanding configurations by a factor of more than 10 while losing less than 1% of Fisher information in all parameters. This indicates that this simple idea could be used as a preprocessing step of a more extensive signal compression approach.

Throughout this chapter, we made several simplifying assumptions that could be lifted in future work. For instance, one could use the FFTLog algorithm to go beyond the flat sky approximation used here [226]. One could also improve the theoretical modelling by taking more orders of the perturbative expansion into account in order to push k_{\max} higher and closer to the more optimistic value mentioned above (see section 5.4.4 and Tables 5.4, 5.5).

Chapter 6

Conclusion

The tightest constraints on cosmological parameters come largely from the two-point analysis of CMB temperature anisotropies and galaxy clustering. While this was sufficient to extract most information from the past CMB and LSS data sets, future LSS surveys will measure a vast amount of non-Gaussian information too. To fully utilise this information, it is necessary to go confidently beyond two-point statistics. In this thesis, we have studied two such statistics that capture non-Gaussian information while being robust against RSDs, one of the leading sources of systematic uncertainties.

Summary

In chapter 3, we presented a first application of the counts-in-cells (CIC) statistics to neutral hydrogen (HI). The CIC statistics combine spherical collapse with large deviation theory to predict the (joint) PDF of densities in spheres on non-linear scales that cannot be accessed perturbatively. Our results are particularly exciting because future intensity-mapping surveys will map vast volumes which will give access to accurate measurements of the PDF, in particular of the tails that are sensitive to dark energy and the total neutrino masses, for instance. We compared two different bias functions for the relation between neutral hydrogen and dark matter and proposed to use a second-order approximation to a non-parametric bias function based on cumulative distribution functions. Using this bias relation, we were able to establish a fully predictive model for the HI density in spheres PDF that matches the measured HI PDF in the IllustrisTNG simulation at a few percent accuracy at scale $R = 5 h^{-1} \text{Mpc}$ from redshift $z=5$ to $z=1$. Furthermore, the bias relation allowed us to predict a density-dependent HI-clustering signal that is consistent with simulations. Without further modelling, we find that the bias relation is also able to capture RSDs and predictions only slightly worsen.

Working with projected power spectra and bispectra poses several challenges. Firstly, for a tomographic galaxy clustering analysis, the amount of resolved radial information is (roughly) given by the inverse projection depth and not the cut-off of the perpendicular modes. This leads to inconsistencies when comparing 2D analyses with varying number of tomographic bins. Theoretical uncertainties would offer a solution to the problem which is why we discussed in chapter 4 a consistent way to project theoretical uncertainties from 3D to 2D. For the theoretical uncertainties of the projected power spectrum, one can separate the required projection integrals which allows us to cache subresults and makes the problem tractable. We use this model to consistently compare 3D and 2D power spectrum Fisher forecasts by formally including all modes into both analysis. This enabled us to empirically validate the model and we found, for instance, that for projection depths of $\sim 10 h^{-1} \text{Mpc}$, one can recover 99% of the 3D power spectrum Fisher information. For the theoretical uncertainties of the projected bispectrum, no such approximation exists and they remain intractable.

In chapter 4, we presented also a fast method to compute projected bispectra in the flat sky approximation. The core idea of the implementation is to separate the multi-dimensional projection integrals and cache reoccurring subresults. Caching essentially simplifies the required projection integrals from 2D to 1D which makes the problem tractable. For the projected galaxy bispectrum, one can further reduce the number of required computations from cubic to quadratic in the number of tomographic bins in case the window functions have the same shape and are equidistantly spaced. For cases where integrals cannot be solved by the FFT method, we introduce analytical results that can be used to approximate the integrals.

In chapter 5, we investigated the relation between biases in the estimated parameters and inaccurate modelling of non-linear redshift-space distortions for the power spectrum and bispectrum of projected galaxy density fields and lensing convergence using the code introduced in chapter 4. We control the amount of small-scale information that enters the analysis by means of a projection depth dependent cut-off $k_{\text{max}}^{2\text{D}}$. We define it by requiring that the relative biases arising from the perturbative matter predictions compared to a fully non-linear model are below 20%. Those cut-offs allow us to recover up to 90% and 66% of the 3D information of the power spectrum and bispectrum respectively in a CMASS-like survey. Using these cut-offs, we find that non-linear redshift-space distortion become only relevant for bins thinner than $150 h^{-1} \text{Mpc}$. Access to higher radial resolution and a sufficiently accurate non-linear RSD model improves constraints by up to 20% when using bins of size $60 h^{-1} \text{Mpc}$. The improvement for local non-Gaussianities is smaller, since it is best constrained from large scales. Even thinner redshift bins lead to worse constraints since the cut-off scale decreases dramatically. As a first step to shrink the bispectrum signal vector,

we introduced a simple compression technique - discarding cross-correlations of distant bins - that allows us to speed up the most demanding computations by up to a factor 10 while still recovering 99% of the available Fisher information.

Outlook

Confronting real data with the CIC formalism would be very interesting. It has been shown that the matter CIC statistic at different scales and redshifts is highly complementary to the matter power spectrum and could thus tighten constraints on cosmological parameters and the total neutrino mass considerably [231]. Moreover, a novel way to constrain PNGs from the bulk of the joint matter CIC PDF at two scales was introduced in [232]. In both cases, the next step towards confronting data would be the inclusion of a bias model for the observed tracers. In the analysis, one could then use the scale-dependent clustering signal to break the new amplitude-bias degeneracies. Another way to break the amplitude-bias degeneracies would be a joint analysis of the CMB-lensing convergence in cylinders and tracer densities in spheres that can be predicted via the CIC formalism. In addition, considerations are needed to take observational effects (e.g. foreground cleaning) into account.

Using the projected bispectra code presented in chapter 4 to analyze real data would provide conservative yet powerful constraints for cosmological parameters and PNGs. In particular, a joint galaxy-clustering CMB-lensing analysis of two- and three-point correlators would break all parameter degeneracies between cosmological and nuisance parameters and thus provide tight constraints. As shown in chapter 5, tuning the projection depth would allow us to control biases from non-linear redshift-space distortions from the galaxy-clustering. However, to perform this analysis, several challenges have to be overcome: The analytic results need to be extended to be able to work with realistic galaxy selection functions, the flat sky approximation has to be relaxed in order to support actual survey geometries and, importantly, the data compression method needs considerable improvements. As a first step towards data compression, our simple but physically motivated method described in chapter 5 worked very well but other simple ideas such as increasing the k -space binning or ignoring certain configurations of the auto-bispectra are likely to become very lossy. A more promising route is the extension of the MODAL framework [233, 234]. There, one would need to find a set of 3D eigen-bispectra that can be easily projected while capturing most of the bispectrum signal in 3D and 2D independently of the projection depth and kernel. This set of eigen-bispectra would simplify projections greatly since one only needs to extract the coefficients of the eigen-bispectra in 2D and can directly translate those to the corresponding 3D eigen-bispectra. Alternatively, one could think about reconstructing a 3D bispectrum from the observed projected spectra in order to perform the analysis there.

As an intermediate step towards a full cosmological analysis, one could perform a MCMC analysis for amplitude-like parameters such as the bias parameters, σ_8 and f_{NL} assuming a Gaussian covariance. However, to go beyond the Gaussian covariance would also require a drastic reduction of the size of the bispectrum signal vector due to the needed simulations to estimate empirical covariances.

References

- [1] O. Leicht, C. Uhlemann, F. Villaescusa-Navarro, S. Codis, L. Hernquist and S. Genel, *Extreme spheres: counts-in-cells for 21cm intensity mapping*, **484** (2019) 269 [1808.09968].
- [2] O. Leicht, T. Baldauf, J. Fergusson and P. Shellard, *Projected two- and three-point statistics: Forecasts and mitigation of non-linear RSDs*, *arXiv e-prints* (2020) arXiv:2008.10199 [2008.10199].
- [3] M. Betoule, R. Kessler, J. Guy, J. Mosher, D. Hardin, R. Biswas et al., *Improved cosmological constraints from a joint analysis of the SDSS-II and SNLS supernova samples*, **568** (2014) A22 [1401.4064].
- [4] R. E. Smith, J. A. Peacock, A. Jenkins, S. D. M. White, C. S. Frenk, F. R. Pearce et al., *Stable clustering, the halo model and non-linear cosmological power spectra*, **341** (2003) 1311 [astro-ph/0207664].
- [5] J. A. Peacock and R. E. Smith, *HALOFIT: Nonlinear distribution of cosmological mass and galaxies*, Feb., 2014.
- [6] J. A. Peacock, S. Cole, P. Norberg, C. M. Baugh, J. Bland-Hawthorn, T. Bridges et al., *A measurement of the cosmological mass density from clustering in the 2dF Galaxy Redshift Survey*, **410** (2001) 169 [astro-ph/0103143].
- [7] C. Uhlemann, S. Codis, C. Pichon, F. Bernardeau and P. Reimberg, *Back in the saddle: large-deviation statistics of the cosmic log-density field*, **460** (2016) 1529 [1512.05793].
- [8] S. Codis, F. Bernardeau and C. Pichon, *The large-scale correlations of multicell densities and profiles: implications for cosmic variance estimates*, **460** (2016) 1598 [1602.03562].
- [9] A. Friedmann, *Über die Krümmung des Raumes*, *Zeitschrift für Physik* **10** (1922) 377.
- [10] G. Lemaître, *Un Univers homogène de masse constante et de rayon croissant rendant compte de la vitesse radiale des nébuleuses extra-galactiques*, *Annales de la Sociéacute;teacute; Scientifique de Bruxelles* **47** (1927) 49.
- [11] E. Hubble, *A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae*, *Proceedings of the National Academy of Science* **15** (1929) 168.
- [12] A. A. Penzias and R. W. Wilson, *A Measurement of Excess Antenna Temperature at 4080 Mc/s.*, **142** (1965) 419.

- [13] A. H. Guth, *Inflationary universe: A possible solution to the horizon and flatness problems*, **23** (1981) 347.
- [14] F. Zwicky, *Die Rotverschiebung von extragalaktischen Nebeln*, *Helvetica Physica Acta* **6** (1933) 110.
- [15] V. C. Rubin and J. Ford, W. Kent, *Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions*, **159** (1970) 379.
- [16] V. C. Rubin, J. Ford, W. K. and N. Thonnard, *Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605 ($R=4kpc$) to UGC 2885 ($R=122kpc$)*, **238** (1980) 471.
- [17] A. G. Riess, A. V. Filippenko, P. Challis, A. Clocchiatti, A. Diercks, P. M. Garnavich et al., *Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant*, **116** (1998) 1009 [astro-ph/9805201].
- [18] S. Perlmutter, G. Aldering, G. Goldhaber, R. A. Knop, P. Nugent, P. G. Castro et al., *Measurements of Ω and Λ from 42 High-Redshift Supernovae*, **517** (1999) 565 [astro-ph/9812133].
- [19] E. Komatsu, K. M. Smith, J. Dunkley, C. L. Bennett, B. Gold, G. Hinshaw et al., *Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation*, **192** (2011) 18 [1001.4538].
- [20] Planck Collaboration, N. Aghanim, Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi et al., *Planck 2018 results. VI. Cosmological parameters*, *arXiv e-prints* (2018) arXiv:1807.06209 [1807.06209].
- [21] A. G. Riess, L. M. Macri, S. L. Hoffmann, D. Scolnic, S. Casertano, A. V. Filippenko et al., *A 2.4% Determination of the Local Value of the Hubble Constant*, **826** (2016) 56 [1604.01424].
- [22] S. Alam, M. Ata, S. Bailey, F. Beutler, D. Bizyaev, J. A. Blazek et al., *The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: cosmological analysis of the DR12 galaxy sample*, **470** (2017) 2617 [1607.03155].
- [23] C. Dvorkin, M. Gerbino, D. Alonso, N. Battaglia, S. Bird, A. Diaz Rivero et al., *Neutrino Mass from Cosmology: Probing Physics Beyond the Standard Model*, **51** (2019) 64 [1903.03689].
- [24] C. Mahony, B. Leistedt, H. V. Peiris, J. Braden, B. Joachimi, A. Korn et al., *Target neutrino mass precision for determining the neutrino hierarchy*, **101** (2020) 083513 [1907.04331].
- [25] P. D. Meerburg, D. Green, R. Flauger, B. Wallisch, M. C. D. Marsh, E. Pajer et al., *Primordial Non-Gaussianity*, **51** (2019) 107 [1903.04409].
- [26] LSST Science Collaboration, P. A. Abell, J. Allison, S. F. Anderson, J. R. Andrew, J. R. P. Angel et al., *LSST Science Book, Version 2.0*, *arXiv e-prints* (2009) arXiv:0912.0201 [0912.0201].

- [27] O. Doré, J. Bock, M. Ashby, P. Capak, A. Cooray, R. de Putter et al., *Cosmology with the SPHEREX All-Sky Spectral Survey*, *arXiv e-prints* (2014) arXiv:1412.4872 [1412.4872].
- [28] L. Amendola, S. Appleby, A. Avgoustidis, D. Bacon, T. Baker, M. Baldi et al., *Cosmology and fundamental physics with the Euclid satellite*, *Living Reviews in Relativity* **21** (2018) 2 [1606.00180].
- [29] DESI Collaboration, A. Aghamousa, J. Aguilar, S. Ahlen, S. Alam, L. E. Allen et al., *The DESI Experiment Part I: Science, Targeting, and Survey Design*, *arXiv e-prints* (2016) arXiv:1611.00036 [1611.00036].
- [30] X. Chen, *The Tianlai Project: a 21CM Cosmology Experiment*, in *International Journal of Modern Physics Conference Series*, vol. 12 of *International Journal of Modern Physics Conference Series*, pp. 256–263, Mar., 2012, DOI [1212.6278].
- [31] R. A. Battye, I. W. A. Browne, C. Dickinson, G. Heron, B. Maffei and A. Pourtsidou, *HI intensity mapping: a single dish approach*, **434** (2013) 1239 [1209.0343].
- [32] K. Bandura, G. E. Addison, M. Amiri, J. R. Bond, D. Campbell-Wilson, L. Connor et al., *Canadian Hydrogen Intensity Mapping Experiment (CHIME) pathfinder*, in , vol. 9145 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, p. 914522, July, 2014, DOI [1406.2288].
- [33] R. Nan, D. Li, C. Jin, Q. Wang, L. Zhu, W. Zhu et al., *The Five-Hundred Aperture Spherical Radio Telescope (fast) Project*, *International Journal of Modern Physics D* **20** (2011) 989 [1105.3794].
- [34] L. B. Newburgh, K. Bandura, M. A. Bucher, T. C. Chang, H. C. Chiang, J. F. Cliche et al., *HIRAX: a probe of dark energy and radio transients*, in , vol. 9906 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, p. 99065X, Aug., 2016, DOI [1607.02059].
- [35] A. Pourtsidou, *HI Intensity Mapping with MeerKAT*, *arXiv e-prints* (2017) arXiv:1709.07316 [1709.07316].
- [36] M. Santos, P. Bull, D. Alonso, S. Camera, P. Ferreira, G. Bernardi et al., *Cosmology from a SKA HI intensity mapping survey*, in *Advancing Astrophysics with the Square Kilometre Array (AASKA14)*, p. 19, Apr., 2015, 1501.03989.
- [37] P. Ade, J. Aguirre, Z. Ahmed, S. Aiola, A. Ali, D. Alonso et al., *The Simons Observatory: science goals and forecasts*, **2019** (2019) 056 [1808.07445].
- [38] K. Abazajian, G. Addison, P. Adshead, Z. Ahmed, S. W. Allen, D. Alonso et al., *CMB-S4 Science Case, Reference Design, and Project Plan*, *arXiv e-prints* (2019) arXiv:1907.04473 [1907.04473].
- [39] H. Reall, *Part 3 general relativity lecture notes*, 2017.
- [40] A. Einstein, *Die Feldgleichungen der Gravitation*, *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin)* (1915) 844.

- [41] A. Einstein, *Kosmologische Betrachtungen zur allgemeinen Relativitätstheorie*, *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin)* (1917) 142.
- [42] R. S. Gonçalves, G. C. Carvalho, C. A. P. Bengaly, J. C. Carvalho and J. S. Alcaniz, *Measuring the scale of cosmic homogeneity with SDSS-IV DR14 quasars*, **481** (2018) 5270 [1809.11125].
- [43] G. Lemaître, *L'Univers en expansion*, *Annales de la Socié;tété Scientifique de Bruxelles* **53** (1933) 51.
- [44] H. P. Robertson, *Kinematics and World-Structure*, **82** (1935) 284.
- [45] H. P. Robertson, *Kinematics and World-Structure II.*, **83** (1936) 187.
- [46] H. P. Robertson, *Kinematics and World-Structure III.*, **83** (1936) 257.
- [47] A. G. Walker, *On Milne's Theory of World-Structure*, *Proceedings of the London Mathematical Society* **42** (1937) 90.
- [48] B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley et al., *GW151226: Observation of Gravitational Waves from a 22-Solar-Mass Binary Black Hole Coalescence*, **116** (2016) 241103 [1606.04855].
- [49]
- [50] D. Baumann, *TASI Lectures on Inflation*, *arXiv e-prints* (2009) arXiv:0907.5424 [0907.5424].
- [51] A. D. Linde, *A new inflationary universe scenario: A possible solution of the horizon, flatness, homogeneity, isotropy and primordial monopole problems*, *Physics Letters B* **108** (1982) 389.
- [52] A. Albrecht and P. J. Steinhardt, *Cosmology for Grand Unified Theories with Radiatively Induced Symmetry Breaking*, **48** (1982) 1220.
- [53] V. Desjacques and U. Seljak, *Primordial non-Gaussianity from the large-scale structure*, *Classical and Quantum Gravity* **27** (2010) 124011 [1003.5020].
- [54] Planck Collaboration, Y. Akrami, F. Arroja, M. Ashdown, J. Aumont, C. Baccigalupi et al., *Planck 2018 results. IX. Constraints on primordial non-Gaussianity*, *arXiv e-prints* (2019) arXiv:1905.05697 [1905.05697].
- [55] J. R. Fergusson, D. M. Regan and E. P. S. Shellard, *Rapid separable analysis of higher order correlators in large-scale structure*, **86** (2012) 063511 [1008.1730].
- [56] A. Gangui, F. Lucchin, S. Matarrese and S. Mollerach, *The Three-Point Correlation Function of the Cosmic Microwave Background in Inflationary Models*, **430** (1994) 447 [astro-ph/9312033].
- [57] L. Verde, L. Wang, A. F. Heavens and M. Kamionkowski, *Large-scale structure, the cosmic microwave background and primordial non-Gaussianity*, **313** (2000) 141 [astro-ph/9906301].

- [58] E. Komatsu and D. N. Spergel, *Acoustic signatures in the primary microwave background bispectrum*, **63** (2001) 063002 [astro-ph/0005036].
- [59] J. Maldacena, *Non-gaussian features of primordial fluctuations in single field inflationary models*, *Journal of High Energy Physics* **2003** (2003) 013 [astro-ph/0210603].
- [60] P. Creminelli and M. Zaldarriaga, *A single-field consistency relation for the three-point function*, **2004** (2004) 006 [astro-ph/0407059].
- [61] Planck Collaboration, P. A. R. Ade, N. Aghanim, C. Armitage-Caplan, M. Arnaud, M. Ashdown et al., *Planck 2013 results. XXIV. Constraints on primordial non-Gaussianity*, **571** (2014) A24 [1303.5084].
- [62] P. Creminelli, A. Nicolis, L. Senatore, M. Tegmark and M. Zaldarriaga, *Limits on non-Gaussianities from WMAP data*, **2006** (2006) 004 [astro-ph/0509029].
- [63] L. Senatore, K. M. Smith and M. Zaldarriaga, *Non-Gaussianities in single field inflation and their optimal limits from the WMAP 5-year data*, **2010** (2010) 028 [0905.3746].
- [64] A. Lewis, A. Challinor and A. Lasenby, *Efficient Computation of Cosmic Microwave Background Anisotropies in Closed Friedmann-Robertson-Walker Models*, **538** (2000) 473 [astro-ph/9911177].
- [65] A. Challinor and A. Lewis, *Linear power spectrum of observed source number counts*, **84** (2011) 043516 [1105.5292].
- [66] J. Lesgourgues, *The Cosmic Linear Anisotropy Solving System (CLASS) I: Overview*, *arXiv e-prints* (2011) arXiv:1104.2932 [1104.2932].
- [67] F. Bernardeau, S. Colombi, E. Gaztañaga and R. Scoccimarro, *Large-scale structure of the Universe and cosmological perturbation theory*, **367** (2002) 1 [astro-ph/0112551].
- [68] V. Desjacques, D. Jeong and F. Schmidt, *Large-scale galaxy bias*, **733** (2018) 1 [1611.09787].
- [69] W. L. W. Sargent and E. L. Turner, *A statistical method for determining the cosmological density parameter from the redshifts of a complete sample of galaxies.*, **212** (1977) L3.
- [70] N. Kaiser, *Clustering in real space and in redshift space*, **227** (1987) 1.
- [71] R. Scoccimarro, H. M. P. Couchman and J. A. Frieman, *The Bispectrum as a Signature of Gravitational Instability in Redshift Space*, **517** (1999) 531 [astro-ph/9808305].
- [72] J. C. Jackson, *A critique of Rees's theory of primordial gravitational radiation*, **156** (1972) 1P [0810.3908].
- [73] J. A. Peacock, *Statistics of Cosmological Density Fields*, in *New Insights into the Universe*, V. J. Martinez, M. Portilla and D. Saez, eds., vol. 408, p. 1, (1992), DOI.

- [74] P. J. E. Peebles, *A Cosmic Virial Theorem*, **45** (1976) 3.
- [75] C. Park, M. S. Vogeley, M. J. Geller and J. P. Huchra, *Power Spectrum, Correlation Function, and Tests for Luminosity Bias in the CfA Redshift Survey*, **431** (1994) 569.
- [76] J. A. Peacock and S. J. Dodds, *Reconstructing the Linear Power Spectrum of Cosmological Mass Fluctuations*, **267** (1994) 1020 [astro-ph/9311057].
- [77] W. E. Ballinger, J. A. Peacock and A. F. Heavens, *Measuring the cosmological constant with redshift surveys*, **282** (1996) 877 [astro-ph/9605017].
- [78] H. Gil-Marín, C. Wagner, J. Noreña, L. Verde and W. Percival, *Dark matter and halo bispectrum in redshift space: theory and applications*, **2014** (2014) 029 [1407.1836].
- [79] H.-J. Seo and D. J. Eisenstein, *Probing Dark Energy with Baryonic Acoustic Oscillations from Future Large Galaxy Redshift Surveys*, **598** (2003) 720 [astro-ph/0307460].
- [80] T. Matsubara, *Resumming cosmological perturbations via the Lagrangian picture: One-loop results in real space and in redshift space*, **77** (2008) 063530 [0711.2521].
- [81] T. Matsubara, *Nonlinear perturbation theory with halo bias and redshift-space distortions via the Lagrangian picture*, **78** (2008) 083519 [0807.1733].
- [82] A. Taruya, T. Nishimichi and S. Saito, *Baryon acoustic oscillations in 2D: Modeling redshift-space power spectrum from perturbation theory*, **82** (2010) 063522 [1006.0699].
- [83] P. McDonald and A. Roy, *Clustering of dark matter tracers: generalizing bias for the coming era of precision LSS*, **2009** (2009) 020 [0902.0991].
- [84] T. Baldauf, U. Seljak, V. Desjacques and P. McDonald, *Evidence for quadratic tidal tensor bias from the halo bispectrum*, **86** (2012) 083540 [1201.4827].
- [85] K. C. Chan, R. Scoccimarro and R. K. Sheth, *Gravity and large-scale nonlocal bias*, **85** (2012) 083509 [1201.3614].
- [86] T. Lazeyras and F. Schmidt, *Beyond LIMD bias: a measurement of the complete set of third-order halo bias parameters*, **2018** (2018) 008 [1712.07531].
- [87] M. M. Abidi and T. Baldauf, *Cubic halo bias in Eulerian and Lagrangian space*, **2018** (2018) 029 [1802.07622].
- [88] R. Scoccimarro, R. K. Sheth, L. Hui and B. Jain, *How Many Galaxies Fit in a Halo? Constraints on Galaxy Formation Efficiency from Spatial Clustering*, **546** (2001) 20 [astro-ph/0006319].
- [89] F. Schmidt, *Towards a self-consistent halo model for the nonlinear large-scale structure*, **93** (2016) 063512 [1511.02231].
- [90] T. Baldauf, U. Seljak, R. E. Smith, N. Hamaus and V. Desjacques, *Halo stochasticity from exclusion and nonlinear clustering*, **88** (2013) 083507 [1305.2917].

- [91] N. Dalal, O. Doré, D. Huterer and A. Shirokov, *Imprints of primordial non-Gaussianities on large-scale structure: Scale-dependent bias and abundance of virialized objects*, **77** (2008) 123514 [0710.4560].
- [92] P. McDonald, *Primordial non-Gaussianity: Large-scale structure signature in the perturbative bias model*, **78** (2008) 123519 [0806.1061].
- [93] L. Verde and S. Matarrese, *Detectability of the Effect of Inflationary Non-Gaussianity on Halo Bias*, **706** (2009) L91 [0909.3224].
- [94] T. Giannantonio and C. Porciani, *Structure formation from non-Gaussian initial conditions: Multivariate biasing, statistics, and comparison with N-body simulations*, **81** (2010) 063530 [0911.0017].
- [95] T. Baldauf, U. Seljak and L. Senatore, *Primordial non-Gaussianity in the bispectrum of the halo density field*, **2011** (2011) 006 [1011.1513].
- [96] V. Assassi, D. Baumann and F. Schmidt, *Galaxy bias and primordial non-Gaussianity*, **2015** (2015) 043 [1510.03723].
- [97] F. Schmidt and M. Kamionkowski, *Halo clustering with nonlocal non-Gaussianity*, **82** (2010) 103002 [1008.0638].
- [98] M. Tellarini, A. J. Ross, G. Tasinato and D. Wands, *Non-local bias in the halo bispectrum with primordial non-Gaussianity*, **2015** (2015) 004 [1504.00324].
- [99] M. Tellarini, A. J. Ross, G. Tasinato and D. Wands, *Galaxy bispectrum, primordial non-Gaussianity and redshift space distortions*, **2016** (2016) 014 [1603.06814].
- [100] F. Bernardeau, *The Gravity-induced Quasi-Gaussian Correlation Hierarchy*, **392** (1992) 1.
- [101] F. Bernardeau, C. Pichon and S. Codis, *Statistics of cosmic density profiles from perturbation theory*, **90** (2014) 103519 [1310.8134].
- [102] C. Lacey and S. Cole, *Merger Rates in Hierarchical Models of Galaxy Formation - Part Two - Comparison with N-Body Simulations*, **271** (1994) 676 [astro-ph/9402069].
- [103] R. K. Sheth and G. Tormen, *Large-scale bias and the peak background split*, **308** (1999) 119 [astro-ph/9901122].
- [104] A. G. Doroshkevich, *The space structure of perturbations and the origin of rotation of galaxies in the theory of fluctuation.*, *Astrofizika* **6** (1970) 581.
- [105] J. M. Bardeen, J. R. Bond, N. Kaiser and A. S. Szalay, *The Statistics of Peaks of Gaussian Random Fields*, **304** (1986) 15.
- [106] R. K. Sheth, H. J. Mo and G. Tormen, *Ellipsoidal collapse and an improved model for the number and spatial distribution of dark matter haloes*, **323** (2001) 1 [astro-ph/9907024].

- [107] J. Tinker, A. V. Kravtsov, A. Klypin, K. Abazajian, M. Warren, G. Yepes et al., *Toward a Halo Mass Function for Precision Cosmology: The Limits of Universality*, **688** (2008) 709 [0803.2706].
- [108] M. LoVerde, A. Miller, S. Shandera and L. Verde, *Effects of scale-dependent non-Gaussianity on cosmological structures*, **2008** (2008) 014 [0711.4126].
- [109] V. Desjacques, D. Jeong and F. Schmidt, *Accurate predictions for the scale-dependent galaxy bias from primordial non-Gaussianity*, **84** (2011) 061301 [1105.3476].
- [110] V. Desjacques, D. Jeong and F. Schmidt, *Non-Gaussian Halo Bias Re-examined: Mass-dependent Amplitude from the Peak-Background Split and Thresholding*, **84** (2011) 063512 [1105.3628].
- [111] F. Schmidt, D. Jeong and V. Desjacques, *Peak-background split, renormalization, and galaxy clustering*, **88** (2013) 023515 [1212.0868].
- [112] D. Karagiannis, A. Lazanu, M. Liguori, A. Raccanelli, N. Bartolo and L. Verde, *Constraining primordial non-Gaussianity with bispectrum and power spectrum from upcoming optical and radio surveys*, **478** (2018) 1341 [1801.09280].
- [113] P. Coles and B. Jones, *A lognormal model for the cosmological mass distribution.*, **248** (1991) 1.
- [114] F. Bernardeau and P. Reimberg, *Large deviation principle at play in large scale structure cosmology*, **94** (2016) 063520 [1511.08641].
- [115] P. Valageas, *Dynamics of gravitational clustering. II. Steepest-descent method for the quasi-linear regime*, **382** (2002) 412 [astro-ph/0107126].
- [116] S. R. S. Varadhan, *Large Deviations and Applications*. SIAM, 1984, 10.1137/1.9781611970241.
- [117] F. den Hollander, *Large Deviations*, vol. 14. AMS, 2000, 10.1090/fim/014.
- [118] C. Uhlemann, S. Codis, J. Kim, C. Pichon, F. Bernardeau, D. Pogosyan et al., *Beyond Kaiser bias: mildly non-linear two-point statistics of densities in distant spheres*, **466** (2017) 2067 [1607.01026].
- [119] N. Kaiser, *On the spatial correlations of Abell clusters.*, **284** (1984) L9.
- [120] F. Bernardeau, *The large-scale gravitational bias from the quasi-linear regime.*, **312** (1996) 11 [astro-ph/9602072].
- [121] O. Doré, M. W. Werner, M. Ashby, P. Banerjee, N. Battaglia, J. Bauer et al., *Science Impacts of the SPHEREx All-Sky Optical to Near-Infrared Spectral Survey: Report of a Community Workshop Examining Extragalactic, Galactic, Stellar and Planetary Science*, *arXiv e-prints* (2016) arXiv:1606.07039 [1606.07039].
- [122] P. Madau, A. Meiksin and M. J. Rees, *21 Centimeter Tomography of the Intergalactic Medium at High Redshift*, **475** (1997) 429 [astro-ph/9608010].

- [123] S. Bharadwaj and S. K. Sethi, *HI Fluctuations at Large Redshifts: I–Visibility correlation*, *Journal of Astrophysics and Astronomy* **22** (2001) 293 [astro-ph/0203269].
- [124] R. A. Battye, R. D. Davies and J. Weller, *Neutral hydrogen surveys for high-redshift galaxy clusters and protoclusters*, **355** (2004) 1339 [astro-ph/0401340].
- [125] R. Barkana and A. Loeb, *A Method for Separating the Physics from the Astrophysics of High-Redshift 21 Centimeter Fluctuations*, **624** (2005) L65 [astro-ph/0409572].
- [126] U.-L. Pen, L. Staveley-Smith, J. B. Peterson and T.-C. Chang, *First detection of cosmic structure in the 21-cm intensity field*, **394** (2009) L6 [0802.3239].
- [127] T.-C. Chang, U.-L. Pen, K. Bandura and J. B. Peterson, *An intensity map of hydrogen 21-cm emission at redshift $z \sim 0.8$* , **466** (2010) 463.
- [128] K. W. Masui, E. R. Switzer, N. Banavar, K. Bandura, C. Blake, L. M. Calin et al., *Measurement of 21 cm Brightness Fluctuations at $z \sim 0.8$ in Cross-correlation*, **763** (2013) L20 [1208.0331].
- [129] J. D. Bowman, A. E. E. Rogers, R. A. Monsalve, T. J. Mozdzen and N. Mahesh, *An absorption profile centred at 78 megahertz in the sky-averaged spectrum*, **555** (2018) 67 [1810.05912].
- [130] E. R. Switzer, K. W. Masui, K. Bandura, L. M. Calin, T. C. Chang, X. L. Chen et al., *Determination of $z \sim 0.8$ neutral hydrogen fluctuations using the 21cm intensity mapping autocorrelation.*, **434** (2013) L46 [1304.3712].
- [131] P. Bull, P. G. Ferreira, P. Patel and M. G. Santos, *Late-time Cosmology with 21 cm Intensity Mapping Experiments*, **803** (2015) 21 [1405.1452].
- [132] A. Pourtsidou, D. Bacon and R. Crittenden, *HI and cosmological constraints from intensity mapping, optical and CMB surveys*, **470** (2017) 4251 [1610.04189].
- [133] E. D. Kovetz, M. P. Viero, A. Lidz, L. Newburgh, M. Rahman, E. Switzer et al., *Line-Intensity Mapping: 2017 Status Report*, *arXiv e-prints* (2017) arXiv:1709.09066 [1709.09066].
- [134] F. Villaescusa-Navarro, P. Bull and M. Viel, *Weighing Neutrinos with Cosmic Neutral Hydrogen*, **814** (2015) 146 [1507.05102].
- [135] F. Villaescusa-Navarro, M. Viel, K. K. Datta and T. R. Choudhury, *Modeling the neutral hydrogen distribution in the post-reionization Universe: intensity mapping*, **2014** (2014) 050 [1405.6713].
- [136] I. P. Carucci, F. Villaescusa-Navarro, M. Viel and A. Lapi, *Warm dark matter signatures on the 21cm power spectrum: intensity mapping forecasts for SKA*, **2015** (2015) 047 [1502.06961].
- [137] A. Obuljen, E. Castorina, F. Villaescusa-Navarro and M. Viel, *High-redshift post-reionization cosmology with 21cm intensity mapping*, **2018** (2018) 004 [1709.07893].

- [138] I. P. Carucci, F. Villaescusa-Navarro and M. Viel, *The cross-correlation between 21 cm intensity mapping maps and the Ly α forest in the post-reionization era*, **2017** (2017) 001 [1611.07527].
- [139] F. Villaescusa-Navarro, M. Viel, D. Alonso, K. K. Datta, P. Bull and M. G. Santos, *Cross-correlating 21cm intensity maps with Lyman Break Galaxies in the post-reionization era*, **2015** (2015) 034 [1410.7393].
- [140] T.-C. Chang, U.-L. Pen, J. B. Peterson and P. McDonald, *Baryon Acoustic Oscillation Intensity Mapping of Dark Energy*, **100** (2008) 091303 [0709.3672].
- [141] F. Villaescusa-Navarro, D. Alonso and M. Viel, *Baryonic acoustic oscillations from 21 cm intensity mapping: the Square Kilometre Array case*, **466** (2017) 2736 [1609.00019].
- [142] I. Harrison, S. Camera, J. Zuntz and M. L. Brown, *SKA weak lensing - I. Cosmological forecasts and the power of radio-optical cross-correlations*, **463** (2016) 3674 [1601.03947].
- [143] A. Bonaldi, I. Harrison, S. Camera and M. L. Brown, *SKA weak lensing- II. Simulated performance and survey design considerations*, **463** (2016) 3686 [1601.03948].
- [144] S. Foreman, P. D. Meerburg, A. van Engelen and J. Meyers, *Lensing reconstruction from line intensity maps: the impact of gravitational nonlinearity*, **2018** (2018) 046 [1803.04975].
- [145] E. Schaan, S. Ferraro and D. N. Spergel, *Weak lensing of intensity mapping: The cosmic infrared background*, **97** (2018) 123539 [1802.05706].
- [146] R. Scoccimarro, M. Zaldarriaga and L. Hui, *Power Spectrum Correlations Induced by Nonlinear Clustering*, **527** (1999) 1 [astro-ph/9901099].
- [147] M. Schmittfull, Y. Feng, F. Beutler, B. Sherwin and M. Y. Chu, *Eulerian BAO reconstructions and N-point statistics*, **92** (2015) 123522 [1508.06972].
- [148] G. J. A. Harker, S. Zaroubi, R. M. Thomas, V. Jelić, P. Labropoulos, G. Mellema et al., *Detection and extraction of signals from the epoch of reionization using higher-order one-point statistics*, **393** (2009) 1449 [0809.2428].
- [149] K. Ichikawa, R. Barkana, I. T. Iliev, G. Mellema and P. R. Shapiro, *Measuring the history of cosmic reionization using the 21-cm probability distribution function from simulations*, **406** (2010) 2521 [0907.2932].
- [150] S. Baek, B. Semelin, P. Di Matteo, Y. Revaz and F. Combes, *Reionization by UV or X-ray sources*, **523** (2010) A4 [1003.0834].
- [151] H. Shimabukuro, S. Yoshiura, K. Takahashi, S. Yokoyama and K. Ichiki, *Studying 21cm power spectrum with one-point statistics*, **451** (2015) 467 [1412.3332].
- [152] P. Kittiwisit, J. D. Bowman, D. C. Jacobs, A. P. Beardsley and N. Thyagarajan, *Sensitivity of the Hydrogen Epoch of Reionization Array and its build-out stages to one-point statistics from redshifted 21 cm observations*, **474** (2018) 4487 [1708.00036].

- [153] S. Codis, C. Pichon, F. Bernardeau, C. Uhlemann and S. Prunet, *Encircling the dark: constraining dark energy via cosmic density in spheres*, **460** (2016) 1549 [1603.03347].
- [154] I. Szapudi and J. Pan, *On Recovering the Nonlinear Bias Function from Counts-in-Cells Measurements*, **602** (2004) 26 [astro-ph/0308525].
- [155] M. Manera and E. Gaztañaga, *The local bias model in the large-scale halo distribution*, **415** (2011) 383 [0912.0446].
- [156] C. Uhlemann, E. Pajer, C. Pichon, T. Nishimichi, S. Codis and F. Bernardeau, *Hunting high and low: disentangling primordial and late-time non-Gaussianity with cosmic densities in spheres*, **474** (2018) 2853 [1708.02206].
- [157] O. Friedrich, D. Gruen, J. DeRose, D. Kirk, E. Krause, T. McClintock et al., *Density split statistics: Joint model of counts and lensing in cells*, **98** (2018) 023508 [1710.05162].
- [158] A. I. Salvador, F. J. Sánchez, A. Pagul, J. García-Bellido, E. Sanchez, A. Pujol et al., *Measuring linear and non-linear galaxy bias using counts-in-cells in the Dark Energy Survey Science Verification data*, **482** (2019) 1435 [1807.10331].
- [159] C. Uhlemann, M. Feix, S. Codis, C. Pichon, F. Bernardeau, B. L’Huillier et al., *A question of separation: disentangling tracer bias and gravitational non-linearity with counts-in-cells statistics*, **473** (2018) 5098 [1705.08901].
- [160] E. Castorina and F. Villaescusa-Navarro, *On the spatial distribution of neutral hydrogen in the Universe: bias and shot-noise of the H I power spectrum*, **471** (2017) 1788 [1609.05157].
- [161] F. Villaescusa-Navarro, S. Genel, E. Castorina, A. Obuljen, D. N. Spergel, L. Hernquist et al., *Ingredients for 21 cm Intensity Mapping*, **866** (2018) 135 [1804.09180].
- [162] P. C. Breysse, E. D. Kovetz, P. S. Behroozi, L. Dai and M. Kamionkowski, *Insights from probability distribution functions of intensity maps*, **467** (2017) 2996 [1609.01728].
- [163] A. Pillepich, D. Nelson, L. Hernquist, V. Springel, R. Pakmor, P. Torrey et al., *First results from the IllustrisTNG simulations: the stellar mass content of groups and clusters of galaxies*, **475** (2018) 648 [1707.03406].
- [164] V. Springel, R. Pakmor, A. Pillepich, R. Weinberger, D. Nelson, L. Hernquist et al., *First results from the IllustrisTNG simulations: matter and galaxy clustering*, **475** (2018) 676 [1707.03397].
- [165] D. Nelson, A. Pillepich, V. Springel, R. Weinberger, L. Hernquist, R. Pakmor et al., *First results from the IllustrisTNG simulations: the galaxy colour bimodality*, **475** (2018) 624 [1707.03395].

- [166] F. Marinacci, M. Vogelsberger, R. Pakmor, P. Torrey, V. Springel, L. Hernquist et al., *First results from the IllustrisTNG simulations: radio haloes and magnetic fields*, **480** (2018) 5113 [1707.03396].
- [167] J. P. Naiman, A. Pillepich, V. Springel, E. Ramirez-Ruiz, P. Torrey, M. Vogelsberger et al., *First results from the IllustrisTNG simulations: a tale of two elements - chemical evolution of magnesium and europium*, **477** (2018) 1206 [1707.03401].
- [168] M. Vogelsberger, S. Genel, V. Springel, P. Torrey, D. Sijacki, D. Xu et al., *Introducing the Illustris Project: simulating the coevolution of dark and visible matter in the Universe*, **444** (2014) 1518 [1405.2921].
- [169] M. Vogelsberger, S. Genel, V. Springel, P. Torrey, D. Sijacki, D. Xu et al., *Properties of galaxies reproduced by a hydrodynamic simulation*, **509** (2014) 177 [1405.1418].
- [170] S. Genel, M. Vogelsberger, V. Springel, D. Sijacki, D. Nelson, G. Snyder et al., *Introducing the Illustris project: the evolution of galaxy populations across cosmic time*, **445** (2014) 175 [1405.3749].
- [171] V. Springel, *E pur si muove: Galilean-invariant cosmological hydrodynamical simulations on a moving mesh*, **401** (2010) 791 [0901.4107].
- [172] R. Pakmor, A. Bauer and V. Springel, *Magnetohydrodynamics on an unstructured moving grid*, **418** (2011) 1392 [1108.1792].
- [173] M. Vogelsberger, S. Genel, D. Sijacki, P. Torrey, V. Springel and L. Hernquist, *A model for cosmological simulations of galaxy formation physics*, **436** (2013) 3031 [1305.2913].
- [174] P. Torrey, M. Vogelsberger, S. Genel, D. Sijacki, V. Springel and L. Hernquist, *A model for cosmological simulations of galaxy formation physics: multi-epoch validation*, **438** (2014) 1985 [1305.4931].
- [175] R. Weinberger, V. Springel, L. Hernquist, A. Pillepich, F. Marinacci, R. Pakmor et al., *Simulating galaxy formation with black hole driven thermal and kinetic feedback*, **465** (2017) 3291 [1607.03486].
- [176] A. Pillepich, V. Springel, D. Nelson, S. Genel, J. Naiman, R. Pakmor et al., *Simulating galaxy formation with the IllustrisTNG model*, **473** (2018) 4077 [1703.02970].
- [177] R. Weinberger, V. Springel, R. Pakmor, D. Nelson, S. Genel, A. Pillepich et al., *Supermassive black holes and their feedback effects in the IllustrisTNG simulation*, **479** (2018) 4056 [1710.04659].
- [178] S. Genel, D. Nelson, A. Pillepich, V. Springel, R. Pakmor, R. Weinberger et al., *The size evolution of star-forming and quenched galaxies in the IllustrisTNG simulation*, **474** (2018) 3976 [1707.05327].
- [179] M. Vogelsberger, F. Marinacci, P. Torrey, S. Genel, V. Springel, R. Weinberger et al., *The uniformity and time-invariance of the intra-cluster metal distribution in galaxy clusters from the IllustrisTNG simulations*, **474** (2018) 2073 [1707.05318].

- [180] P. Torrey, M. Vogelsberger, F. Marinacci, R. Pakmor, V. Springel, D. Nelson et al., *The evolution of the mass-metallicity relation and its scatter in IllustrisTNG*, **484** (2019) 5587 [1711.05261].
- [181] B. Diemer, A. R. H. Stevens, J. C. Forbes, F. Marinacci, L. Hernquist, C. d. P. Lagos et al., *Modeling the Atomic-to-molecular Transition in Cosmological Simulations of Galaxy Formation*, **238** (2018) 33 [1806.02341].
- [182] M. Davis, G. Efstathiou, C. S. Frenk and S. D. M. White, *The evolution of large-scale structure in a universe dominated by cold dark matter*, **292** (1985) 371.
- [183] U. Seljak, N. Hamaus and V. Desjacques, *How to Suppress the Shot Noise in Galaxy Surveys*, **103** (2009) 091303 [0904.2963].
- [184] F. Bernardeau and R. van de Weygaert, *A new method for accurate estimation of velocity field statistics*, **279** (1996) 693 [astro-ph/9508125].
- [185] U. Abbas and R. K. Sheth, *Strong clustering of underdense regions and the environmental dependence of clustering from Gaussian initial conditions*, **378** (2007) 641 [astro-ph/0703391].
- [186] C. Uhlemann et al., *in prep.*, .
- [187] Y. Sigad, E. Branchini and A. Dekel, *Measuring the Nonlinear Biasing Function from a Galaxy Redshift Survey*, **540** (2000) 62 [astro-ph/0002170].
- [188] I. Jee, C. Park, J. Kim, Y.-Y. Choi and S. S. Kim, *A Second-order Bias Model for the Logarithmic Halo Mass Density*, **753** (2012) 11 [1204.5573].
- [189] M. C. Neyrinck, I. Szapudi and A. S. Szalay, *Rejuvenating the Matter Power Spectrum: Restoring Information with a Logarithmic Density Mapping*, **698** (2009) L90 [0903.4693].
- [190] J. Carron and I. Szapudi, *Optimal non-linear transformations for large-scale structure statistics*, **434** (2013) 2961 [1306.1230].
- [191] H. Padmanabhan and G. Kulkarni, *Constraints on the evolution of the relationship between HI mass and halo mass in the last 12 Gyr*, **470** (2017) 340 [1608.00007].
- [192] H. Padmanabhan and A. Refregier, *Constraining a halo model for cosmological neutral hydrogen*, **464** (2017) 4008 [1607.01021].
- [193] L. Wolz, S. G. Murray, C. Blake and J. S. Wyithe, *Intensity mapping cross-correlations II: HI halo models including shot noise*, **484** (2019) 1007 [1803.02477].
- [194] C. Uhlemann, O. Friedrich, F. Villaescusa-Navarro, A. Banerjee and S. r. Codis, *Fisher for complements: Extracting cosmology and neutrino mass from the counts-in-cells PDF*, (2020) [1911.11158].

- [195] R. Emami, T. Broadhurst, P. Jimeno, G. Smoot, R. Angulo, J. Lim et al., *Evidence of Neutrino Enhanced Clustering in a Complete Sample of Sloan Survey Clusters, Implying $\sum m_\nu = 0.119 \pm 0.034$ eV*, *arXiv e-prints* (2017) arXiv:1711.05210 [1711.05210].
- [196] M. Suginoara, T. Suginoara and D. N. Spergel, *Detecting $z > 10$ Objects through Carbon, Nitrogen, and Oxygen Emission Lines*, **512** (1999) 547 [astro-ph/9803236].
- [197] J. Fonseca, M. B. Silva, M. G. Santos and A. Cooray, *Cosmology with intensity mapping techniques using atomic and molecular lines*, **464** (2017) 1948 [1607.05288].
- [198] H. T. Ihle, D. Chung, G. Stein, M. Alvarez, J. R. Bond, P. C. Breysse et al., *Joint Power Spectrum and Voxel Intensity Distribution Forecast on the CO Luminosity Function with COMAP*, **871** (2019) 75 [1808.07487].
- [199] D. Alonso, P. G. Ferreira and M. G. Santos, *Fast simulations for intensity mapping experiments*, **444** (2014) 3183 [1405.1751].
- [200] E. Sefusatti, M. Crocce, S. Pueblas and R. Scoccimarro, *Cosmology and the bispectrum*, **74** (2006) 023522 [astro-ph/0604505].
- [201] B. Joachimi, X. Shi and P. Schneider, *Bispectrum covariance in the flat-sky limit*, **508** (2009) 1193 [0907.2901].
- [202] G. N. Watson, *A Treatise on the Theory of Bessel Functions*. Cambridge University Press, 1966.
- [203] I. S. Gradshteyn, I. M. Ryzhik, A. Jeffrey and D. Zwillinger, *Table of Integrals, Series, and Products*. 2000.
- [204] T. Baldauf, M. Mirbabayi, M. Simonović and M. Zaldarriaga, *LSS constraints with controlled theoretical uncertainties*, *arXiv e-prints* (2016) arXiv:1602.00674 [1602.00674].
- [205] OEIS Foundation Inc., *The On-Line Encyclopedia of Integer Sequences®*, 2020.
- [206] M. Schmittfull, T. Baldauf and U. Seljak, *Near optimal bispectrum estimators for large-scale structure*, **91** (2015) 043530 [1411.6595].
- [207] A. Eggemeier and R. E. Smith, *Cosmology with phase statistics: parameter forecasts and detectability of BAO*, **466** (2017) 2496 [1611.01160].
- [208] J. Byun, A. Eggemeier, D. Regan, D. Seery and R. E. Smith, *Towards optimal cosmological parameter recovery from compressed bispectrum statistics*, **471** (2017) 1581 [1705.04392].
- [209] D. Gualdi, M. Manera, B. Joachimi and O. Lahav, *Maximal compression of the redshift-space galaxy power spectrum and bispectrum*, **476** (2018) 4045 [1709.03600].

- [210] R. Scoccimarro, *The Bispectrum: From Theory to Observations*, **544** (2000) 597 [astro-ph/0004086].
- [211] Z. Slepian and D. J. Eisenstein, *Modelling the large-scale redshift-space 3-point correlation function of galaxies*, **469** (2017) 2059 [1607.03109].
- [212] N. Agarwal, V. Desjacques, D. Jeong and F. Schmidt, *Information content in the redshift-space galaxy power spectrum and bispectrum*, *arXiv e-prints* (2020) arXiv:2007.04340 [2007.04340].
- [213] V. Yankelevich and C. Porciani, *Cosmological information in the redshift-space bispectrum*, **483** (2019) 2078 [1807.07076].
- [214] M. Schmittfull and U. Seljak, *Parameter constraints from cross-correlation of CMB lensing with galaxy clustering*, **97** (2018) 123540 [1710.09465].
- [215] N. Schöneberg, M. Simonović, J. Lesgourgues and M. Zaldarriaga, *Beyond the traditional line-of-sight approach of cosmological angular statistics*, **2018** (2018) 047 [1807.09540].
- [216] H. S. Grasshorn Gebhardt and D. Jeong, *Fast and accurate computation of projected two-point functions*, **97** (2018) 023504 [1709.02401].
- [217] M. Simonović, T. Baldauf, M. Zaldarriaga, J. J. Carrasco and J. A. Kollmeier, *Cosmological perturbation theory using the FFTLog: formalism and connection to QFT loop integrals*, **2018** (2018) 030 [1708.08130].
- [218] Z. Slepian, *On Decoupling the Integrals of Cosmological Perturbation Theory*, *arXiv e-prints* (2018) arXiv:1812.02728 [1812.02728].
- [219] X. Fang, E. Krause, T. Eifler and N. MacCrann, *Beyond Limber: efficient computation of angular power spectra for galaxy clustering and weak lensing*, **2020** (2020) 010 [1911.11947].
- [220] Z. Slepian, Y. Li, M. Schmittfull and Z. Vlah, *Rotation method for accelerating multiple-spherical Bessel function integrals against a numerical source function*, *arXiv e-prints* (2019) arXiv:1912.00065 [1912.00065].
- [221] J. E. Campagne, S. Plaszczynski and J. Neveu, *The Galaxy Count Correlation Function in Redshift Space Revisited*, **845** (2017) 28 [1703.02818].
- [222] V. Tansella, C. Bonvin, R. Durrer, B. Ghosh and E. Sellentin, *The full-sky relativistic correlation function and power spectrum of galaxy number counts. Part I: theoretical aspects*, **2018** (2018) 019 [1708.00492].
- [223] V. Tansella, G. Jelic-Cizmek, C. Bonvin and R. Durrer, *COFFE: a code for the full-sky relativistic galaxy correlation function*, **2018** (2018) 032 [1806.11090].
- [224] J. E. Campagne, J. Neveu and S. Plaszczynski, *Angpow: a software for the fast computation of accurate tomographic power spectra*, **602** (2017) A72 [1701.03592].

- [225] A. Moradinezhad Dizgah, H. Lee, M. Schmittfull and C. Dvorkin, *Capturing non-Gaussianity of the large-scale structure with weighted skew-spectra*, **2020** (2020) 011 [1911.05763].
- [226] V. Assassi, M. Simonović and M. Zaldarriaga, *Efficient evaluation of angular power spectra and bispectra*, **2017** (2017) 054 [1705.05022].
- [227] M. Tegmark, A. N. Taylor and A. F. Heavens, *Karhunen-Loève Eigenvalue Problems in Cosmology: How Should We Tackle Large Data Sets?*, **480** (1997) 22 [astro-ph/9603021].
- [228] R. E. Smith, D. S. Reed, D. Potter, L. Marian, M. Crocce and B. Moore, *Precision cosmology in muddy waters: cosmological constraints and N-body codes*, **440** (2014) 249 [1211.6434].
- [229] Y.-S. Song, A. Taruya and A. Oka, *Cosmology with anisotropic galaxy clustering from the combination of power spectrum and bispectrum*, **2015** (2015) 007 [1502.03099].
- [230] K. C. Chan and L. Blot, *Assessment of the information content of the power spectrum and bispectrum*, **96** (2017) 023528 [1610.06585].
- [231] C. Uhlemann, O. Friedrich, F. Villaescusa-Navarro, A. Banerjee and S. r. Codis, *Fisher for complements: extracting cosmology and neutrino mass from the counts-in-cells PDF*, **495** (2020) 4006 [1911.11158].
- [232] O. Friedrich, C. Uhlemann, F. Villaescusa-Navarro, T. Baldauf, M. Manera and T. Nishimichi, *Primordial non-Gaussianity without tails - how to measure f_{NL} with the bulk of the density PDF*, (2020) [1912.06621].
- [233] J. R. Fergusson, M. Liguori and E. P. S. Shellard, *General CMB and primordial bispectrum estimation: Mode expansion, map making, and measures of F_{NL}* , **82** (2010) 023502 [0912.5516].
- [234] J. Hung, J. R. Fergusson and E. P. S. Shellard, *Advancing the matter bispectrum estimation of large-scale structure: a comparison of dark matter codes*, *arXiv e-prints* (2019) arXiv:1902.01830 [1902.01830].