

Copyright
by
Donghui Jeong
2010

The Dissertation Committee for Donghui Jeong
certifies that this is the approved version of the following dissertation:

Cosmology with high ($z > 1$) redshift galaxy surveys

Committee:

Eiichiro Komatsu, Supervisor

Volker Bromm

Gary Hill

Uros Seljak

Paul Shapiro

Cosmology with high ($z > 1$) redshift galaxy surveys

by

Donghui Jeong, B.S., M.S.

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN

August 2010

To Hyun Seol

Acknowledgments

First and foremost, I would like to express my deepest gratitude and appreciation to my advisor, Professor Eiichiro Komatsu, for everything he provided to me: kind and careful guidance, encouragement, and an infinite amount of support for both research and everyday life. His deep and broad knowledge about cosmology always awakes me and his deep insight always inspires me. Above all, Eiichiro is my hero and my role model for a cosmologist.

I am also grateful to my committee members, Professor Paul Shapiro, Professor Volker Bromm, Doctor Gary Hill, and Professor Uros Seljak, for their interest in my work, and many useful comments and suggestions.

I am indebted to many of my HETDEX collaborators. In particular, Professor Karl Gebhardt and Doctor Gary Hill continuously show interest in my work and support the work in many ways. Josh Adams and Guillermo Blanc always satisfy my thirst for knowledge on observation. I feel very lucky to work for such a promising project and with such superb colleagues.

The TVD code I used to test validity of perturbation theory in Chapter 3 was developed by Professor Dongsu Ryu. I thank him for kindly providing the code and answering my questions about the code. I also thank Doctor Kyungjin Ahn for giving me step-by-step instructions on how to run the TVD code.

At an earlier stage of the Reloaded II project, discussed in Chapter 4, Doctor Ilian Iliev (now Professor in Sussex University) kindly provided his N-body simulation data from which I developed my analysis code.

Chapter 6 of this dissertation was accomplished in collaboration with Professor Bhuvnesh Jain. I thank Bhuv for teaching me about many issues relating to large scale galaxy-galaxy weak gravitational lensing.

I would like to thank Doctor Emiliano Sefusatti for teaching me about the new method of measuring bispectrum, which is written about in Chapter 7 of this dissertation.

I am grateful to Professor Jai-chan Hwang and Doctor Hyerim Noh, Doctor Jinn-Ouk Gong, and Doctor Ki-Young Choi. Although my papers with them were not explicitly

included as a chapter in this dissertation, working with them enlightened me about the new topics of non-linear evolution in the General Theory of Relativity and the isocurvature perturbation in Multi-field inflation models.

I also benefited from my UT cosmology group members and office mates. I would especially like to thank Kyungjin Ahn, Marcelo Alvarez, Jun Koda, Jun Zhang, Yi Mao, Yuki Watanabe, Joshua Adams, Masatoshi Shoji, Jonathan Ganc and Chi-Ting Chiang.

I am indebted to my fellow graduate students, Hyo-Jeong Kim, Sehyun Hwang, Myungwon Jeon and Mimi Song. Thanks to them, my life in Austin became happier and richer.

This dissertation would not be possible without the help of the fabulous staff at the Astronomy Department. I owe a lot to two of my graduate coordinators, Stephanie Croch, Charmarie Burke, and the omnipotent Gordon Oris.

Words are insufficient to express my appreciation to my father, Ki-Bong Jeong, mother, Nan-Sun Lee, and brother, Hyung-Hee Jeong, for their warm love and ceaseless encouragement. I am also grateful to my father-in-law, Jae-Young Park, and mother-in-law Yong-Gi Huh for their belief in me and for letting Hyun Seol live with me here in the U.S.

Last, but not least, I would like to express my loving thanks to my wife, Hyun Seol Park, who is always with me and is the meaning of my life.

Cosmology with high ($z > 1$) redshift galaxy surveys

Publication No. _____

Donghui Jeong, Ph.D.

The University of Texas at Austin, 2010

Supervisor: Eiichiro Komatsu

Galaxy redshift surveys are powerful probes of cosmology. Yet, in order to fully exploit the information contained in galaxy surveys, we need to improve upon our understanding of the structure formation in the Universe. Galaxies are formed/observed at late times when the density field is no longer linear so that understanding non-linearities is essential. In this thesis, we show that, at high redshifts, we can accurately model the galaxy power spectrum in redshift space by using the standard cosmological perturbation theory.

Going beyond the power spectrum, we can use the three-point function, or the bispectrum, to gain important information on the early universe as well as on the galaxy formation via measurements of primordial non-Gaussianity and galaxy bias. We show that the galaxy bispectrum is more sensitive to primordial non-Gaussianities than previously recognized, making high-redshift galaxy surveys a particularly potent probe of the physics of inflation.

Weak lensing offers yet another way of probing cosmology. By cross correlating the angular position of galaxies with the shear measurement from galaxy lensing or CMB lensing, we also show that one can obtain the information on cosmological distance scale, the galaxy bias, and the primordial non-Gaussianity from weak lensing method.

Table of Contents

Acknowledgments	v
Abstract	vii
List of Tables	xvi
List of Figures	xvii
Chapter 1. Introduction	1
Chapter 2. Eulerian cosmological Perturbation Theory	6
2.1 Eulerian Perturbation Theory solution	8
2.1.1 Linear solution for density field and velocity field	9
2.1.2 Non-linear solution for density field and velocity field	14
2.2 Statistics of the cosmological density field: Gaussian vs. non-Gaussian	17
2.3 Next-to-leading order power spectrum from the perturbative solution: the theoretical template	21
2.3.1 One-loop Power spectrum with Gaussian linear density field	22
2.3.2 One-loop Power spectrum with non-Gaussian linear density field	23
2.3.2.1 Local primordial non-Gaussianity	24
2.4 Nonlinear matter power spectrum in real space	25
2.4.1 Gaussian case	26
2.4.2 Non-Gaussianity case	26
2.5 Nonlinear galaxy power spectrum in real space	27
2.5.1 Gaussian case	28
2.5.2 Galaxy-matter cross power spectrum	31
2.5.3 non-Gaussian case	33
2.6 Nonlinear mater power spectrum in redshift space	34
2.6.1 Gaussian case	38
2.6.2 non-Gaussian case	40

2.7	Nonlinear galaxy power spectrum in redshift space	40
2.7.1	Gaussian case	42
2.7.2	non-Gaussian case	45
2.8	Summary	46
2.8.1	On the smoothing of density field	46
2.8.2	Summary of equations	46
Chapter 3. Perturbation Theory Reloaded: analytical calculation of the non-linear matter power spectrum in real and redshift space		49
3.1	N -body simulations and analysis method	49
3.2	Non-linear Matter Power spectrum in Real Space	50
3.2.1	Two theories: Halo Fit vs. Perturbation Theory	50
3.2.2	Results	53
3.2.3	Convergence test	58
3.3	Non-linear Matter Power spectrum in Redshift Space	61
3.3.1	Three theories: Kaiser, Scoccimarro, and 3PT	61
3.3.2	Result	63
3.3.3	Modeling Finger of God effect	66
3.4	Discussion and Conclusions	77
Chapter 4. Perturbation Theory Reloaded II: Non-linear Bias, Baryon Acoustic Oscillations and Millennium Simulation In Real Space		79
4.1	Non-linear galaxy power spectrum from perturbation theory	79
4.1.1	Locality Assumption	79
4.1.2	3rd-order PT galaxy power spectrum	80
4.1.3	Why we do not care about the precise values of bias parameters	81
4.2	Dark Matter Power spectrum from Millennium Simulation	82
4.2.1	Millennium Simulation	83
4.2.2	3rd-order PT versus Millennium Simulation: Dark Matter Power Spectrum	84
4.3	HALO/GALAXY POWER SPECTRUM AND THE NON-LINEAR BIAS MODEL	90
4.3.1	Analysis method	90
4.3.2	Halo power spectra	93

4.3.2.1	Measuring non-linear halo bias parameters	93
4.3.2.2	Degeneracy of bias parameters	100
4.3.2.3	Comparison with the halo model predictions	101
4.3.2.4	Comments on the bispectrum	102
4.3.2.5	Effects on BAOs	102
4.3.3	Galaxy power spectra	103
4.3.3.1	Measuring non-linear galaxy bias parameters	103
4.3.3.2	Comparison with the simplest HOD predictions	108
4.3.3.3	Effects on BAOs	109
4.3.4	Mass dependence of bias parameters and effects on BAOs	109
4.4	Cosmological parameter estimation with the non-linear bias model	112
4.4.1	Measuring Distance Scale	113
4.4.1.1	Background	113
4.4.1.2	Method: Measuring “Box Size” of the Millennium Simulation	115
4.4.1.3	Results: Unbiased Extraction of the distance scale from the Millennium Simulation	118
4.4.1.4	Optimal estimation of the distance scale	118
4.4.1.5	Forecast for a HETDEX-like survey	121
4.5	Discussion and Conclusions	123

Chapter 5. Primordial non-Gaussianity, scale-dependent bias, and the bispectrum of galaxies 128

5.1	Bispectrum of Dark Matter Halos	131
5.1.1	Matarrese-Lucchin-Bonometto (MLB) method	131
5.1.2	Alternative derivation	133
5.2	Effects of Local-type Primordial non-Gaussianity on The Halo Bispectrum	136
5.2.1	Known Terms	138
5.2.1.1	Formula	138
5.2.1.2	Shape Dependence: Non-linear Gravitational Evolution and Non-linear Galaxy Bias	139
5.2.1.3	Shape Dependence: f_{NL} Term	144
5.2.2	New Term	145
5.2.2.1	Formula	145

5.2.2.2	Shape Dependence	150
5.2.3	Scale Dependence	159
5.2.4	Redshift Dependence	162
5.3	Discussion and Conclusions	162
5.4	Summary of Equations	165
Chapter 6.	Galaxy-CMB and galaxy-galaxy lensing on large scales: sensitivity to primordial non-Gaussianity	169
6.1	Halo-mass correlation from galaxy-galaxy lensing	171
6.1.1	Formula	171
6.1.2	Results	175
6.1.3	Covariance matrix of the mean tangential shear	179
6.1.4	Detectability of the mean tangential shear	181
6.2	Harmonic Space Approach	186
6.2.1	Formula	186
6.2.2	Result	188
6.2.3	Covariance matrix of the galaxy-convergence cross power spectrum . .	191
6.3	Halo-mass correlation from galaxy-CMB lensing	195
6.3.1	Formula	195
6.3.2	Results	197
6.3.3	Covariance matrix of the galaxy-CMB lensing	199
6.4	Discussion and Conclusions	207
Chapter 7.	Measuring Power spectrum and Bispectrum	209
7.1	Power spectrum and Bispectrum from N-body simulation	209
7.1.1	From particle to grid	209
7.1.1.1	NGP	211
7.1.1.2	CIC	212
7.1.1.3	TSC	212
7.1.1.4	3D window function	213
7.1.2	Power spectrum and bispectrum: the estimators	213
7.1.2.1	Power spectrum estimator: direct sampling	214
7.1.2.2	Bispectrum estimator: direct sampling	215

7.1.2.3	Estimating power spectrum II	215
7.1.2.4	Estimating Bispectrum II	217
7.1.2.5	Counting the number of triangles	218
7.1.3	Deconvolution	219
7.1.3.1	Deconvolving only window function	219
7.1.3.2	Deconvolving window function and aliasing	219
7.1.4	Poisson shot noise	221
7.1.4.1	Poisson sampling and underlying density field	222
7.1.4.2	Power spectrum and Bispectrum of discrete particles	222
7.1.4.3	Bispectrum of discrete particles	224
7.2	Power spectrum from Galaxy surveys	225
7.2.1	The FKP estimator	226
7.2.1.1	The power spectrum with weighting function $w(\mathbf{r})$	226
7.2.2	The variance of the power spectrum	228
7.2.3	Optimal weighting	229
7.3	Implementing the FKP estimator	230
7.3.1	The estimator	234
7.3.1.1	Constant weighting	234
7.3.1.2	FKP optimal weighting	235

Appendices 237

Appendix A. Fourier transforms 238

A.1	Continuous Fourier transform	238
A.1.1	Delta function	238
A.1.2	Convolution theorem	239
A.2	From Fourier Transform to Fourier Series	240
A.2.1	Sampling and aliasing	242
A.3	Fourier Series to Discrete Fourier Transform	244
A.4	Discrete Fourier Transform with FFTW	246
A.4.1	Output array of <code>dft_r2c_1d</code>	248
A.4.2	Output array of <code>dft_r2c_2d</code>	249
A.4.3	Output array of <code>dft_r2c_3d</code>	252

Appendix B. Derivation of one-loop power spectrum	254
B.1 Gaussian linear density field	254
B.2 Non-Gaussian linear density field	256
Appendix C. Derivation of the primordial bispectrum and trispectrum from local non-Gaussianity	258
Appendix D. $P_{gs,22}^{(rest)}$ and $P_{gs,13}^{(rest)}$	262
Appendix E. Lagrangian Perturbation Theory and Initial condition for Cosmological N-body Simulation	266
E.1 Lagrangian perturbation theory formalism	266
E.2 Linear Lagrangian perturbation theory	268
E.2.1 Zel'dovich approximation	269
E.3 Second order Lagrangian perturbation theory (2LPT)	270
E.4 Generating initial condition using Linear solution	271
E.5 Starting redshifts, initial condition generators, and convergence tests	274
Appendix F. Error on power spectrum and bispectrum	279
F.1 Error on power spectrum	279
F.2 Error on bispectrum	280
Appendix G. Analytical marginalization of the likelihood function over \tilde{b}_1^2 and P_0	285
Appendix H. distribution of errors on the distance scale	287
Appendix I. Theory on statistics of density peaks: Mass function and halo bias	289
I.1 Mass function of halos	289
I.1.1 Variables	289
I.1.2 The random walk view	290
I.1.3 Press-Schechter mass function	291
I.1.3.1 ST convention	292
I.1.4 Mass function from excursion set approach	292
I.1.5 Halo merger rate	294

I.1.5.1	Conditional mass-function	294
I.1.5.2	Halo merger rate	294
I.1.6	Sheth-Tormen mass function	295
I.1.6.1	ST convention	296
I.2	Halo bias with Gaussian linear density field	296
I.2.1	Kaiser (1984)	296
I.2.1.1	Gaussian density field	296
I.2.1.2	General pdf : peak/background split method	299
I.2.2	Linear bias with Press-Schechter mass function	299
I.2.3	Linear bias parameters from excursion set approach	300
I.2.4	Linear bias with Shech-Tormen mass function	302
I.2.5	Nonlinear halo bias and its mass dependence	303
I.3	Halo bias with local type primordial non-Gaussianity	304
I.3.1	Basic idea	304
I.3.2	Calculation	305
Appendix J. Integration of T_R^{1112}		309
Appendix K. The three-point correlation function (bispectrum) of density peaks		312
K.1	Probability of finding N distinct peaks	313
K.1.1	Calculation of P_1	315
K.1.2	Calculation of P_2	317
K.1.3	Calculation of P_3	319
K.2	The two point correlation function of peaks	321
K.3	The three point correlation function of peaks	323
Appendix L. Mean Tangential Shear and its covariance		328
L.1	Derivation of the mean tangential shear	328
L.2	Derivation of the covariance matrix of the mean tangential shear	329
Appendix M. On the accuracy of Limber's approximation		332
Appendix N. Cosmic Microwave Background lensing reconstruction: Quadratic Estimator with the flat-sky approximation		346

Bibliography	353
Vita	363

List of Tables

3.1	The best-fit $[\sigma_p^2]$ parameter for 6 different redshifts and 3 different k_{max} 's for each redshift. Note that in most of cases, k_{max} 's are set by the resolution of each simulation boxes : $k_{max} = 0.24, 0.5, 1.4 h \text{ Mpc}^{-1}$ for 512, 256, 128 $h^{-1} \text{ Mpc}$, respectively. However, when k_{max} set by the resolution is higher than the k_{max} we get from the valid k region of the nonlinear PT in real space, we choose the later. The higher values of χ_{red}^2 in the case of $k < 1.4$ at $z = 2$, and $k < 0.5, k < 1.4$ at $z = 1$ show the failure of the fitting in the highly nonlinear region. Note that $[\sigma_v^2]$ from equation (3.40) is always larger than the best-fit value when perturbation theory ansatz give rise to the small χ_{red}^2 .	70
4.1	Maximum wavenumbers, k_{max} , for the Millennium Simulation	84
4.2	Summary of six snapshots from the Millennium Simulation	90
4.3	Non-linear halo bias parameters and the corresponding 68% interval estimated from the MPA halo power spectra	99
4.4	Non-linear halo bias parameters and the corresponding 68% interval estimated from the MPA galaxy power spectra	103
4.5	Non-linear halo bias parameters and the corresponding 68% interval estimated from the Durham galaxy power spectra	108
4.6	Mass dependence of non-linear halo bias parameters (MPA halos)	126
4.7	Mass dependence of non-linear galaxy bias parameters (MPA galaxies) . . .	127

List of Figures

2.1	The linear growth factor, $D(a)$, for three different cosmologies: sCDM ($\Omega_m = 1, \Omega_\Lambda = 0$) Λ CDM ($\Omega_m = 0.277, \Omega_\Lambda = 0.723$) oCDM ($\Omega_m = 0.277, \Omega_\Lambda = 0$) . . .	10
2.2	The logarithmic derivative of the linear growth factor, $f(a) \equiv d \ln D / d \ln a$, for three different cosmologies: sCDM ($\Omega_m = 1, \Omega_\Lambda = 0$) Λ CDM ($\Omega_m = 0.277, \Omega_\Lambda = 0.723$) oCDM ($\Omega_m = 0.277, \Omega_\Lambda = 0$)	11
3.1	Power spectrum at $z = 1, 2, 3, 4, 5$ and 6 (from top to bottom), derived from N -body simulations (dashed lines), perturbation theory (solid lines), and linear theory (dot-dashed lines). We plot the simulation data from 512, 256, 128, and $64 h^{-1}$ Mpc simulations at $k \leq 0.24 h \text{ Mpc}^{-1}$, $0.24 < k \leq 0.5 h \text{ Mpc}^{-1}$, $0.5 < k \leq 1.4 h \text{ Mpc}^{-1}$, and $1.4 < k \leq 5 h \text{ Mpc}^{-1}$, respectively. Note that we did not run $64 h^{-1}$ Mpc simulations at $z = 1$ or 2	54
3.2	Dimensionless power spectrum, $\Delta^2(k)$. The solid and dashed lines show perturbation theory calculations and N -body simulations, respectively. The dotted lines show the predictions from halo approach (Smith et al., 2003). The dot-dashed lines show the linear power spectrum.	55
3.3	Residuals of Figure 3.2. The errorbars show the N -body data divided by the perturbation theory predictions minus one, while the solid curves show the halo model calculations given in Smith et al. (2003) divided by the perturbation theory predictions minus one. The perturbation theory predictions agree with simulations to better than 1% accuracy for $\Delta^2(k) \lesssim 0.4$	56
3.4	Non-linearity in baryonic acoustic oscillations. All of the power spectra have been divided by a smooth power spectrum without baryonic oscillations from equation (29) of Eisenstein & Hu (1998). The errorbars show N -body simulations, while the solid lines show perturbation theory calculations. The dot-dashed lines show the linear theory predictions. Perturbation theory describes non-linear distortion on baryonic oscillations very accurately at $z > 1$. Note that different redshift bins are not independent, as they have grown from the same initial conditions. The N -body data at $k < 0.24$ and $k > 0.24 h \text{ Mpc}^{-1}$ are from 512 and $256 h^{-1}$ Mpc box simulations, respectively.	57
3.5	Non-linearity and the amplitude of matter fluctuations, σ_8 . In each panel the lines show the linear spectrum and non-linear spectrum with $\sigma_8 = 0.7, 0.8, 0.9$ and 1.0 from bottom to top.	58
3.6	Convergence test. Fractional differences between the power spectra from N -body simulations in $L_{\text{box}} = 512, 256$, and $128 h^{-1}$ Mpc box (from bottom to top lines) and the perturbation theory predictions in $k < 1.5 h \text{ Mpc}^{-1}$	59

- 3.7 Convergence test. The same as left panel, but for simulations in $L_{\text{box}} = 512, 256, 128,$ and $64 h^{-1} \text{ Mpc}$ box (from bottom to top lines) in the expanded range in wavenumber, $k < 5 h \text{ Mpc}^{-1}$ 60
- 3.8 Comparing the angular averaged redshift space power spectrum for three different models at redshift $z = 1, 2, 3, 4, 5, 6$ (top to bottom). dashed line (*blue*) is the redshift space power spectrum calculated from the N-body simulations. Linear Kaiser $P(k)$ (dot dashed, *green* line), Scoccimarro's non-linear $P(k)$ (dotted, *olive* line), and full 3rd order non-linear redshift space $P(k)$ (solid, *red* line) are compared. Note that Finger of God effect is not yet included in the model and one can clearly see the suppression of the power on smaller scales. We plot the N-body power spectrum using the same data we used to draw Figure 3.1, so the k -range and the box size are all the same. 64
- 3.9 Dimensionless power spectrum, $\Delta_{red}^2(k)$, of redshift space for six different redshifts. The solid and dashed lines show full non-linear PT calculations and N-body simulations, respectively. The dotted line shows the Scoccimarro's formula and the dot-dashed line represents the linear kaiser spectrum. PT prediction agree with simulations very well at redshift $z = 6$ up to $k \simeq 0.8 h \text{ Mpc}^{-1}$, and at redshift $z = 5$ up to $k \simeq 0.5 h \text{ Mpc}^{-1}$. On low redshifts, however, N-body result shows power suppression relative to the PT power spectrum. It may be the Finger of God suppression. Note that the Scoccimarro's formula and linear kaiser spectrum are lower than the N-body result in the intermediate scale. 65
- 3.10 Nonlinearity in baryonic acoustic oscillations due to the redshift space distortion. Note that we don't include the Finger of God effect in the modeling power spectrum yet. All the power spectra have been divided by a smooth power spectrum without baryonic oscillations from equation (29) of Eisenstein & Hu (1998). The errorbars show N-body simulations, while solid line show PT calculations. The dot-dashed line show the linear theory predictions, and dotted line show the power spectrum from Scoccimarro's formula. PT describes nonlinear distortion on BAO in the redshift space very accurately at $z > 3$, up to third peak, and at $z = 2$ up to second peak. Although peak position is preserved, at $z = 1$, we loose almost all the information on the amplitude. It clearly shows that we need to model the power suppression. The N-body data at $k < 0.24$ and $> 0.24 h \text{ Mpc}^{-1}$ are from 512 and 256 $h \text{ Mpc}^{-1}$ box simulations, respectively. 67
- 3.11 Result of 2-dimensional fitting of σ_p^2 using 512 $h^{-1} \text{ Mpc}$ box. N-body power spectrum (blue line) in redshift space, PT power spectrum + Finger of God ansatz (red line), and PT power spectrum (red dot-dashed line) is shown. Note that we fit only for $|k| < 0.24 h \text{ Mpc}^{-1}$, except for $z = 1$, where we use $k_{max} = 0.15 h \text{ Mpc}^{-1}$ which is the maximum wavenumber below which PT works well [Section 3.2.2] 72
- 3.12 Same as Figure 3.11, but for 256 $h^{-1} \text{ Mpc}$ box. Note that we fit only for $|k| < 0.5 h \text{ Mpc}^{-1}$, except for $z = 2$ ($k_{max} = 0.33 h \text{ Mpc}^{-1}$). 73

3.13	Same as Figure 3.11, but for $128 h^{-1}\text{Mpc}$ box. Note that we fit only for $ k < 1.4 h \text{ Mpc}^{-1}$, except for $z = 3$ ($k_{max} = 1.2 h \text{ Mpc}^{-1}$).	74
3.14	Same as Figure 3.8, but with the PT+Finger of God fitting result (red, solid line). Finger of God modeling improves the agreement between PT and N-body power spectrum dramatically.	75
3.15	FoG effect on the angular averaged baryonic oscillations. This figure is the same as Figure 3.10, except for the PT+FoG plot (solid line). Note that the overall suppression of baryonic oscillation at $z = 1$ is modeled quite correctly up to the second peak.	76
4.1	Matter power spectrum at $z = 0, 1, 2, 3, 4, 5$ and 6 (<i>from top to bottom</i>) derived from the Millennium Simulation (<i>dashed lines</i>), the 3rd-order PT (<i>solid lines</i>), and the linear PT (<i>dot-dashed lines</i>).	85
4.2	Dimensionless matter power spectrum, $\Delta^2(k)$, at $z = 1, 2, 3, 4, 5$, and 6 . The dashed and solid lines show the Millennium Simulation data and the 3rd-order PT calculation, respectively. The dot-dashed lines show the linear power spectrum.	86
4.3	Fractional difference between the matter power spectra from the 3rd-order PT and that from the Millennium Simulation, $P_m^{sim}(k)/P_m^{PT} - 1$ (dots with errorbars). The solid lines show the perfect match, while the dashed lines show $\pm 2\%$ accuracy. We also show $k_{max}(z)$, below which we trust the prediction from the 3rd-order PT, as a vertical dotted line.	87
4.4	Distortion of BAOs due to non-linear matter clustering. All of the power spectra have been divided by a smooth power spectrum without baryonic oscillations from eq. (29) of Eisenstein & Hu (1998). The error bars show the simulation data, while the solid lines show the PT calculations. The dot-dashed lines show the linear theory calculations. The power spectrum data shown here have been taken from Figure 6 of Springel et al. (2005).	88
4.5	Halo power spectra from the Millennium Simulation at $z = 1, 2, 3, 4, 5$, and 6 . Also shown in smaller panels are the residual of fits. The points with errorbars show the measured halo power spectra, while the solid, dashed, and dot-dashed lines show the best-fitting non-linear bias model (Eq. (4.2)), the best-fitting linear bias with the non-linear matter power spectrum, and the best-fitting linear bias with the linear matter power spectrum, respectively. Both linear models have been fit for $k_{max,linear} = 0.15 [h \text{ Mpc}^{-1}]$, whereas $k_{max}(z)$ given in Table 4.1 (also marked in each panel) have been used for the non-linear bias model.	94
4.6	One-dimensional marginalized distribution of non-linear bias parameters at $z = 6$: from top to bottom panels, P_0 , b_2 , and b_1 . Different lines show the different values of k_{max} used for the fits. The dashed and solid lines correspond to $0.3 \leq k_{max}/[h \text{ Mpc}^{-1}] \leq 1.0$ and $1.0 < k_{max}/[h \text{ Mpc}^{-1}] \leq 1.5$, respectively. The double-peak structure disappears for higher k_{max}	95

4.7	Same as Figure 4.6, but for a Monte Carlo simulation of a galaxy survey with a bigger box size, $L_{box} = 1.5 \text{ Gpc}/h$	96
4.8	One-dimensional marginalized constraints and two-dimensional joint marginalized constraint of $2\text{-}\sigma$ (95.45% CL) range for bias parameters (b_1, b_2, P_0) . Covariance matrices are calculated from the Fisher information matrix (Eq. (4.9)) with the best-fitting bias parameters for halo at $z = 4$	97
4.9	Distortion of BAOs due to non-linear matter clustering and non-linear halo bias. All of the power spectra have been divided by a smooth power spectrum without baryonic oscillations from equation (29) of Eisenstein & Hu (1998). The errorbars show the Millennium Simulation, while the solid lines show the PT calculations. The dashed lines show the linear bias model with the non-linear matter power spectrum, and the dot-dashed lines show the linear bias model with the linear matter power spectrum. Therefore, the difference between the solid lines and the dashed lines shows the distortion solely due to non-linear halo bias.	98
4.10	Same as Figure 4.5, but for the MPA galaxy catalogue (Mgalaxy).	104
4.11	Same as Figure 4.5, but for the Durham galaxy catalogue (Dgalaxy).	105
4.12	Same as Figure 4.9, but for the MPA galaxy power spectrum (Mgalaxy).	106
4.13	Same as Figure 4.9, but for the Durham galaxy power spectrum (Dgalaxy).	107
4.14	Mass dependence of distortion of BAOs due to non-linear bias. Four mass bins, $M < 5 \times 10^{10} M_\odot/h$, $5 \times 10^{10} M_\odot/h < M < 10^{11} M_\odot/h$, $10^{11} M_\odot/h < M < 5 \times 10^{11} M_\odot/h$, and $5 \times 10^{11} M_\odot/h < M < 10^{12} M_\odot/h$, are shown. (M_{10} stands for $M/(10^{10} M_\odot)$.) All of the power spectra have been divided by a smooth power spectrum without baryonic oscillations from equation (29) of Eisenstein & Hu (1998). The errorbars show the Millennium Simulation data, while the solid lines show the PT calculation.	110
4.15	Same as Figure 4.14, but for the MPA galaxy catalogue (Mgalaxy).	111
4.16	Distance scale extracted from the Millennium Simulation using the 3rd-order PT galaxy power spectrum given by Eq. (4.2), divided by the true value. The mean of the likelihood (<i>stars</i>), and the maximum likelihood values (<i>filled circles</i>) and the corresponding $1\text{-}\sigma$ intervals (<i>errorbars</i>), are shown as a function of maximum wavenumbers used in the fits, k_{max} . We find $D/D_{true} = 1$ to within the $1\text{-}\sigma$ errors from all the halo/galaxy catalogues (“halo,” “Mgalaxy,” and “Dgalaxy”) at all redshifts, provided that we use k_{max} estimated from the matter power spectra, $k_{max} = 0.15, 0.25, 1.0, 1.2, 1.3$, and 1.5 at $z = 1, 2, 3, 4, 5$, and 6 , respectively (see Table 4.1). Note that the errors on D do not decrease as k_{max} increases due to degeneracy between D and the bias parameters. See Figure (4.18) and (4.19) for further analysis.	116
4.17	Same as Figure 4.8, but including the distance scale D/D_{true}	117

4.18	Same as Figure 4.16, but with b_1 and b_2 fixed at the best-fitting values. The 1- σ ranges for D are 1.5% and 0.15% for $k_{max} = 0.2 \text{ h/Mpc}$ and $k_{max} = 1.5 \text{ h/Mpc}$, respectively. The errors on D decrease as k_{max} increases, but the scaling is still milder than $1/\sqrt{\sum_{k < k_{max}} N_k}$	119
4.19	Same as Figure 4.16, but with b_1 , b_2 and P_0 fixed at the best-fitting values. The 1- σ ranges for D are 0.8% and 0.05% for $k_{max} = 0.2 \text{ h/Mpc}$ and $k_{max} = 1.5 \text{ h/Mpc}$, respectively. The errors on D decrease as k_{max} increases as $1/\sqrt{\sum_{k < k_{max}} N_k}$	120
4.20	Projected constraints on D at $z = 3$ from a HETDEX-like survey with the survey volume of $(1.5 \text{ Gpc/h})^3$. We have used the best-fitting 3rd-order PT power spectrum of MPA halos in the Millennium Simulation for generating a mock simulation data. We show the results for the number of objects of $N_{galaxy} = 2 \times 10^5$, 10^6 , 2×10^6 , and 10^9 , from the top to bottom panels, respectively, for which we find the projected 1- σ errors of 2.5%, 1.5%, 1%, and 0.3%, respectively.	122
5.1	Visual representations of triangles forming the bispectrum, $B(k_1, k_2, k_3)$, with various combinations of wavenumbers satisfying $k_3 \leq k_2 \leq k_1$	141
5.2	Shape of the bispectrum, $B(k_1, k_2, k_3)$. Each panel shows the amplitude of the bispectrum as a function of k_2/k_1 and k_3/k_1 for a given k_1 , with a condition that $k_3 \leq k_2 \leq k_1$ is satisfied. The amplitude is normalized such that it is unity at the point where the bispectrum takes on the maximum value. For the visual representations of the triangle names such as the squeezed, elongated, folded, isosceles, and equilateral, see Fig. 5.1. (Top Left) The bispectrum from the non-linear gravitational evolution, B_m^G (Eq. (5.23)), for $k_1 = 0.01 \text{ h Mpc}^{-1}$. (Top Right) B_m^G for $k_1 = 0.05 \text{ h Mpc}^{-1}$. (Bottom Left) The bispectrum from the non-linear galaxy biasing, $P_R(k_1)P_R(k_2) + (2 \text{ cyclic})$ (the second term in Eq. (5.16)), for $k_1 = 0.01 \text{ h Mpc}^{-1}$. (Bottom Right) $P_R(k_1)P_R(k_2) + (2 \text{ cyclic})$ for $k_1 = 0.05 \text{ h Mpc}^{-1}$	142
5.3	Same as the top panels of Fig. 5.2, but for $B_m^G/[P_R(k_1)P_R(k_2) + (2 \text{ cyclic})]$ (Eq. (5.25)).	142
5.4	Same as Fig. 5.2, but for the terms proportional to f_{NL} . (Top) the B_m^{nG} term (Eq. (5.35)), (Middle) the $B_{f_{NL}}^{nG0}$ term (Eq. (5.24)), and (Bottom) the $B_{f_{NL}}^{nG1}$ term (Eq. (5.41)). Note that the non-Gaussian terms diverge in the exact squeezed limit, $k_3 \rightarrow 0$; thus, we show these terms normalized to be unity at $k_3/k_1 = 10^{-2}$. In order to facilitate the comparison better, we draw the dotted contour for all six panels.	143
5.5	Same as Fig. 5.2, but for (Top) the g_{NL} term (Eq. (5.32)), and (Bottom) the f_{NL}^2 term (Eq. (5.33)). Note that the non-Gaussian terms diverge in the exact squeezed limit, $k_3 \rightarrow 0$; thus, we show these terms normalized to be unity at $k_3/k_1 = 10^{-2}$. In order to facilitate the comparison better, we draw the dotted contour for top panels.	144

5.6	Shape of the function, $\mathcal{F}_R(k)$, defined in Eq. (5.37). We show $\mathcal{F}_R(k)$ for four different smoothing lengths: $R = 1, 2, 5, 10$ Mpc/ h	150
5.7	Shape of the function, $\mathcal{G}_R(k)$, defined in Eq. (5.38). We show $\mathcal{G}_R(k)$ for four different smoothing lengths: $R = 1, 2, 5, 10$ Mpc/ h	151
5.8	Large-scale asymptotic value of $\mathcal{G}_R(k)$ as a function of the smoothing scale R . The value for $R = 1$ [Mpc/ h], which is used for generating Figs. 5.10 to 5.14, is 0.3718.	151
5.9	Shape of the integration that appears in the dominant term of $B_{f_{\text{NL}}^2}^{nG}$ and $B_{g_{\text{NL}}}^{nG}$, Eq. (5.50). We use four different smoothing scales: $R = 1, 2, 5$, and 10 Mpc/ h	155
5.10	Scale and shape dependence of the galaxy bispectrum terms that are linearly proportional to f_{NL} , as a function of k_1 . Except for the bottom-right panel, we use $R = 1$ Mpc/ h . (Top Left) The squeezed triangles with $k_1 = k_2 = 100k_3$, (Top Right) the elongated triangles with $k_1 = k_2 + k_3$ and $k_2 = 3k_3$, (Middle Left) the folded triangles with $k_1 = 2k_2 = 2k_3$, (Middle Right) the isosceles triangles with $3k_1 = 4k_2 = 4k_3$, and (Bottom Left) the equilateral triangles with $k_1 = k_2 = k_3$. The thick dot-dashed, dashed, solid, and dotted lines show the contributions from the primordial non-Gaussianity: the $B_{f_{\text{NL}}}^{nG0}$ (Eq. (5.24)), $\tilde{b}_2/b_1 B_m^{nG}$ (Eq. (5.35)), $4(\tilde{b}_2/b_1)[\mathcal{G}_R(k_1) + \mathcal{G}_R(k_2) + \mathcal{G}_R(k_3)] B_{f_{\text{NL}}}^{nG0}$ ($\mathcal{G}_R(k)$ defined in Eq. (5.38)), and $\tilde{b}_2/b_1 B_{f_{\text{NL}}}^{nG1}$ (Eq. (5.41)) terms, respectively. The thin dotted and dashed lines show the non-linear effects: B_m^G (Eq. (5.23)) and the non-linear bias (the second term in Eq. (5.11)), respectively. We use the standard value of $\tilde{b}_2/b_1 \equiv \delta_c \simeq 1.686$ from spherical collapse model. (Bottom Right) Dependence of the squeezed bispectrum on the smoothing scale, R , showing that the dependence is negligible for $k_1 \ll 1/R$	156
5.11	Scale and shape dependence of various bispectrum terms, $B(k_1, k_2, k_3)$, as a function of k_1 . For the figure except for the bottom right, we use $R = 1$ Mpc/ h . (Top Left) The squeezed triangles with $k_1 = k_2 = 100k_3$, (Top Right) the elongated triangles with $k_1 = k_2 + k_3$ and $k_2 = 3k_3$, (Middle Left) the folded triangles with $k_1 = 2k_2 = 2k_3$, (Middle Right) the isosceles triangles with $3k_1 = 4k_2 = 4k_3$, and (Bottom Left) the equilateral triangles with $k_1 = k_2 = k_3$. The thick dot-dashed, triple-dot-dashed, and solid lines show the contributions from the primordial non-Gaussianity: the $f_{\text{NL}} B_{f_{\text{NL}}}^{tot}$ (Eq. (5.53)), $\tilde{b}_2/b_1 g_{\text{NL}} B_{g_{\text{NL}}}^{nG}$ (Eq. (5.32)), and $\tilde{b}_2/b_1 f_{\text{NL}}^2 B_{f_{\text{NL}}^2}^{nG}$ (Eq. (5.33)) terms, respectively. The thin dotted and dashed lines show the non-linear effects: B_m^G (Eq. (5.23)) and the non-linear bias (the second term in Eq. (5.11)), respectively. We use the standard value of $\tilde{b}_2/b_1 \equiv \delta_c \simeq 1.686$ from spherical collapse model. (Bottom Right) Dependence of the squeezed bispectrum on the smoothing scale, R , showing that the dependence is negligible for $k_1 \ll 1/R$	157

5.12	Same as Fig. 5.11, but for squeezed triangles with different ratios: $\alpha = 50$ and $\alpha = 10$. (Top) All the parameters are the same as in Fig. 5.11. (Middle) $z = 3$ and $b_2/b_1 = 1.5$. The non-Gaussianity parameters, $f_{\text{NL}} = 40$ and $g_{\text{NL}} = 10^4$, are the same as in Fig. 5.11. (Bottom) $z = 3$ and $b_2/b_1 = 1.5$. The non-Gaussianity parameters, $f_{\text{NL}} = 4$ and $g_{\text{NL}} = 100$	158
5.13	Same as Fig. 5.11, but for $z = 3$ and $b_2/b_1 = 1.5$. The non-Gaussianity parameters, $f_{\text{NL}} = 40$ and $g_{\text{NL}} = 10^4$, are the same as in Fig. 5.11.	163
5.14	Same as Fig. 5.13, but for smaller non-Gaussianity parameters, $f_{\text{NL}} = 4$ and $g_{\text{NL}} = 100$	166
6.1	Coordinate system and γ_1 and γ_2 . The shear along \mathbf{e}_1 has $\gamma_1 > 0$ and $\gamma_2 = 0$, whereas the shear along \mathbf{e}_2 has $\gamma_1 < 0$ and $\gamma_2 = 0$. The shear along $\mathbf{e}_1 + \mathbf{e}_2$ has $\gamma_1 = 0$ and $\gamma_2 > 0$, whereas The shear along $\mathbf{e}_1 - \mathbf{e}_2$ has $\gamma_1 = 0$ and $\gamma_2 < 0$	0.170
6.2	Critical surface density, $\Sigma_c(z_L; z_S)$, as a function of the source redshift, z_S , for various lens redshifts that roughly correspond to the Two Degree Field Galaxy Redshift Survey (2dFGRS; $z_L = 0.1$, solid), the main sample of the Sloan Digital Sky Survey (SDSS; $z_L = 0.2$, dotted), the Luminous Red Galaxies (LRGs) of SDSS ($z_L = 0.3$, dashed), and the Large Synoptic Survey Telescope (LSST; $z_L = 0.5$ and 0.8 , dot-dashed and triple-dot-dashed, respectively).	174
6.3	The baryonic feature in the matter power spectrum, as seen in the galaxy-galaxy lensing, $\Delta\Sigma(R)$, for several populations of lens galaxies with $b_1 = 2$ at $z_L = 0.3$ (similar to SDSS LRGs, solid), $b_1 = 2$ at $z_L = 0.5$ (higher- z LRGs, dotted), $b_1 = 2$ at $z_L = 0.8$ (galaxies that can be observed by LSST, dashed), and $b_1 = 5$ at $z_L = 0.8$ (clusters of galaxies that can be observed by LSST, dot-dashed). The vertical line shows the location of the baryonic feature, $R_{\text{BAO}} = 106.9 h^{-1}$ Mpc, calculated from the “WMAP+BAO+SN ML” parameters in Table 1 of Komatsu et al. (2009). Note that we have used the linear matter power spectrum and the Gaussian initial condition ($f_{\text{NL}} = 0$) for this calculation.	176
6.4	Imprints of the local-type primordial non-Gaussianity in the galaxy-galaxy lensing, $\Delta\Sigma(R)$, for the same populations of lens galaxies as in Fig. 6.3. The solid, dashed, and dotted lines show $f_{\text{NL}} = 0, \pm 50$, and ± 100 , respectively.	177
6.5	Fractional differences between $\Delta\Sigma(R)$ from non-Gaussian initial conditions and the Gaussian initial condition, $ \Delta\Sigma(R; f_{\text{NL}})/\Delta\Sigma(R; f_{\text{NL}} = 0) - 1 $, calculated from the curves shown in Fig. 6.4. The dot-dashed, dashed, and dotted lines show $f_{\text{NL}} = \pm 10, \pm 50$, and ± 100 , respectively, while the thin solid line shows $\propto R^2$ with an arbitrary normalization.	178

- 6.6 Same as Fig. 6.3, but with the expected $1\text{-}\sigma$ uncertainties for full-sky lens surveys and a single lens redshift. Adjacent bins are highly correlated, with the correlation coefficients shown in Fig. 6.7. The open (filled) boxes show the binned uncertainties with (without) the cosmic variance term due to the cosmic shear field included. See Eq. (6.26) and (6.27) for the formulae giving open and filled boxes, respectively. We use the radial bin of size $\Delta R = 5 h^{-1}$ Mpc. For comparison, we also show $\Delta\Sigma(R)$ computed from the smooth power spectrum without the baryonic feature (Eisenstein & Hu, 1998) (dashed lines). Note that the uncertainties are calculated for a single lens redshift slice, and thus they will go down as we add more lens redshift slices. 182
- 6.7 The cross-correlation-coefficient matrix, $r_{ij} \equiv C_{ij}/\sqrt{C_{ii}C_{jj}}$, where C_{ij} is the covariance matrix given in Eq. (6.25), for a radial bin of $\Delta R = 5 h^{-1}$ Mpc. We show r_{ij} for the same populations of lens galaxies as shown in Fig. 6.3 and 6.6. We use the same number of source galaxies and the same shape noise as in Fig. 6.6. The neighboring bins are highly correlated for $\Delta R < 10 h^{-1}$ Mpc. 183
- 6.8 Same as Fig. 6.4, but with the expected $1\text{-}\sigma$ uncertainties for full-sky lens surveys and a single lens redshift. Adjacent bins are highly correlated. The open (filled) boxes show the binned uncertainties with (without) the cosmic variance term due to the cosmic shear field included. See Eq. (6.26) and (6.27) for the formulae giving open and filled boxes, respectively. We use logarithmic bins with $\Delta R = R/10$. Note that the uncertainties are calculated for a single lens redshift slice, and thus they will go down as we add more lens redshift slices. 184
- 6.9 Angular power spectrum of the galaxy-convergence cross correlation, $C_l^{h\kappa}$, at various multipoles as a function of the lens redshift, z_L , for two effective source redshifts, $z_s = 1$ (top) and 2 (bottom). We have divided $C_l^{h\kappa}$ by its maximum value. The solid, dotted, dashed, dot-dashed, and triple-dot-dashed lines show $l = 10, 50, 100, 350$, and 1000, respectively. 187
- 6.10 Imprints of the local-type primordial non-Gaussianity in the galaxy-convergence cross power spectrum, $l(l+1)C_l^{h\kappa}/(2\pi)$, for the same populations of lens galaxies as in Fig. 6.3. The solid, dashed, and dotted lines show $f_{\text{NL}} = 0, \pm 50$, and ± 100 , respectively. 189
- 6.11 Fractional differences between $C_l^{h\kappa}$ from non-Gaussian initial conditions and the Gaussian initial condition, calculated from the curves shown in Fig. 6.10. These differences are equal to $|\Delta b(l = k/d_A, z_L)|/b_1(z_L)$. The dashed and dotted lines show $f_{\text{NL}} = \pm 50$ and ± 100 , respectively, while the thin solid lines show l^{-2} with an arbitrary normalization. 190

- 6.12 Angular power spectra of the galaxy-galaxy correlation, C_l^h (thick dotted lines), the galaxy-convergence cross-correlation, $C_l^{h\kappa}$ (thick solid lines), and the convergence-convergence correlation, C_l^κ (thick dashed lines) for the Gaussian initial condition ($f_{\text{NL}} = 0$). The four panels show the same populations of galaxies and clusters of galaxies as in Fig. 6.10. We also show the galaxy shot noise, $1/n_L$ (thin dotted lines) as well as the source shape noise, σ_γ^2/n_S (thin dashed lines), for $N_L = 10^6$, $\sigma_\gamma = 0.3$, and $n_S = 3.5 \times 10^8 \text{ sr}^{-1}$. We find $1/n_L \ll C_l^h$ and $\sigma_\gamma^2/n_S \ll C_l^\kappa$ for $l \lesssim 100$ 192
- 6.13 Same as Fig. 6.10, with the expected $1\text{-}\sigma$ uncertainties for full-sky lens surveys and a single lens redshift. Adjacent bins are uncorrelated. The open (filled) boxes show the binned uncertainties with (without) the cosmic variance term due to the cosmic shear field included. We used Eq. (6.32) for the open boxes, and Eq. (6.32) with $C_l^{h\kappa} = 0 = C_l^\kappa$ for the filled boxes. We use logarithmic bins of $\Delta l = 0.23l$. Note that the uncertainties are calculated for a single lens redshift slice, and thus they will go down as we add more lens redshift slices. 193
- 6.14 Angular power spectrum of the galaxy-CMB lensing, $C_l^{h\kappa}$, at various multipoles as a function of the lens redshift, z_L . We have divided $C_l^{h\kappa}$ by its maximum value. The solid, dotted, dashed, dot-dashed, and triple-dot-dashed lines show $l = 10, 50, 100, 350$, and 1000 , respectively. 196
- 6.15 Imprints of the local-type primordial non-Gaussianity in the galaxy-CMB lensing power spectrum, $l(l+1)C_l^{h\kappa}/(2\pi)$, for the same populations of lens galaxies as in Fig. 6.3. The solid, dashed, and dotted lines show $f_{\text{NL}} = 0, \pm 50$, and ± 100 , respectively. 198
- 6.16 Same as Fig. 6.15, but for high- z lens galaxies with $b_1 = 2$ at $z_L = 2$ (top-left), $b_1 = 2.5$ at $z_L = 3$ (top-right), $b_1 = 3$ at $z_L = 4$ (bottom-left), and $b_1 = 3.5$ at $z_L = 5$ (bottom-right). 200
- 6.17 Same as Fig. 6.11, but for high- z lens galaxies with $b_1 = 2$ at $z_L = 2$ (top-left), $b_1 = 2.5$ at $z_L = 3$ (top-right), $b_1 = 3$ at $z_L = 4$ (bottom-left), and $b_1 = 3.5$ at $z_L = 5$ (bottom-right). 201
- 6.18 Angular power spectra of the galaxy-galaxy correlation, C_l^h (thick dotted lines), the galaxy-convergence cross-correlation, $C_l^{h\kappa}$ (thick solid lines), and the convergence-convergence correlation, C_l^κ (thick dashed lines) for the Gaussian initial condition ($f_{\text{NL}} = 0$). The four panels show the same populations of galaxies and clusters of galaxies as in Fig. 6.15. We also show the galaxy shot noise, $1/n_L$ (thin dotted lines) as well as the lens reconstruction noise, N_l^κ (thin dashed lines), for $N_L = 10^6$ and $N_l^\kappa \simeq 6 \times 10^{-8} \text{ sr}^{-1}$ (for multipoles much smaller than that corresponds to the beam size of $4'$). We find $1/n_L \ll C_l^h$ and $N_l^\kappa \ll C_l^\kappa$ for $l \lesssim 100$ 202

6.19	Same as Fig. 6.15, but with 1-sigma uncertainty due to the shape noise of source galaxies (filled box, Eq. (6.27)) and full error budget (empty box, diagonal of Eq. (6.25)) including the cosmic variance. We use the multipole bins of size $\Delta l = 0.23l$. For uncertainty of CMB lensing reconstruction, We assume the nearly-perfect reference experiment of Hu & Okamoto (2002): white detector noise $\Delta_T = \Delta_P/\sqrt{2} = 1 \mu K$ arcmin, and FWHM of the beam $\sigma = 4'$	203
6.20	Same as Fig. 6.18, but for the high redshift lens galaxies shown in Fig. 6.16. For these populations (and with $N_L = 10^6$), the shot noise is about the same as the galaxy power spectrum, i.e., $C_l^h \simeq 1/n_L$	204
6.21	Same as Fig. 6.19, but for the high redshift lens galaxies shown in Fig. 6.16.	205
A.1	Result of FFTW stored in the complex array with half size. Shaded region shows DC component which has a real value. This diagram is for an odd N	250
A.2	Same as Figure A.1, but for an even N . DC components and Nyquist components are shaded with grey.	251
E.1	Comparison between the power spectra calculated from 5 different ZA runs using different starting redshifts, $z_{start} = 50, 100, 150, 300$, and 400 (from bottom to top). The power spectra are divided by the one from 2LPT with $z_{start} = 300$, to facilitate comparison.	277
E.2	Comparison between the power spectra calculated from 5 different 2LPT runs using different starting redshifts, $z_{start} = 50, 100, 150, 300$, and 400 (from bottom to top). The power spectra are divided by the one from 2LPT with $z_{start} = 300$, to facilitate comparison. Note that the power spectra with $z_{start} = 300$ and 400 agree very well, which suggests convergence at $z_{start} = 300$	278
F.1	Standard deviation of the matter power spectrum: analytical versus simulations. The symbols show the standard deviations directly measured from 120 independent N -body simulations whose box sizes are $L = 512 \text{ Mpc}/h$ (60 realizations for $k < 0.24h/\text{Mpc}$) and $L = 256 \text{ Mpc}/h$ (60 realizations for $0.24 < k < 0.5h/\text{Mpc}$). Each simulation contains 256^3 particles. The solid and dot-dashed lines show the analytical formula (Eq. (F.7)) with the 3rd-order PT non-linear $P(k)$ and the linear $P(k)$, respectively. Note that the graph is discontinuous at $k = 0.24h/\text{Mpc}$ because the number of k modes, N_k , for a given wavenumber k is different for different box sizes.	281
F.2	Residuals. We divide both analytical estimation and simulation results by the analytical formula (Eq. (F.7)) with the 3rd-order PT nonlinear $P(k)$	282

H.1	Histogram for the $1\text{-}\sigma$ errors on D calculated from 1000 Monte Carlo realizations generated with the best-fitting bias parameters of halo power spectrum of Millennium Simulation with $k_{max} = 1.5 \text{ h/Mpc}$ at $z = 6$. The error derived from the Fisher matrix is close to the mean, while the error from the marginalized one-dimensional likelihood function of Millennium Simulation is on the tail of the distribution. The probability of having an error on D greater than that from the Millennium Simulation is about 6%.	288
M.1	Top: Convergence-convergence angular power spectrum from two different methods: the exact calculation (Eq. M.14, symbols) and Limber's approximation (Eq. 6.33, solid lines). Bottom: Fractional differences between Limber's approximation and the exact integration. Symbols are the same as the top panel. Grey symbols show the absolute values of negative values.	335
M.2	Same as Fig. M.1, but for the galaxy-convergence cross angular power spectrum with $f_{NL} = 0$ and $b_1 = 1$	336
M.3	Same as Fig. M.1, but for the non-Gaussian correction (i.e., the term proportional to $\Delta b(k)$) to the galaxy-convergence cross angular power spectrum. We show the corrections with $f_{NL} = 1$ and $b_1 = 2$	337
M.4	Same as Fig. M.2, but for the galaxy-CMB lensing.	338
M.5	Same as Fig. M.3, but for the galaxy-CMB lensing.	339
M.6	Top: Same as Fig. 6.3, but also showing the exact result (Eq. M.13, thick lines) on top of the result from Limber's approximation (Eq. 6.30, thin lines). Bottom: Fractional difference of Limber's approximation relative to the exact result.	340
M.7	Same as Fig. M.6, but for larger R . Thick lines are the results of the exact integration, while the thin lines are Limber's approximation. The Limber approximation overpredicts $\Delta\Sigma(R)$ for large R , but the error is at most 5% for $R < 500 \text{ h}^{-1} \text{ Mpc}$. The error is the largest for the lowest z_L , as a physical separation R at a lower redshift corresponds to a larger angular separation on the sky.	342
M.8	Fractional differences in the non-Gaussian correction terms, $\Delta\Sigma_{nG}$, from Limber's approximation and the exact integration. Using Limber's approximation, we overpredict the non-Gaussian correction by $\sim 20\%$ at $R = 300 \text{ h}^{-1} \text{ Mpc}$ for $z_L = 0.3$	343
M.9	Same as Fig. 6.4, but with the exact integration instead of Limber's approximation.	344

N.1	Noise power spectra of Cosmic Microwave Background lensing reconstruction. We show the noise power spectrum of $d_{\Theta\Theta}$ (<i>solid line</i>), $d_{\Theta E}$ (<i>dashed line</i>), $d_{\Theta B}$ (<i>dot-dashed line</i>), d_{EE} (<i>dots-dashed line</i>), d_{EB} (<i>long-dashed line</i>), and minimum variance estimator d_{mv} (<i>orange solid line</i>). For comparison, we also show the convergence power spectrum C_l^κ (<i>red line</i>). Noise power spectrum is calculated for Planck satellite: $\Delta_T = 35.4 \mu\text{Karcmin}$, $\Delta_P = 63.1 \mu\text{Karcmin}$, and $\sigma = 7 \text{ arcmin}$ as described in Appendix A of Zaldarriaga et al. (2008). .	351
N.2	Same as Figure N.1, but for the <i>nearly perfect</i> experiment quoted by Hu & Okamoto (2002). We use $\Delta_T = 1 \mu\text{Karcmin}$, $\Delta_P = \sqrt{2} \mu\text{Karcmin}$, and $\sigma = 4 \text{ arcmin}$. In Chapter6, we estimate the noise power spectrum (N_l^κ) by minimum variance estimator (<i>orange solid line</i>).	352

Chapter 1

Introduction

We are living in a golden age of cosmology. By ‘golden age’, we mean not only that all observations converge to the concordance cosmology model and that we can measure most of the cosmological parameters accurately (Komatsu et al., 2010), but also that we face many theoretical challenges, mostly on the nature of the constituents of the concordance model. Among them, the nature of dark matter, of dark energy, and of inflation are the three biggest questions in modern cosmology.

Inflation (Guth, 1981; Sato, 1981; Albrecht & Steinhardt, 1982; Linde, 1982), which presumably took place in the very early stages of the universe, was a phase of accelerated expansion that enlarged the universe by at least 10^{27} times. With such a large amount of expansion, inflation naturally provided a flat, homogeneous, and isotropic universe. At the same time, inflation continually stretched the quantum vacuum fluctuations outside of the Hubble horizon, seeding the cosmic structures such as galaxies (Guth & Pi, 1982; Hawking, 1982; Starobinsky, 1982; Bardeen et al., 1983). While this idea is attractive, we still do not know the underlying theory behind such an accelerated expansion. The situation is the same for dark energy; we, again, do not have any clues as to what it is. The only thing we know about dark energy is that it is responsible for the current accelerating expansion of the Universe (Frieman et al., 2008, for a review). For both cases, the clues have to come from observations.

The large scale structure of the universe is a promising probe of both periods of cosmic acceleration. The reason is simple: inflation sets the initial conditions, and dark energy controls the distances and the growth of cosmic structures. 1) The seed perturbations for cosmic structures are predicted by most inflation models to be nearly scale invariant and to obey Gaussian statistics. However, the extent to which they deviate from perfectly scale invariant, Gaussian fluctuations depends on the details of the model. Therefore, by accurately measuring deviations from scale invariance as well as from Gaussianity, we can constrain the physics of inflation. 2) Dark energy controls the expansion history of the universe. Expansion history, in turn, affects how we observe galaxies on the sky: galaxies are

cataloged by their angular positions and redshifts, and the relation between the coordinates and physical distances depends on the expansion rate and angular diameter distance. Also, the expansion rate slows down the gravitational evolution of cosmic structure. Therefore, by measuring the distances and growth cosmic structures, we can constrain the properties of dark energy.

In this *dissertation*, we shall develop the necessary theoretical tools to extract the information about inflation and dark energy from the large scale structure of the Universe. Our work is motivated by *Hobby-Eberly Telescope Dark Energy eXperiment (HETDEX)* (Hill et al., 2008), which will observe about million Ly- α emitters at high redshifts: $1.9 < z < 3.5$. *HETDEX* will start observing from Fall 2011, and the method we develop throughout this dissertation will be used to analyze the data from the survey.

The biggest challenge for studying inflation and dark energy from the large scale structure is the accurate modeling of the non-linearities. While density fluctuations generated from inflation are very small, $\sim 10^{-5}$, the subsequent evolution powered by gravitational instability magnifies them so that the linear theory is applicable only for the very large scales ($k \lesssim 0.1$ [h/Mpc] at $z = 0$).

In order to model the nonlinearities, we shall use Eulerian perturbation theory. Over the last two decades, the non-linear perturbation theory, including modeling of non-linear galaxy power spectra, had been studied actively (see Bernardeau et al., 2002, for a review). In particular, a lot of efforts have been devoted into understanding the non-linear power spectrum at $z \sim 0$. However, these earlier works only show that perturbation approach would not provide accurate descriptions of the power spectrum at $z \sim 0$ due to too strong non-linearity. Therefore, Perturbation Theory had never been applied to the real data such as *Two-degree Field Galaxy Redshift Survey (2dFGRS)* (Cole et al., 2005) or the *Sloan Digital Sky Survey (SDSS)* (Tegmark et al., 2004; Seljak et al., 2005), as non-linearities are too strong for Perturbation Theory to be valid at low redshifts, $z < 1$ (e.g., Meiksin et al., 1999). For those low redshift surveys, we have to rely on the empirical fitting formula calibrated to the N-body simulations (Peacock & Dodds, 1996; Smith et al., 2003).

On the other hand, Perturbation Theory is expected to perform better at high redshifts, i.e., $z > 1$, because of weaker non-linearity. Plus, Perturbation Theory provides the natural framework to take into account the nonlinearities in the redshift space distortion and in the galaxy bias (Chapter 2). Therefore, modeling the nonlinearities at high redshift Universe with Perturbation Theory is one of the main themes of this dissertation.

Throughout this dissertation, we shall focus on the statistical correlations of the galaxy distribution. It is because, from theory, we can only predict the statistical properties of galaxies thanks to the stochastic nature of the cosmological perturbation.

The galaxy two-point correlation function $\xi(\mathbf{r})$ is defined as an expected mean excess number of ‘pairs’ at a given separation \mathbf{r} . That is, the probability P_2 of finding two galaxies separated by a distance \mathbf{r} is given by (Peebles, 1980)

$$P_2(\mathbf{r}) = P_1^2 [1 + \xi(\mathbf{r})], \quad (1.1)$$

where P_1 is the probability of finding one galaxy at a position in the universe. The galaxy power spectrum, the Fourier transform of the two-point correlation function of galaxy, is defined as

$$\langle \delta(\mathbf{k})\delta(\mathbf{k}') \rangle = (2\pi)^3 P(\mathbf{k}) \delta^D(\mathbf{k} + \mathbf{k}'), \quad (1.2)$$

where $\delta(\mathbf{k})$ is the Fourier transform of $\delta(\mathbf{x}) \equiv \rho(\mathbf{x})/\bar{\rho} - 1$ and δ^D is the Dirac delta function. The galaxy power spectrum has been widely used for extracting cosmological information from the galaxy survey data.

Yet, we will not be able to exploit all the information in the large scale structure if we use the galaxy power spectrum only. The higher order correlation functions also contains valuable cosmological information. We define the galaxy three-point correlation in a similar manner to the two-point correlation function. The probability P_3 finding three galaxies at \mathbf{x}_1 , \mathbf{x}_2 and \mathbf{x}_3 is given by the sum of the cosmic mean probability P_1^3 , the probability of having two ‘clustered’ galaxies and one random galaxy $P_1^3 \xi$, and having three ‘clustered’ galaxies $P_1^3 \zeta$ as (Peebles, 1980)

$$P_3(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = P_1^3 [1 + \xi(\mathbf{x}_{12}) + \xi(\mathbf{x}_{23}) + \xi(\mathbf{x}_{31}) + \zeta(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)]. \quad (1.3)$$

The galaxy bispectrum, the Fourier transform of the galaxy three-point correlation function, is defined as

$$\langle \delta(\mathbf{k}_1)\delta(\mathbf{k}_2)\delta(\mathbf{k}_3) \rangle = (2\pi)^3 B(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) \delta^D(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3), \quad (1.4)$$

and is expected to be a powerful probe of inflation, non-linear structure formation, and astrophysics such as galaxy formation.

Along with the correlation functions of galaxies, weak gravitational lensing also provides a powerful probe of the growth of structure as the cosmic shear field is a proxy for the total matter distribution along the line of sight. Therefore, by using weak lensing method, we can get a direct information about the total matter distribution without using

tracers (e.g. galaxies), or by combining the galaxy distribution with weak lensing, we can test the bias scheme itself.

This dissertation is organized as follows. In the first part (Chapter 2 to 4), we present the Eulerian perturbation theory modeling of the galaxy power spectrum from the redshift surveys.

In chapter 2, we review Eulerian cosmological perturbation theory, and calculate the non-linear galaxy power spectrum we shall observe from galaxy surveys. There are three non-linearities that changes the galaxy power spectrum different from the linear theory prediction: namely, non-linear growth of density field, non-linear bias, and non-linear redshift space distortion. We first study these non-linearities separately, and combine the effect at the end of the chapter. For all cases, we also calculate the leading order correction to the power spectrum by primordial local non-Gaussianity.

In chapter 3, we test the validity and applicability of Eulerian perturbation theory at high redshifts by comparing the resulting non-linear matter power spectrum of Eulerian perturbation theory to the power spectrum measured from a series of N-body simulations, in both real and redshift space.

In chapter 4, we compare the non-linear galaxy power spectrum calculated from perturbation theory and local bias ansatz to the Millennium Simulation. As we fit the power spectrum with the three bias parameters, we also study the effect of such a fitting on cosmological parameter estimation with a specific example of the distance measurement.

Chapter 5 and 6 consist of the second part of the dissertation, where we discuss about the signature of the primordial non-Gaussianity in the large scale structure of the universe.

In chapter 5, we calculate the galaxy bispectrum by using two different methods: the Matarrese-Lucchin-Bonometto (MLB) formula and the locality of galaxy bias. We include all the dominant terms of the galaxy bispectrum in the presence of, physically-motivated, local form of primordial non-Gaussianity in the curvature perturbation, and show that the signature of the local type primordial non-Gaussianity is much stronger than previous calculation. This result indicate that the galaxy bispectrum is one of the most powerful probe of the primordial non-Gaussianity.

In chapter 6, we study the galaxy-galaxy, and galaxy-CMB weak lensing method on large angular separation for both Gaussian and non-Gaussian initial conditions.

Finally, in the third part (chapter 7) we present the method of estimating the galaxy power spectrum and the galaxy bispectrum from both N-body simulation and real galaxy distribution from galaxy surveys.

Chapter 2

Eulerian cosmological Perturbation Theory

Cosmological Perturbation Theory (Bernardeau et al., 2002, and references therein) provides the unique theoretical framework of studying the evolution of the density and velocity fields of matter fluctuation in the Universe. While the non-linear gravitational instability breaks down the validity of linear Perturbation Theory on smaller scales ($k \gtrsim 0.1 [h/\text{Mpc}]$ at present), we expect to model the non-linear evolution of cosmic matter field by using higher order Perturbation Theory. Yet, there is a fundamental limitation of Perturbation Theory: it improves upon the linear theory only in the very small region when non-linearity is too strong (this happens around $z \sim 0$), and breaks down on the scales where non-perturbative effects such as shell-crossing and violent relaxation take place. Therefore, we define *quasi-nonlinear regime* where higher-order Perturbation Theory correctly models the non-linear evolution of cosmic matter field.

Quasi-nonlinear regime in standard Perturbation Theory satisfies following three conditions.

- [1] Quasi-nonlinear regime is small compare to the Hubble length so that evolution of cosmic matter field is governed by Newtonian fluid equations.
- [2] Quasi-nonlinear regime is large enough to neglect baryonic pressure so that we can treat dark matter and baryon as a single component of pressureless matter.
- [3] In quasi-nonlinear regime, vorticity developed by non-linear gravitational interaction is negligibly small.

With these three conditions, we can approximate cosmic matter field as a pressureless, single component Newtonian fluid which is completely described by its density contrast and velocity gradient. In Section 2.1, we shall present the Perturbation Theory calculation of the non-linear evolution of cosmic matter field based on these conditions.

Extended studies of standard Perturbation Theory by relaxing one of these conditions are also available in literature. Noh & Hwang (2008) have studied the single fluid

equation in the full General Relativistic context and show that, if one use the proper gauge (temporal comoving gauge, to be specific), the General Relativistic perturbation equations *exactly* coincide with their Newtonian counterparts up to 2nd order; thus, the General Relativistic correction appears from third order in density perturbation. In Noh et al. (2009), we showed that the 3rd order General Relativistic correction term is sub-dominant on sub-horizon scales, so that the Newtonian PT approach is valid on quasi-nonlinear regime. It is because the purely General Relativistic effect comes through the gravitational potential, and the gravitational potential is much smaller than the density field on sub-horizon scales. At the same time, this correction term increases on large scales comparable to the Hubble radius, because gravitational potential sharply increases as $P_\phi \propto k^{n_s-4}$; thus, it eventually exceeds the linear power spectrum near horizon scale. Shoji & Komatsu (2009) have included a pressureful component to the analysis and have found a perturbative solution of double-fluid equations up to 3rd order. Finally, Pueblas & Scoccimarro (2009) measures the vorticity power spectrum from N-body simulations, and show that vorticity effect on density power spectrum is indeed negligible in the quasi-nonlinear regime.

These studies have indicated that non-linear effects coming from violating three conditions are not significant on scales which are most relevant for upcoming high redshift galaxy surveys. One notable exception is when including massive neutrinos. Massive neutrinos suppress the linear power spectrum below the neutrino free streaming scale (Takada et al., 2006), and change nonlinear matter power spectrum, correspondingly. Although the non-linear effect to the matter power spectrum is marginal due to the small energy fraction of neutrino, $f_\nu \equiv \Omega_\nu/\Omega_m$, this effect has to be included in order to measure neutrino mass from galaxy surveys (Shoji & Komatsu, 2009; Saito et al., 2009).

Once we model the non-linear evolution of density field and velocity field of cosmic matter fluctuation, we can calculate the galaxy power spectrum we would observe from galaxy surveys. Here, we have to model two more non-linearities: nonlinear redshift space distortion and nonlinear bias. In order to understand those non-linear effects separately, we first present the non-linear galaxy power spectrum in real space in Section 2.5, then present the non-linear redshift space matter power spectrum in Section 2.6. We combine all the non-linearities and present the non-linear galaxy power spectrum in redshift space in Section 2.7. For each section, we also analyze the effect of primordial non-Gaussianity on the power spectrum of large scale structure in the Perturbation Theory framework.

While we focus on the Eulerian Perturbation Theory in this chapter, Lagrangian Perturbation Theory provide yet another intuition on the non-linear growth of the structure.

In particular, Lagrangian perturbation theory (or its linear version which is also known as *Zel'dovich approximation*) is widely used to generate the initial condition for cosmological N-body simulations. We review the Lagrangian Perturbation Theory in Appendix E.

Although the material in this chapter is self-contained, we by no means aim for the complete review. For more detailed review on Perturbation Theory, we refer readers to Bernardeau et al. (2002).

2.1 Eulerian Perturbation Theory solution

We review calculation of non-linear Eulerian Perturbation Theory following the pioneering work in the literature (Vishniac, 1983; Fry, 1984; Goroff et al., 1986; Suto & Sasaki, 1991; Makino et al., 1992; Jain & Bertschinger, 1994; Scoccimarro & Frieman, 1996). We treat dark matter and baryons as pressureless dust particles, as we are interested in the scales much larger than the Jeans length. We also assume that peculiar velocity is much smaller than the speed of light, which is always an excellent approximation, and that the fluctuations we are interested in are deep inside the horizon; thus, we treat the system as Newtonian. Then, the evolution of the matter fluctuation, $\delta(\mathbf{x}, \tau) \equiv \rho(\mathbf{x}, \tau)/\bar{\rho}(\tau) - 1$, follows Newtonian fluid equations in expanding universe:

$$\dot{\delta} + \nabla \cdot [(1 + \delta)\mathbf{v}] = 0, \quad (2.1)$$

$$\dot{\mathbf{v}} + (\mathbf{v} \cdot \nabla)\mathbf{v} = -\mathcal{H}\mathbf{v} - \nabla\phi, \quad (2.2)$$

$$\nabla^2\phi = 4\pi G a^2 \bar{\rho} \delta, \quad (2.3)$$

where the dots denote $\partial/\partial\tau$ (τ is the conformal time), ∇ denotes $\partial/\partial\mathbf{x}$ (\mathbf{x} is the comoving coordinate), $\mathbf{v} = d\mathbf{x}/d\tau$ is the peculiar velocity field, and ϕ is the peculiar gravitational potential field from density fluctuations, and $\mathcal{H} \equiv d\ln a/d\tau = aH$. As we ignore the vorticity, \mathbf{v} is curl-free, which motivates our using $\theta \equiv \nabla \cdot \mathbf{v}$, the velocity divergence field.

In Fourier space, the Newtonian fluid equations become two coupled integro-differential equations for $\delta_{\mathbf{k}}(\tau)$ and $\theta_{\mathbf{k}}(\tau)$. Using equation (2.3) and the Friedmann equation, we write

the continuity equation [Eq. (2.1)] and the Euler equation [Eq. (2.2)] in Fourier space

$$\begin{aligned} & \frac{\partial \delta_{\mathbf{k}}}{\partial \tau}(\tau) + \theta_{\mathbf{k}}(\tau) \\ &= - \int \frac{d^3 k_1}{(2\pi)^3} \int d^3 k_2 \delta^D(\mathbf{k}_1 + \mathbf{k}_2 - \mathbf{k}) \frac{\mathbf{k} \cdot \mathbf{k}_1}{k_1^2} \theta_{\mathbf{k}_1}(\tau), \delta_{\mathbf{k}_2}(\tau), \end{aligned} \quad (2.4)$$

$$\begin{aligned} & \frac{\partial \theta_{\mathbf{k}}}{\partial \tau}(\tau) + \mathcal{H}(\tau) \theta_{\mathbf{k}}(\tau) + \frac{3}{2} \mathcal{H}^2(\tau) \Omega_{\text{m}}(\tau) \delta_{\mathbf{k}}(\tau) \\ &= - \int \frac{d^3 k_1}{(2\pi)^3} \int d^3 k_2 \delta^D(\mathbf{k}_1 + \mathbf{k}_2 - \mathbf{k}) \frac{k^2 (\mathbf{k}_1 \cdot \mathbf{k}_2)}{2k_1^2 k_2^2} \theta_{\mathbf{k}_1}(\tau) \theta_{\mathbf{k}_2}(\tau), \end{aligned} \quad (2.5)$$

respectively. Note that left hand side of equations above are linear in perturbation variables, and non-linear evolution is described by the right hand side as coupling between different Fourier modes.

2.1.1 Linear solution for density field and velocity field

When density and velocity fluctuations are small, we can neglect the mode coupling terms in the right hand side of equation (2.4) and equation (2.5). Then, the continuity and the Euler equation are linearized as

$$\frac{\partial \delta_1(\mathbf{k}, \tau)}{\partial \tau} + \theta_1(\mathbf{k}, \tau) = 0, \quad (2.6)$$

$$\frac{\partial \theta_1(\mathbf{k}, \tau)}{\partial \tau} + \mathcal{H}(\tau) \theta_1(\mathbf{k}, \tau) + \frac{3}{2} \mathcal{H}^2(\tau) \Omega_{\text{m}}(\tau) \delta_1(\mathbf{k}, \tau) = 0. \quad (2.7)$$

Combining these two equations, we have a second order differential equation for $\delta_1(\mathbf{k}, \tau)$ as

$$\frac{\partial^2 \delta_1(\mathbf{k}, \tau)}{\partial \tau^2} + \mathcal{H}(\tau) \frac{\partial \delta_1(\mathbf{k}, \tau)}{\partial \tau} + \frac{3}{2} \mathcal{H}^2(\tau) \Omega_{\text{m}}(\tau) \delta_1(\mathbf{k}, \tau) = 0, \quad (2.8)$$

whose solution is given by

$$\delta_1(\mathbf{k}, a) = C_+(\mathbf{k}) H(a) \int_0^a \frac{da'}{a'^3 H(a')^3} + C_-(\mathbf{k}) H(a). \quad (2.9)$$

Here, the first term is a growing mode and the second term is a decaying mode.

Let us only consider a growing mode. There are two conventions in the literature about normalizing a growing mode. One normalization convention is requiring that a growing mode is equal to the scale factor in the matter dominated epoch: $D_+(a)|_{\text{EdS}} = a$. Here, EdS stands for the ‘Einstein de-Sitter’ Universe which is a flat, matter dominated universe. Therefore, a growing solution becomes

$$D_+(a) = \frac{5}{2} \Omega_{\text{m}} \frac{H(a)}{H_0} \int_0^a \frac{da'}{[a' H(a')/H_0]^3}, \quad (2.10)$$

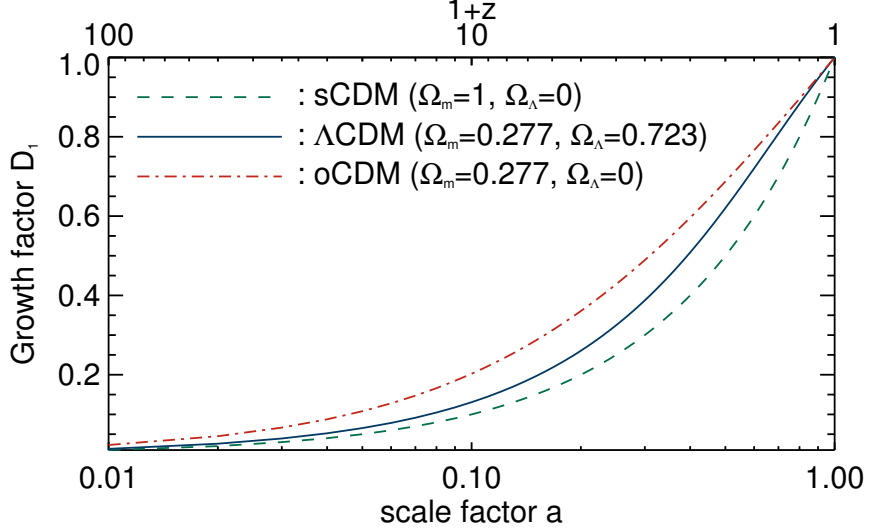


Figure 2.1: The linear growth factor, $D(a)$, for three different cosmologies: sCDM ($\Omega_m = 1$, $\Omega_\Lambda = 0$) Λ CDM ($\Omega_m = 0.277$, $\Omega_\Lambda = 0.723$) oCDM ($\Omega_m = 0.277$, $\Omega_\Lambda = 0$)

where Ω_m takes its present value. Another convention is normalizing its value to be unity at present:

$$D(a) = \frac{D_+(a)}{D_+(a=1)}. \quad (2.11)$$

Throughout this dissertation, we use the later convention, and call $D(a)$ the ‘linear growth factor’. Note that the two different conventions differ by a factor of 0.765 for the cosmological parameters in Table 1 (“WMAP+BAO+SN”) of Komatsu et al. (2009).

Figure 2.1 shows the linear growth factor for three different cosmologies: standard Cold Dark Matter (sCDM) model ($\Omega_m = 1$), Cold Dark Matter with cosmological constant (Λ CDM) model ($\Omega_m = 0.277$, $\Omega_\Lambda = 0.723$), and open Cold Dark Matter (oCDM) model ($\Omega_m = 0.277$, $\Omega_\Lambda = 0$). For given density fluctuations today, at high redshifts, the density fluctuations have to be larger for the oCDM universe, and smaller for sCDM universe compare to the standard Λ CDM universe. It is because in Λ CDM and oCDM universe, energy density is dominated by dark energy and curvature, respectively; both of them retard the growth of structure by speeding up the expansion of universe faster than the sCDM universe.

We calculate the velocity gradient field $\theta_1(\mathbf{k}, \tau)$ as

$$\theta_1(\mathbf{k}, \tau) = -\frac{\partial \delta_1(\mathbf{k}, \tau)}{\partial \tau} = -\frac{\delta_1(\mathbf{k}, \tau)}{D(\tau)} \frac{dD(\tau)}{d\tau} = -f(\tau)\mathcal{H}(\tau)\delta_1(\mathbf{k}, \tau), \quad (2.12)$$

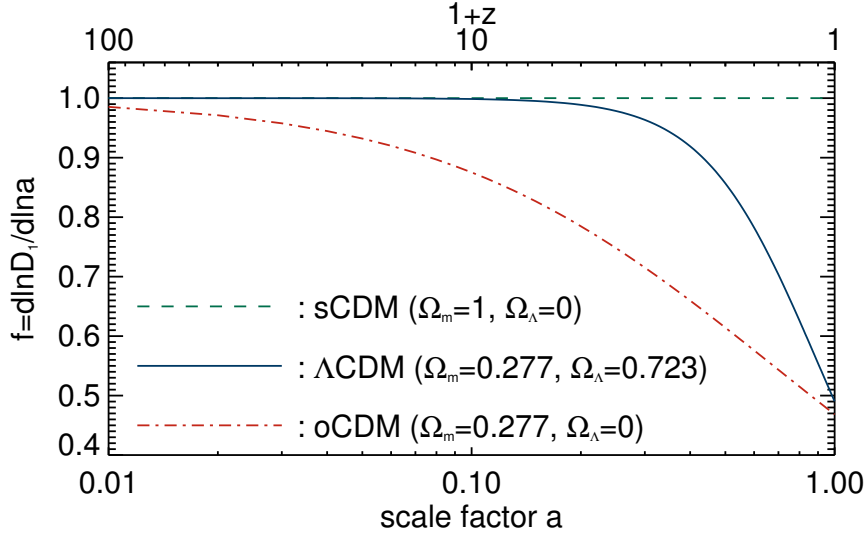


Figure 2.2: The logarithmic derivative of the linear growth factor, $f(a) \equiv d \ln D / d \ln a$, for three different cosmologies: sCDM ($\Omega_m = 1$, $\Omega_\Lambda = 0$) Λ CDM ($\Omega_m = 0.277$, $\Omega_\Lambda = 0.723$) oCDM ($\Omega_m = 0.277$, $\Omega_\Lambda = 0$)

where

$$f(\tau) \equiv \frac{d \ln D}{d \ln a} = \frac{1}{2} \left(\frac{H_0}{aH(a)} \right)^2 \left[\frac{5\Omega_m}{D_+(a)} - \frac{3\Omega_m}{a} - 2(1 - \Omega_m - \Omega_\Lambda) \right], \quad (2.13)$$

is the logarithmic growth rate. Here, Ω_m , Ω_Λ are the values at present.

Figure 2.2 shows the logarithmic growth rate for three different cosmologies: sCDM model, Λ CDM model, and oCDM model. When universe is flat, matter dominated, $f = 1$, and linear growth is slowing down once cosmological constant start to affect the expansion in Λ CDM universe. For oCDM universe, the growth rate is always slower than Λ CDM or sCDM universe.

What about the wave vector \mathbf{k} dependence? We can divided the k -dependence of linear perturbation by two parts: k -dependence due to the generation of primordial perturbation from inflation, and k -dependence due to the subsequent evolution of density perturbation to matter epoch.

Inflation stretches the quantum fluctuation outside of horizon, and generate the primordial curvature perturbation, $\zeta(\mathbf{k})$, which is conserved outside of horizon even if the equation of state $w \equiv P/\rho$ changing (Mukhanov et al., 1992). The Bardeen's potential, a

relativistic generalization of the peculiar gravitational potential¹, $\Phi_H(\mathbf{k})$ is also conserved outside of horizon, but only for constant w , and for the universe dominated by a perfect fluid whose equation of state is w , it is related to the primordial curvature perturbation $\zeta(\mathbf{k})$ by

$$\Phi_H(\mathbf{k}) = \frac{3+3w}{5+3w}\zeta(\mathbf{k}). \quad (2.14)$$

When universe is dominated by radiation or matter, expansion of the universe decelerates, and the wavemodes once stretched outside of horizon by inflation start to re-enter inside of horizon. As $w = 0$ for matter, the Bardeen's potential of the mode which re-enter the horizon during matter era is $\Phi_H(\mathbf{k}) = 3/5\zeta(\mathbf{k})$ at horizon crossing time. Inside of horizon, the Bardeen's potential $\Phi(\mathbf{k}, a)$ is related to the density field² by the Poisson equation:

$$k^2\Phi(\mathbf{k}, a) = 4\pi G a^2 \bar{\rho}(a)\delta_1(\mathbf{k}, a) = \frac{3}{2}H_0^2\Omega_m(1+z)\delta_1(\mathbf{k}, a). \quad (2.15)$$

We denote the time evolution of the peculiar gravitational potential as $g(z)$, and it is apparent from equation (2.15) that

$$g(z) = (1+z)D(z). \quad (2.16)$$

Then, we rewrite the Bardeen's potential at large scales as

$$\Phi(\mathbf{k}, a) = g(z)\Phi(\mathbf{k}), \quad (2.17)$$

where $\Phi(\mathbf{k})$ is the Bardeen's potential extrapolated at present epoch³; thus, it is related to the horizon crossing value as $\Phi(\mathbf{k})/\Phi_H(\mathbf{k}) = 1/g(z)|_{\text{EdS}} = D_+(a=1) \simeq 0.765$. The numerical value is for cosmological parameters in Table 1 ("WMAP5+BAO+SN") of Komatsu et al. (2009).

On the other hand, for the wave modes re-enter horizon during radiation era, as perturbation of dominant component (radiation) cannot grow due to its pressure, peculiar gravitational potential decays and matter density contrast can only grow logarithmically. Therefore, the amplitude of sub-horizon perturbations are suppressed relative to the super-horizon perturbations. Plus, at that time baryons were tightly coupled to photon, and could not contribute to the growth of matter fluctuation.

¹Note that Φ_H has an opposite sign of the Newtonian peculiar gravitational potential ϕ we defined earlier.

²To be precise, this equation holds for *comoving gauge* where $\delta u = 0$.

³Throughout this dissertation, we consistently follow this convention: a dynamical quantity, such as Φ , $\delta_1 P_L$, written without explicit time (redshift) dependence denotes the quantity extrapolated to its present value.

In order to take into account these evolution, we need to solve the perturbed Einstein equation and Boltzmann equation for coupled dark matter, photon, baryon, neutrino system. There are many publically available code for calculating such equation systems; among them, CAMB⁴ and CMBFAST⁵ are most widely used in the cosmology community.

These codes calculate so called the ‘transfer function’ $T(k)$. The transfer function $T(k)$ encodes the evolution of density perturbation throughout the matter-radiation equality and CMB last scattering. Since transfer function is defined as the relative changes of small scale modes (which enter horizon earlier) compared to the large scale modes (which enter horizon during matter dominated epoch), the transfer function is unity on large scales: $T(k) = 1$. Therefore, the effects of the retarded growth in the radiation epoch and tight coupling between baryon-photon can be taken into account by multiplying the transfer function to the left hand side of equation (2.15):

$$\delta_1(\mathbf{k}, z) = \frac{2}{3} \frac{k^2 T(k)}{H_0^2 \Omega_m} D(z) \Phi(\mathbf{k}) \equiv \mathcal{M}(k) D(z) \Phi(\mathbf{k}). \quad (2.18)$$

Primordial curvature perturbation predicted by the most inflationary models, and confirmed by observations such as *WMAP* and *SDSS*, is characterized by *nearly* a scale invariant power spectrum. Therefore, we conventionally parametrize the shape of the primordial curvature power spectrum as

$$P_\zeta(k) = 2\pi^2 \Delta_{\mathcal{R}}^2(k_p) \left(\frac{k}{k_p} \right)^{n_s(k_p) - 4 + \frac{1}{2} \alpha_s \ln\left(\frac{k}{k_p}\right)}, \quad (2.19)$$

where we use three parameters: amplitude of primordial power spectrum $\Delta_{\mathcal{R}}^2$, spectral tilt n_s , and running index α_s . Here, k_p is a pivot wavenumber⁶. Note that the perfectly scale invariant primordial perturbation corresponds to $n_s = 1$, $\alpha_s = 0$.

Combining the primordial power spectrum [Eq (2.19)] and the late time linear evolution [Eq (2.18)], we calculate the linear matter power spectrum as

$$P_L(k) = \frac{8\pi^2}{25} \frac{[D_+(a=1)]^2}{H_0^4 \Omega_m^2} \Delta_{\mathcal{R}}^2(k_p) D^2(z) T^2(k) \left(\frac{k}{k_p} \right)^{n_s(k_p) + \frac{1}{2} \alpha_s \ln\left(\frac{k}{k_p}\right)}. \quad (2.20)$$

⁴<http://camb.info>

⁵<http://www.cmbfast.org>

⁶Different authors, surveys use different value of k_p . Komatsu et al. (2009) uses $k_p \equiv 0.002 \text{ [Mpc}^{-1}]$ for *WMAP*, while Reid et al. (2010) uses $k_p \equiv 0.05 \text{ [Mpc}^{-1}]$ for *SDSS*.

Alternatively, we can also normalize the linear power spectrum by fixing σ_8 , a r.m.s. density fluctuation smoothed by the spherical top-hat filter of radius 8 Mpc/ h , whose explicit formula is given by

$$\sigma_8^2 \equiv \int d \ln k \frac{k^3 P_L(k)}{2\pi^2} W^2(kR), \quad (2.21)$$

where

$$W(kR) = 3 \left[\frac{\sin(kR)}{k^3 R^3} - \frac{\cos(kR)}{k^2 R^2} \right]$$

with $R = 8 \text{ Mpc}/h$.

2.1.2 Non-linear solution for density field and velocity field

Let us come back to the original non-linear equations. In order to solve these coupled integro-differential equations, we shall expand $\delta_{\mathbf{k}}(\tau)$ and $\theta_{\mathbf{k}}(\tau)$ perturbatively by using the n -th power of linear density contrast $\delta_1(\mathbf{k}, \tau)$ as a basis:

$$\begin{aligned} \delta_{\mathbf{k}}(\tau) = & \sum_{n=1}^{\infty} \int \frac{d^3 q_1}{(2\pi)^3} \cdots \int \frac{d^3 q_{n-1}}{(2\pi)^3} \int d^3 q_n \delta^D(\mathbf{k} - \sum_{i=1}^n \mathbf{q}_i) \\ & \times F_n^{(s)}(\mathbf{q}_1, \mathbf{q}_2, \cdots, \mathbf{q}_n, \tau) \delta_1(\mathbf{q}_1, \tau) \cdots \delta_1(\mathbf{q}_n, \tau) \end{aligned} \quad (2.22)$$

$$\begin{aligned} \theta_{\mathbf{k}}(\tau) = & -f(\tau) \mathcal{H}(\tau) \sum_{n=1}^{\infty} \int \frac{d^3 q_1}{(2\pi)^3} \cdots \int \frac{d^3 q_{n-1}}{(2\pi)^3} \int d^3 q_n \delta^D(\mathbf{k} - \sum_{i=1}^n \mathbf{q}_i) \\ & \times G_n^{(s)}(\mathbf{q}_1, \mathbf{q}_2, \cdots, \mathbf{q}_n, \tau) \delta_1(\mathbf{q}_1, \tau) \cdots \delta_1(\mathbf{q}_n, \tau). \end{aligned} \quad (2.23)$$

Here, $f(\tau) = d \ln D / d \ln a$ with the linear growth factor $D(a)$, and $F_n^{(s)}$ and $G_n^{(s)}$ are symmetrized kernels, which characterize coupling between different wave modes. We introduce $f(\tau) \mathcal{H}(\tau)$ factor in equation (2.23) motivated by the linear relation between density contrast and velocity divergence: $\theta_1(\mathbf{k}, \tau) = -f(\tau) \mathcal{H}(\tau) \delta_1(\mathbf{k}, \tau)$. In this definition the first order kernels become unity: $F_1^{(s)} = G_1^{(s)} = 1$.

As we know the evolution of $\delta_1(\mathbf{q}, \tau)$ from the linear theory, calculating $F_n^{(s)}$ and $G_n^{(s)}$ will complete the solution. The standard procedure (Goroff et al., 1986; Jain & Bertschinger, 1994) is calculating the *un-symmetized* kernels F_n and G_n first, then symmetrize them under changing arguments:

$$F_n^{(s)}(\mathbf{q}_1, \cdots, \mathbf{q}_n) = \frac{1}{n!} \sum_{\sigma} F_n(\mathbf{q}_{\sigma_1}, \cdots, \mathbf{q}_{\sigma_n}) \quad (2.24)$$

$$G_n^{(s)}(\mathbf{q}_1, \cdots, \mathbf{q}_n) = \frac{1}{n!} \sum_{\sigma} G_n(\mathbf{q}_{\sigma_1}, \cdots, \mathbf{q}_{\sigma_n}). \quad (2.25)$$

Here, sum is taken for all the permutations $\sigma \equiv (\sigma_1, \dots, \sigma_n)$ of the set $\{1, \dots, n\}$. By substituting the perturbative expansion of equation (2.22) and equation (2.23) back into the original equations, we get the equations of $F_n^{(s)}$ s and $G_n^{(s)}$ s. For example, the continuity equation [Eq. (2.4)] becomes

$$\begin{aligned}
& \sum_{n=1}^{\infty} \int \frac{d^3 q_1}{(2\pi)^3} \cdots \int \frac{d^3 q_{n-1}}{(2\pi)^3} \int d^3 q_n \delta^D(\mathbf{k} - \sum_{i=1}^n \mathbf{q}_i) \delta_1(\mathbf{q}_1, \tau) \cdots \delta_1(\mathbf{q}_n, \tau) \\
& \times \left[\frac{\partial F_n^{(s)}(\mathbf{q}_1, \dots, \mathbf{q}_n, \tau)}{\partial \tau} + n \mathcal{H} f F_n^{(s)}(\mathbf{q}_1, \dots, \mathbf{q}_n, \tau) - \mathcal{H} f G_n^{(s)}(\mathbf{q}_1, \dots, \mathbf{q}_n, \tau) \right] \\
& = (2\pi)^3 \mathcal{H} f \sum_{m=1}^{\infty} \sum_{l=1}^{\infty} \int \frac{d^3 q_{11}}{(2\pi)^3} \cdots \int \frac{d^3 q_{1m}}{(2\pi)^3} \int \frac{d^3 q_{21}}{(2\pi)^3} \cdots \int \frac{d^3 q_{2l}}{(2\pi)^3} \\
& \times \frac{\mathbf{k} \cdot (\mathbf{q}_{11} + \cdots + \mathbf{q}_{1m})}{|\mathbf{q}_{11} + \cdots + \mathbf{q}_{1m}|^2} G_m^{(s)}(\mathbf{q}_{11}, \dots, \mathbf{q}_{1m}) F_l^{(s)}(\mathbf{q}_{21}, \dots, \mathbf{q}_{2l}) \\
& \times \delta^D(\mathbf{k} - \sum_{i=1}^m \mathbf{q}_{1i} - \sum_{j=1}^l \mathbf{q}_{2j}) \delta_{11}(\mathbf{q}_1, \tau) \cdots \delta_{1m}(\mathbf{q}_m, \tau) \delta_{21}(\mathbf{q}_1, \tau) \cdots \delta_{2l}(\mathbf{q}_l, \tau).
\end{aligned}$$

Here, we use an identity from linear theory $\partial \delta_1 / \partial \tau = f \mathcal{H} \delta_1$ to replace the conformal-time (τ) derivatives. In order to isolate the equation for the kernels, we have to identify \mathbf{q}_i s to \mathbf{q}_{1i} s and \mathbf{q}_{2i} s. In principle, one can take all possible identifications and symmetrize the equation itself, but we follow the standard approach where (for given n) indexes are matched as $(\mathbf{q}_{11}, \dots, \mathbf{q}_{1m}) = (\mathbf{q}_1, \dots, \mathbf{q}_m)$, and $(\mathbf{q}_{21}, \dots, \mathbf{q}_{2l}) = (\mathbf{q}_{m+1}, \dots, \mathbf{q}_n)$. By matching indexes this way, the equations are not manifestly symmetric anymore, and as a result the kernels are not symmetric: we drop the superscript (s) and symmetrize kernels later by using equation (2.24) and equation (2.25).

Once \mathbf{q}_i s are identified, we can read off the equation for F_n and G_n from the equation above as

$$\begin{aligned}
& \frac{1}{f(\tau) \mathcal{H}(\tau)} \frac{\partial F_n(\mathbf{k}, \tau)}{\partial \tau} + n F_n(\mathbf{k}, \tau) - G_n(\mathbf{k}, \tau) \\
& = \sum_{m=1}^{n-1} \frac{\mathbf{k} \cdot \mathbf{k}_1}{k_1^2} G_m(\mathbf{k}_1, \tau) F_{n-m}(\mathbf{k}_2, \tau),
\end{aligned} \tag{2.26}$$

where we use the short hand notation of $F_n(\mathbf{k}, \tau) \equiv F_n(\mathbf{q}_1, \dots, \mathbf{q}_n, \tau)$, and $G_n(\mathbf{k}, \tau) \equiv G_n(\mathbf{q}_1, \dots, \mathbf{q}_n, \tau)$ with a constraint $\mathbf{k} = \mathbf{q}_1 + \cdots + \mathbf{q}_n$. Note that from the matching condition of \mathbf{q}_i s, $G_m(\mathbf{k}_1, \tau) = G_n(\mathbf{q}_1, \dots, \mathbf{q}_m, \tau)$ and $F_{n-m}(\mathbf{k}_2, \tau) = F_{n-m}(\mathbf{q}_{m+1}, \dots, \mathbf{q}_n, \tau)$ in the right hand side of equation (2.26). Also, Dirac delta function dictates $\mathbf{k}_1 + \mathbf{k}_2 = \mathbf{k}$. Similarly,

we find

$$\begin{aligned} & \frac{1}{f(\tau)\mathcal{H}(\tau)} \frac{\partial G_n(\mathbf{k}, \tau)}{\partial \tau} + \left[\frac{3}{2} \frac{\Omega_m(\tau)}{f^2(\tau)} + n - 1 \right] G_n(\mathbf{k}, \tau) - \frac{3}{2} \frac{\Omega_m(\tau)}{f^2(\tau)} F_n(\mathbf{k}, \tau) \\ &= \sum_{m=1}^{n-1} \frac{k^2(\mathbf{k}_1 \cdot \mathbf{k}_2)}{2k_1^2 k_2^2} G_m(\mathbf{k}_1, \tau) G_{n-m}(\mathbf{k}_2, \tau), \end{aligned} \quad (2.27)$$

from the Euler equation [Eq. (2.5)].

In general, the kernels F_n and G_n depend on time, and we have to solve complicated differential equations, [Eq. (2.26)] and [Eq. (2.27)]. However, it is well known that the kernels are extremely insensitive to the underlying cosmology, and the next-to-leading order correction to $P(k)$ can be correctly modeled as long as one uses the correct growth factor for $\delta_1(\mathbf{k}, \tau)$ (Bernardeau et al., 2002). Therefore, we shall calculate the kernels in the Einstein de-Sitter (spatially flat, matter-dominated) universe. In Einstein de-Sitter Universe, as $\Omega_m/f^2 = 1$, F_n and G_n are constant in time and equation (2.26) and equation (2.27) reduce to the algebraic equations. Moreover, Takahashi (2008) has calculated the exact solution up to third order in general dark energy model, and has concluded that the difference between the next-to-leading order power spectrum from exact kernels and that from Einstein de-Sitter kernels is extremely small. It is at most sub-percent level at $z = 0$, and decreases as redshift increases to $\sim 10^{-4}$ at $z = 3$. As we are mostly interested in the high redshift ($z > 1$), we can safely ignore such a small difference in the kernels.

In Einstein de-Sitter Universe, equation (2.26) and equation (2.27) become

$$nF_n(\mathbf{k}) - G_n(\mathbf{k}) = \sum_{m=1}^{n-1} \frac{\mathbf{k} \cdot \mathbf{k}_1}{k_1^2} G_m(\mathbf{k}_1) F_{n-m}(\mathbf{k}_2), \quad (2.28)$$

$$(2n+1)G_n(\mathbf{k}) - 3F_n(\mathbf{k}) = \sum_{m=1}^{n-1} \frac{k^2(\mathbf{k}_1 \cdot \mathbf{k}_2)}{k_1^2 k_2^2} G_m(\mathbf{k}_1) G_{n-m}(\mathbf{k}_2). \quad (2.29)$$

By solving the algebraic equations, we find the recursion relations

$$\begin{aligned} F_n(\mathbf{q}_1, \dots, \mathbf{q}_n) &= \sum_{m=1}^{n-1} \frac{G_m(\mathbf{q}_1, \dots, \mathbf{q}_m)}{(2n+3)(n-1)} \left\{ (2n+1) \frac{\mathbf{k} \cdot \mathbf{k}_1}{k_1^2} F_{n-m}(\mathbf{q}_{m+1}, \dots, \mathbf{q}_n) \right. \\ &\quad \left. + \frac{k^2(\mathbf{k}_1 \cdot \mathbf{k}_2)}{k_1^2 k_2^2} G_{n-m}(\mathbf{q}_{m+1}, \dots, \mathbf{q}_n) \right\} \end{aligned} \quad (2.30)$$

$$\begin{aligned} G_n(\mathbf{q}_1, \dots, \mathbf{q}_n) &= \sum_{m=1}^{n-1} \frac{G_m(\mathbf{q}_1, \dots, \mathbf{q}_m)}{(2n+3)(n-1)} \left\{ 3 \frac{\mathbf{k} \cdot \mathbf{k}_1}{k_1^2} F_{n-m}(\mathbf{q}_{m+1}, \dots, \mathbf{q}_n) \right. \\ &\quad \left. + n \frac{k^2(\mathbf{k}_1 \cdot \mathbf{k}_2)}{k_1^2 k_2^2} G_{n-m}(\mathbf{q}_{m+1}, \dots, \mathbf{q}_n) \right\}. \end{aligned} \quad (2.31)$$

With the recursion relations and the kernels for linear theory, $F_1 = G_1 = 1$, we can calculate $F_n^{(s)}, G_n^{(s)}$ for all order. For example, the second order kernels are

$$F_2^{(s)}(\mathbf{k}_1, \mathbf{k}_2) = \frac{5}{7} + \frac{2}{7} \frac{(\mathbf{k}_1 \cdot \mathbf{k}_2)^2}{k_1^2 k_2^2} + \frac{\mathbf{k}_1 \cdot \mathbf{k}_2}{2} \left(\frac{1}{k_1^2} + \frac{1}{k_2^2} \right), \quad (2.32)$$

$$G_2^{(s)}(\mathbf{k}_1, \mathbf{k}_2) = \frac{3}{7} + \frac{4}{7} \frac{(\mathbf{k}_1 \cdot \mathbf{k}_2)^2}{k_1^2 k_2^2} + \frac{\mathbf{k}_1 \cdot \mathbf{k}_2}{2} \left(\frac{1}{k_1^2} + \frac{1}{k_2^2} \right), \quad (2.33)$$

and third order kernels can be calculated from second order kernels as⁷

$$\begin{aligned} F_3^{(s)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = & \frac{2k^2}{54} \left[\frac{\mathbf{k}_1 \cdot \mathbf{k}_{23}}{k_1^2 k_{23}^2} G_2^{(s)}(\mathbf{k}_2, \mathbf{k}_3) + (2 \text{ cyclic}) \right] \\ & + \frac{7}{54} \mathbf{k} \cdot \left[\frac{\mathbf{k}_{12}}{k_{12}^2} G_2^{(s)}(\mathbf{k}_1, \mathbf{k}_2) + (2 \text{ cyclic}) \right] \\ & + \frac{7}{54} \mathbf{k} \cdot \left[\frac{\mathbf{k}_1}{k_1^2} F_2^{(s)}(\mathbf{k}_2, \mathbf{k}_3) + (2 \text{ cyclic}) \right], \end{aligned} \quad (2.34)$$

and

$$\begin{aligned} G_3^{(s)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = & \frac{k^2}{9} \left[\frac{\mathbf{k}_1 \cdot \mathbf{k}_{23}}{k_1^2 k_{23}^2} G_2^{(s)}(\mathbf{k}_2, \mathbf{k}_3) + (2 \text{ cyclic}) \right] \\ & + \frac{1}{18} \mathbf{k} \cdot \left[\frac{\mathbf{k}_{12}}{k_{12}^2} G_2^{(s)}(\mathbf{k}_1, \mathbf{k}_2) + (2 \text{ cyclic}) \right] \\ & + \frac{1}{18} \mathbf{k} \cdot \left[\frac{\mathbf{k}_1}{k_1^2} F_2^{(s)}(\mathbf{k}_2, \mathbf{k}_3) + (2 \text{ cyclic}) \right]. \end{aligned} \quad (2.35)$$

In summary, the solution of Eulerian Perturbation Theory consists of the perturbative expansions (Eq. (2.22) and Eq. (2.23)) with kernels calculated by recursion relations (Eq. (2.30) and Eq. (2.31)) followed by symmetrization (Eq. (2.24) and Eq. (2.25)). By using this solution, we can describe the non-linear growth of density field and velocity field of cosmic matter in the quasi-nonlinear regime.

2.2 Statistics of the cosmological density field: Gaussian vs. non-Gaussian

In the previous section, we calculate the solution for the non-linear evolution of the density and velocity fields in terms of the linear density field. As the direct observable of the

⁷For the compactness of the equation, we adopt the short-hand notation of

$$\mathbf{k}_{ijk\dots} \equiv \mathbf{k}_i + \mathbf{k}_j + \mathbf{k} + \dots$$

throughout this dissertation.

cosmological observation is the statistical correlation function such as power spectrum and bispectrum, in order to compare the non-linear solution to observation, we have to calculate the statistical correlation of density and velocity field.

Inflationary theories predict that the primordial curvature perturbation obeys a *nearly* Gaussian statistics; thus, the linear density field $\delta_1(\mathbf{k})$, which evolves linearly from the curvature perturbation [Eq. (2.18)], also obeys the same statistics. Let us first consider the Gaussian statistics.

In order to describe the statistics of a field, we have to introduce a *probability functional* $\mathcal{P}[\delta_1(\mathbf{x})]$, which describes the probability of having a configuration of density field δ_1 whose value is $\delta_1(\mathbf{x})$ at a point \mathbf{x} . Note that statistical homogeneity dictates $\mathcal{P}[\delta_1(\mathbf{x})]$ to be independent of position \mathbf{x} . For Gaussian case, it is given by (Gabrielli et al., 2005)

$$\mathcal{P}[\delta_1(\mathbf{x})] = \frac{1}{Z} \exp \left[-\frac{1}{2} \int d^3y \int d^3z \delta_1(\mathbf{y}) K(\mathbf{y}, \mathbf{z}) \delta_1(\mathbf{z}) \right], \quad (2.36)$$

where $K(\mathbf{y}, \mathbf{z})$ is corresponding to the inverse of the covariance matrix C^{-1} in the usual Gaussian statistics of discrete variables, and

$$Z = \int [\mathcal{D}\delta_1] \exp \left[-\frac{1}{2} \int d^3y \int d^3z \delta_1(\mathbf{y}) K(\mathbf{y}, \mathbf{z}) \delta_1(\mathbf{z}) \right] \quad (2.37)$$

is the normalization constant, and $[\mathcal{D}\delta_1]$ is the integration measure in the Hilbert space.

By using the probability functional, we calculate the n -point correlation function of $\delta_1(\mathbf{x})$ by taking the expectation value as:

$$\langle \delta_1(\mathbf{x}_1) \cdots \delta_1(\mathbf{x}_n) \rangle = \int [\mathcal{D}\delta_1(\mathbf{x})] \delta_1(\mathbf{x}_1) \cdots \delta_1(\mathbf{x}_n) \mathcal{P}[\delta_1(\mathbf{x})]. \quad (2.38)$$

A usual technique of calculating the expectation value is by introducing the *generating functional*

$$Z[J] \equiv \left\langle e^{i \int d^3x J(\mathbf{x}) \delta_1(\mathbf{x})} \right\rangle, \quad (2.39)$$

and taking successive functional derivatives with respect to $J(\mathbf{x}_i)$. Finally, we get the correlation function by setting $J = 0$. That is, n -point correlation function is

$$\langle \delta_1(\mathbf{x}_1) \cdots \delta_1(\mathbf{x}_n) \rangle = \frac{1}{i^n} \frac{\delta^n Z[J]}{\delta J(\mathbf{x}_1) \cdots \delta J(\mathbf{x}_n)} \Big|_{J=0}. \quad (2.40)$$

Furthermore, we can also calculate the n -point *connected* correlation function by

$$\langle \delta_1(\mathbf{x}_1) \cdots \delta_1(\mathbf{x}_n) \rangle_c = \frac{1}{i^n} \frac{\delta^n \ln Z[J]}{\delta J(\mathbf{x}_1) \cdots \delta J(\mathbf{x}_n)} \Big|_{J=0}. \quad (2.41)$$

For Gaussian case, we can re-write the generating functional in a closed form:

$$Z[J] = \frac{1}{Z} \int [\mathcal{D}f] \exp \left[-\frac{1}{2} \int d^3y \int d^3z f(\mathbf{y}) K(\mathbf{y}, \mathbf{z}) f(\mathbf{z}) + i \int d^3x J(\mathbf{x}) f(\mathbf{x}) \right]. \quad (2.42)$$

Then, we make the coordinate transformation of

$$\delta'_1(\mathbf{x}) = \delta_1(\mathbf{x}) - i \int d^3y K^{-1}(\mathbf{x}, \mathbf{y}) J(\mathbf{y}),$$

where $K^{-1}(\mathbf{x}, \mathbf{y})$ is the functional inverse of $K(\mathbf{x}, \mathbf{y})$, which is defined by

$$\int d^3z K(\mathbf{x}, \mathbf{z}) K^{-1}(\mathbf{z}, \mathbf{y}) = \delta^D(\mathbf{x} - \mathbf{y}).$$

This coordinate transformation does not change the integral measure, as is simple translation in the functional space. Under the coordinate transformation, exponent in equation (2.42) becomes

$$-\frac{1}{2} \int d^3y \int d^3z \delta'_1(\mathbf{y}) K(\mathbf{y}, \mathbf{z}) \delta'_1(\mathbf{z}) - \frac{1}{2} \int d^3x \int d^3y J(\mathbf{x}) K^{-1}(\mathbf{x}, \mathbf{y}) J(\mathbf{y}), \quad (2.43)$$

where the functional integration of the first term is unity; thus, generating functional for Gaussian case becomes

$$Z[J] = \exp \left[-\frac{1}{2} \int d^3x \int d^3y J(\mathbf{x}) K^{-1}(\mathbf{x}, \mathbf{y}) J(\mathbf{y}) \right]. \quad (2.44)$$

Therefore, by using equation (2.41), we calculate the n -point correlation function. By taking the functional derivative twice, we find that

$$\begin{aligned} \frac{\delta Z[J]}{\delta J(\mathbf{x}_1)} &= -\frac{1}{2} \left[\int d^3x \int d^3y \delta^D(\mathbf{x} - \mathbf{x}_1) K^{-1}(\mathbf{x}, \mathbf{y}) J(\mathbf{y}) \right. \\ &\quad \left. + \int d^3x \int d^3y J(\mathbf{x}) K^{-1}(\mathbf{x}, \mathbf{y}) \delta^D(\mathbf{y} - \mathbf{x}_1) \right] Z[J] \\ &= - \left[\int d^3x K^{-1}(\mathbf{x}, \mathbf{x}_1) J(\mathbf{x}_1) \right] Z[J], \end{aligned}$$

and

$$\frac{\delta^2 Z[J]}{\delta J(\mathbf{x}_1) \delta J(\mathbf{x}_2)} = -K^{-1}(\mathbf{x}, \mathbf{x}_1) Z[J].$$

By carefully observing this procedure, we find that any odd-number n -point correlation function has to vanish when setting $J = 0$, and any even-number n -point correlation function is given only by $K^{-1}(\mathbf{x}, \mathbf{y})$. Especially, from the calculation above, we can read that the two point correlation function is given by

$$\xi(\mathbf{x}_1, \mathbf{x}_2) \equiv \langle \delta_1(\mathbf{x}_1) \delta_1(\mathbf{x}_2) \rangle = K^{-1}(\mathbf{x}_1, \mathbf{x}_2). \quad (2.45)$$

This result is so called *Wick's theorem*, which states that “ the average of a product of an even number of δ_1 s is the sum over all ways of pairing δ_1 s with each other of a product of the average values of the pairs:

$$\langle \delta_1(\mathbf{x}_1)\delta_1(\mathbf{x}_2)\cdots \rangle = \sum_{\text{pairings}} \prod_{\text{pairs}} \langle \delta_1\delta_1 \rangle, \quad (2.46)$$

with the sum over pairings not distinguishing those which interchange coordinates in a pair, or which merely interchange pairs” (Weinberg, 2008). What about the connected correlation function? From equation (2.41), it is clear that, for Gaussian case, all higher order ($n > 2$) connected correlation functions vanish:

$$\langle \delta_1(\mathbf{x}_1)\cdots\delta_1(\mathbf{x}_n) \rangle_c = -\frac{1}{2i^n} \frac{\delta^n}{\delta J(\mathbf{x}_1)\cdots\delta J(\mathbf{x}_n)} \int d^3x \int d^3y J(\mathbf{x}) K^{-1}(\mathbf{x}, \mathbf{y}) J(\mathbf{y}).$$

From equation (2.45), $K(\mathbf{x}, \mathbf{y})$ is related to the correlation function by

$$\int d^3z \xi(\mathbf{x}, \mathbf{z}) K(\mathbf{z}, \mathbf{y}) = \delta^D(\mathbf{x} - \mathbf{y}). \quad (2.47)$$

Let us further investigate on the kernel, $K(\mathbf{x}, \mathbf{y})$. Due to the statistical homogeneity and isotropy of universe, $\xi(\mathbf{x}, \mathbf{z})$ and $K(\mathbf{x}, \mathbf{y})$ have to depend only on the separation, i.e. $\xi(\mathbf{x}, \mathbf{z}) = \xi(\mathbf{x} - \mathbf{z})$. Therefore,

$$\int d^3z \xi_f(\mathbf{x} - \mathbf{z}) K(\mathbf{z} - \mathbf{y}) = \delta^D(\mathbf{x} - \mathbf{y}). \quad (2.48)$$

Then, Fourier transform of the equation above leads

$$K(\mathbf{q}) = \frac{1}{P_L(q)}, \quad (2.49)$$

where $P_L(q)$ is the linear power spectrum, which is related to the two-point correlation function by

$$\xi(r) = \int \frac{d^3q}{(2\pi)^3} P_L(q) e^{i\mathbf{q}\cdot\mathbf{r}}. \quad (2.50)$$

That is, $K(\mathbf{x} - \mathbf{y})$ is Fourier transform of the inverse of linear power spectrum $P_L(q)$. One can also show that directly by using Fourier space representation of the linear density field $\delta_1(\mathbf{k})$. By Fourier transforming the exponent of equation (2.36), we find

$$\begin{aligned} & -\frac{1}{2} \int d^3y \int d^3z \delta_1(\mathbf{y}) K(\mathbf{y} - \mathbf{z}) \delta_1(\mathbf{z}) \\ &= -\frac{1}{2} \int d^3q_1 \int d^3q_2 \delta_1(\mathbf{q}_1) \left[\frac{K(\mathbf{q}_1)}{(2\pi)^3} \delta^D(\mathbf{q}_1 + \mathbf{q}_2) \right] \delta_1(\mathbf{q}_2). \end{aligned} \quad (2.51)$$

As the structure of the probability functional is the same as that for the real space, we can simply relate $K(\mathbf{q})$ to the power spectrum by using the analogy. Because the functional inversion of the Fourier space kernel is

$$\left[\frac{K(\mathbf{q}_1)\delta^D(\mathbf{q}_1 + \mathbf{q}_2)}{(2\pi)^3} \right]^{-1} = \frac{(2\pi)^3}{K(\mathbf{q}_1)}\delta^D(\mathbf{q}_1 + \mathbf{q}_2), \quad (2.52)$$

the power spectrum is given by

$$\langle \delta_1(\mathbf{q}_1)\delta_1(\mathbf{q}_2) \rangle \equiv (2\pi)^3 P(q_1)\delta^D(\mathbf{q}_1 + \mathbf{q}_2) = \frac{(2\pi)^3}{K(\mathbf{q}_1)}\delta^D(\mathbf{q}_1 + \mathbf{q}_2). \quad (2.53)$$

By using the second equality, we reproduce equation (2.49). Politzer & Wise (1984) extended this method to calculate the n -point correlation function of peaks.

As we have shown here, when the linear density field follows the Gaussian statistics, all higher order statistical properties are solely determined by the two point correlation function, or power spectrum. For non-Gaussian case, however, the connected n -point correlation functions do not vanish, and we have to know the probability functional or generating functional in order to calculate the statistical quantities. Conversely, if we know the n -point correlation functions to infinite order, then, we can calculate the generating functional by inverting equation (2.41):

$$\ln Z[J] = \sum_{n=2}^{\infty} \frac{i^n}{n!} \int d^3y_1 \cdots \int d^3y_n \langle \delta_1(\mathbf{y}_1) \cdots \delta_1(\mathbf{y}_n) \rangle_c J(\mathbf{y}_1) \cdots J(\mathbf{y}_n), \quad (2.54)$$

thus fully specify the statistics of the linear density field. Matarrese et al. (1986) derives the n -point correlation function of peaks by using this method (MLB formula), and we shall calculate the galaxy bispectrum in the presence of primordial non-Gaussianity in Chapter 5 based on this formula. The full derivation of MLB formula is also shown in Appendix K.

2.3 Next-to-leading order power spectrum from the perturbative solution: the theoretical template

In the following sections, we shall calculate the power spectrum of density contrast of matter in real space (Section 2.4), matter in redshift space (Section 2.6), galaxies in real space (Section 2.5), and galaxies in redshift space (Section 2.7). For all cases, we shall calculate corresponding density contrasts as a function of the matter density contrast $\delta_{\mathbf{k}}(\tau)$ and velocity gradient $\theta_{\mathbf{k}}(\tau)$ for which we know the perturbation theory solution.

Therefore, in most general form, we can also write a quantity $X(\mathbf{k}, \tau)$, of which we want to calculate power spectrum, as a sum of the perturbative series

$$\begin{aligned} X(\mathbf{k}, \tau) &= X^{(1)}(\mathbf{k}, \tau) + X^{(2)}(\mathbf{k}, \tau) + X^{(3)}(\mathbf{k}, \tau) + \dots \\ &= \sum_{n=1}^{\infty} \int \frac{d^3 q_1}{(2\pi)^3} \dots \int \frac{d^3 q_{n-1}}{(2\pi)^3} \int d^3 q_n \delta^D(\mathbf{k} - \sum_{i=1}^n \mathbf{q}_i) \\ &\quad \times K_n^{(s)}(\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n) \delta_1(\mathbf{q}_1, \tau) \dots \delta_1(\mathbf{q}_n, \tau), \end{aligned} \quad (2.55)$$

with a general symmetrized kernel $K_n^{(s)}(\mathbf{q}_1, \dots, \mathbf{q}_n)$, which is given by a combination of $F_n^{(s)}$ and $G_n^{(s)}$. Here, $X^{(n)}(\mathbf{k})$ denote that the quantity is n -th order in linear density contrast, $\delta_1(\mathbf{k}, \tau)$. For a quantity which can be expanded as equation (2.55), we can calculate next-to-leading order (called *one-loop*) power spectrum and bispectrum as a function of linear order (called *tree-level*) quantities. In this section, we present these general formulas which we shall use in the rest of this chapter.

As the power spectrum, $P_X(\mathbf{k}, \tau)$, is a quadratic quantity of $X(\mathbf{k}, \tau)$ in Fourier space,

$$\langle X(\mathbf{k})X(\mathbf{k}') \rangle = (2\pi)^3 P_X(\mathbf{k}) \delta^D(\mathbf{k} + \mathbf{k}'), \quad (2.56)$$

for the expansion in equation (2.55), the left hand side of equation (2.56) becomes

$$\begin{aligned} &\langle X(\mathbf{k}, \tau)X(\mathbf{k}', \tau) \rangle \\ &= \langle X^{(1)}(\mathbf{k}, \tau)X^{(1)}(\mathbf{k}', \tau) \rangle + 2\langle \text{Re} \left[X^{(1)}(\mathbf{k}, \tau)X^{(2)}(\mathbf{k}', \tau) \right] \rangle \\ &\quad + \langle X^{(2)}(\mathbf{k}, \tau)X^{(2)}(\mathbf{k}', \tau) \rangle + 2\langle \text{Re} \left[X^{(1)}(\mathbf{k}, \tau)X^{(3)}(\mathbf{k}', \tau) \right] \rangle + \dots \end{aligned} \quad (2.57)$$

Here, we use the Hermitianity (Reality) condition,

$$X^{(i)}(\mathbf{k}, \tau) = \left[X^{(i)}(-\mathbf{k}, \tau) \right]^* = \left[X^{(i)}(\mathbf{k}', \tau) \right]^*,$$

and Re takes the real part of a complex number. In equation (2.57), we explicitly write down the terms up to order $\mathcal{O}(\delta_1^4)$.

2.3.1 One-loop Power spectrum with Gaussian linear density field

When linear density contrast, δ_1 , is Gaussian, odd power of δ_1 vanishes when taking an ensemble average. Therefore, the next-to-leading order nonlinear power spectrum consists of $\langle X^{(1)}(\mathbf{k}, \tau)X^{(1)}(\mathbf{k}', \tau) \rangle$, $\langle X^{(2)}(\mathbf{k}, \tau)X^{(2)}(\mathbf{k}', \tau) \rangle$ and $\langle X^{(1)}(\mathbf{k}, \tau)X^{(3)}(\mathbf{k}', \tau) \rangle$ which is often

denoted as $P_{X,11}(\mathbf{k}, \tau)$, $P_{X,22}(\mathbf{k}, \tau)$ and $P_{X,13}(\mathbf{k}, \tau)$, respectively. We calculate the next-to-leading order power spectrum by using Wick's theorem as⁸

$$\begin{aligned}
& P_X(\mathbf{k}, \tau) \\
& \equiv P_{X,11}(\mathbf{k}, \tau) + P_{X,22}(\mathbf{k}, \tau) + 2P_{X,13}(\mathbf{k}, \tau) \\
& = \left[K_1^{(s)}(\mathbf{k}) \right]^2 P_L(k, \tau) + 2 \int \frac{d^3 q}{(2\pi)^3} P_L(q, \tau) P_L(|\mathbf{k} - \mathbf{q}|, \tau) \left[K_2^{(s)}(\mathbf{q}, \mathbf{k} - \mathbf{q}) \right]^2 \\
& \quad + 6 K_1^{(s)}(\mathbf{k}) P_L(k, \tau) \int \frac{d^3 q}{(2\pi)^3} P_L(q, \tau) K_3^{(s)}(\mathbf{q}, -\mathbf{q}, \mathbf{k}), \tag{2.58}
\end{aligned}$$

where $P_L(k)$ is the linear power spectrum.

2.3.2 One-loop Power spectrum with non-Gaussian linear density field

When linear density contrast follows non-Gaussian distribution, the second term in equation (2.57), $\langle X^{(1)}(\mathbf{k}) X^{(2)}(\mathbf{k}') \rangle$, is non-zero, and becomes the leading order non-Gaussian correction to the power spectrum,

$$\begin{aligned}
\Delta P_{X,NG}(\mathbf{k}, \tau) &= 2P_{12}(\mathbf{k}, \tau) \\
&= 2K_1^{(s)}(\mathbf{k}) \int \frac{d^3 q}{(2\pi)^3} K_2^{(s)}(\mathbf{q}, \mathbf{k} - \mathbf{q}) B_L(-\mathbf{k}, \mathbf{q}, \mathbf{k} - \mathbf{q}, \tau). \tag{2.59}
\end{aligned}$$

Here, B_L is the bispectrum of linear density field. There also are the non-Gaussian correction term coming from $P_{X,22}(\mathbf{k}, \tau)$ and $P_{X,13}(\mathbf{k}, \tau)$ proportional to the linear trispectrum, $T_L(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4)$ ⁹:

$$\begin{aligned}
\Delta P_{X,22}(\mathbf{k}, \tau) &= \int \frac{d^3 q_1}{(2\pi)^3} \int \frac{d^3 q_2}{(2\pi)^3} K_2^{(s)}(\mathbf{q}_1, \mathbf{k} - \mathbf{q}_1) K_2^{(s)}(\mathbf{q}_2, -\mathbf{k} - \mathbf{q}_2) \\
&\quad \times T_L(\mathbf{q}_1, \mathbf{k} - \mathbf{q}_1, \mathbf{q}_2, -\mathbf{k} - \mathbf{q}_2, \tau) \tag{2.60}
\end{aligned}$$

$$\begin{aligned}
\Delta P_{X,13}(\mathbf{k}, \tau) &= K_1^{(s)}(\mathbf{k}) \int \frac{d^3 q_1}{(2\pi)^3} \int \frac{d^3 q_2}{(2\pi)^3} K_3^{(s)}(\mathbf{q}_1, \mathbf{q}_2, \mathbf{k} - \mathbf{q}_1 - \mathbf{q}_2) \\
&\quad \times T_L(-\mathbf{k}, \mathbf{q}_1, \mathbf{q}_2, \mathbf{k} - \mathbf{q}_1 - \mathbf{q}_2, \tau) \tag{2.61}
\end{aligned}$$

Therefore, in order to calculate the effect of primordial non-Gaussianity fully in one-loop level of perturbation theory, we need to know both linear bispectrum and the linear trispectrum. However, linear trispectrum ($T_L \sim \mathcal{M}^4 P_\phi^3$) is parametrically smaller

⁸For the derivation of $P_{X,22}$ and $P_{X,13}$, see Appendix B.1.

⁹For derivation, see Appendix B.2

than linear bispectrum ($B_L \sim \mathcal{M}^3 P_\phi^1$) in the quasi-nonlinear regime, as the primordial curvature power spectrum P_ϕ sharply decreases as k increases, $P_\phi(k) \sim k^{n_s-4}$. In other words, linear trispectrum arises from the two-loop contribution of the primordial curvature perturbation, we shall neglect $\Delta P_{X,22}$ and $\Delta P_{X,13}$ correction terms (Taruya et al., 2008). However, note that the trispectrum term may be important on the very large scales. Indeed $\Delta P_{X,22}$ term include a term proportional to $f_{\text{NL}}^2/(k^4 T^2(k))$ which sharply peaks on large scales (Desjacques & Seljak, 2010). In the rest of the chapter, we shall use equation (2.59) to calculate the non-Gaussian correction to the one-loop power spectrum.

2.3.2.1 Local primordial non-Gaussianity

Although we can apply equation (2.59) to any kind of non-Gaussianities, we shall focus on the local type non-Gaussianity. The local type non-Gaussianity is particularly interesting because of the consistency relation (Maldacena, 2003; Acquaviva et al., 2003; Creminelli & Zaldarriaga, 2004), which states that the coefficient of the squeezed bispectrum [Eq. (2.64)], f_{NL} , has to satisfy

$$f_{\text{NL}} = \frac{5}{12}(1 - n_s) \quad (2.62)$$

for *any* kind of single field inflation model. Here, n_s is the spectral tilt of the primordial power spectrum [Eq. (2.19)], whose current best estimation is $n_s = 0.963 \pm 0.012$ (68% CL) from WMAP7 (Komatsu et al., 2010). Therefore, any significant detection of f_{NL} above the value dictated by the consistency relation $f_{\text{NL}} \sim 0.017$ will rule out the single field inflationary models. The current limit from the WMAP 7-year data is $f_{\text{NL}} = 32 \pm 21$ (68% CL) (Komatsu et al., 2010), and from the SDSS is $f_{\text{NL}} = 31^{+16}_{-27}$ (68% CL) (Slosar et al., 2008). We shall further study the effect of local primordial non-Gaussianity on the galaxy bispectrum and the galaxy-galaxy, galaxy-CMB weak gravitational lensing in Chapter 5 and Chapter 6, respectively.

Local type primordial non-Gaussianity is defined by the Bardeen's potential $\Phi(\mathbf{x})$ [Eq. (2.14)] in real space (Salopek & Bond, 1990; Gangui et al., 1994; Verde et al., 2000; Komatsu & Spergel, 2001):

$$\Phi(\mathbf{x}) = \phi(\mathbf{x}) + f_{\text{NL}} [\phi^2(\mathbf{x}) - \langle \phi^2 \rangle] + \mathcal{O}(\phi^3), \quad (2.63)$$

where $\phi(\mathbf{x})$ is a Gaussian random field. The primordial bispectrum generated by the local

non-Gaussianity¹⁰ is

$$B_\Phi(k_1, k_2, k_3) = 2f_{\text{NL}} [P_\phi(k_1)P_\phi(k_2) + (2 \text{ cyclic})]. \quad (2.64)$$

We can find the linear matter bispectrum by linearly evolving the primordial bispectrum [Eq. (2.64)] to present using equation (2.18):

$$B_L(k_1, k_2, k_3; z) = 2f_{\text{NL}} D^3(z) \mathcal{M}(k_1) \mathcal{M}(k_2) \mathcal{M}(k_3) [P_\phi(k_1)P_\phi(k_2) + (2 \text{ cyclic})]. \quad (2.65)$$

Using the linear bispectrum above, we calculate the non-Gaussian correction to the nonlinear power spectrum as

$$\begin{aligned} \Delta P_{X, \text{nG}}(k, z) = & 4f_{\text{NL}} D^3(z) K_1^{(s)}(\mathbf{k}) \mathcal{M}(k) \int \frac{d^3 q}{(2\pi)^3} \mathcal{M}(q) \mathcal{M}(|\mathbf{k} - \mathbf{q}|) \\ & \times K_2^{(s)}(\mathbf{q}, \mathbf{k} - \mathbf{q}) P_\phi(q) [2P_\phi(k) + P_\phi(|\mathbf{k} - \mathbf{q}|)]. \end{aligned} \quad (2.66)$$

Note that we ignore the nonlinearity in the linear power spectrum generated by equation (2.63), and use a linear approximation as

$$P_L(k, z) \simeq \mathcal{M}^2(k) D^2(z) P_\phi(k). \quad (2.67)$$

This approximation is valid up to slight rescaling of amplitude and slope of the primordial curvature power spectrum. For more discussion, see, Section II of McDonald (2008).

2.4 Nonlinear matter power spectrum in real space

For a nonlinear matter power spectrum in real space, we can simply use the perturbative solution for $\delta_{\mathbf{k}}(\tau)$ in equation (2.22). That is,

$$\begin{aligned} K_1^{(s)}(\mathbf{q}_1) &= 1 \\ K_2^{(s)}(\mathbf{q}_1, \mathbf{q}_2) &= F_2^{(s)}(\mathbf{q}_1, \mathbf{q}_2) \\ K_3^{(s)}(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) &= F_3^{(s)}(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3), \end{aligned}$$

where $F_2^{(s)}$ and $F_3^{(s)}$ are presented in equation (2.32), and equation (2.34), respectively.

¹⁰For derivation, see Appendix C.

2.4.1 Gaussian case

By substituting this kernels and equation (2.58), we calculate the matter power spectrum in real space as

$$P_m(k, z) = D^2(z)P_L(k) + D^4(z)[P_{m,22}(k) + 2P_{m,13}(k)], \quad (2.68)$$

where $D(z)$ is the linear growth factor and

$$P_{m,22}(k) = \frac{1}{98} \frac{k^3}{(2\pi)^2} \int dr P_L(kr) \int_{-1}^1 dx P_L(k\sqrt{1+r^2-2rx}) \left[\frac{7x+3r-10rx^2}{1+r^2-2rx} \right]^2 \quad (2.69)$$

$$P_{m,13}(k) = \frac{1}{504} \frac{k^3}{(2\pi)^2} P_L(k) \int dr P_L(kr) \times \left[\frac{12}{r^2} - 158 + 100r^2 - 42r^4 + \frac{3}{r^3}(r^2-1)^3(7r^2+2) \ln \left(\frac{r+1}{|r-1|} \right) \right]. \quad (2.70)$$

Here, $P_L(k)$ is calculated at present where linear growth factor is normalized to be unity. This form is practical useful as, for given linear power spectrum, we only need to calculate the integration once for a give redshift. Then, the non-linear matter power spectrum for different redshifts can be calculated by simple rescaling of P_{11} , P_{22} and P_{13} with corresponding powers of linear growth factor.

We show that, in the quasi-nonlinear regime at high redshift, this *analytic* expression accurately models the nonlinear evolution of the matter power spectrum from a series of N-body simulations we run in Chapter 3. We also verify the result against the matter power spectrum from Millennium Simulation (Springel et al., 2005) in Section 4.2.

2.4.2 Non-Gaussianity case

We also calculate the leading order non-Gaussian term due to the local type primordial non-Gaussianity from equation (2.66).

$$\begin{aligned} \Delta P_{m,nG}(k, z) = & \frac{3}{7} f_{NL} H_0^2 \Omega_m D^3(z) \frac{k}{(2\pi)^2} \int \frac{dr}{r} \frac{P_L(kr)}{T(kr)} \int_{-1}^1 dx \left(\frac{7x+3r-10rx^2}{1+r^2-2rx} \right) \\ & \times \left[2P_L(k)(1+r^2-2rx) \frac{T(k\sqrt{1+r^2-2rx})}{T(k)} \right. \\ & \left. + \frac{P_L(k\sqrt{1+r^2-2rx})}{1+r^2-2rx} \frac{T(k)}{T(k\sqrt{1+r^2-2rx})} \right] \end{aligned} \quad (2.71)$$

This equation is first derived from Taruya et al. (2008), and they find that non-Gaussianity signal in matter power spectrum is so tiny that gigantic space based survey with survey volume of $100 [\text{Gpc}^3/h^3]$ only detect with large uncertainty ($\Delta f_{NL} \simeq 300$).

2.5 Nonlinear galaxy power spectrum in real space

In galaxy surveys, what we observe are galaxies, not a matter fluctuation. Since galaxies are the biased tracers of the underlying matter fluctuation, we have to understand how galaxy distribution and the matter fluctuation are related. This relation is known to be very complicated, because we have to understand the complex galaxy formation processes as well as the dark matter halo formation processes for given matter fluctuation in order to calculate the relation from the first principle. Both of which are the subject of the forefront research and need to be investigated further.

We simplify the situation by assuming that the galaxy formation and halo formation are local processes. This assumption is valid on large enough scale, which may include the *quasi nonlinear* scale where PT models the nonlinear evolution very well. Then, the galaxy over/under density at a given position depends only on the matter fluctuation at the same position. Therefore, the galaxy density contrast $\delta_g(\mathbf{x})$ can be Taylor-expanded with respect to the smoothed matter density contrast

$$\delta_R(\mathbf{x}) = \int d^3y \delta(\mathbf{y}) \mathcal{W}_R(\mathbf{x} - \mathbf{y})$$

as

$$\delta_g(\mathbf{x}) = \epsilon + c_1 \delta_R(\mathbf{x}) + \frac{c_2}{2} [\delta_R^2(\mathbf{x}) - \langle \delta_R^2 \rangle] + \frac{c_3}{6} \delta_R^3(\mathbf{x}) + \dots, \quad (2.72)$$

where $\langle \delta^2 \rangle$ is subtracted in order to ensure $\langle \delta_g \rangle = 0$ (McDonald, 2006). Here, $\mathcal{W}_R(r)$ is the smoothing (filtering) function, and $\bar{W}_R(k)$ is its Fourier transform¹¹. We also introduce the stochastic parameter ϵ which quantifies the “stochasticity” of galaxy bias, i.e. the relation between $\delta_g(\mathbf{x})$ and $\delta_R(\mathbf{x})$ is not completely deterministic, but contains some noise with zero mean, $\langle \epsilon \rangle = 0$ (e.g., Yoshikawa et al. (2001), and reference therein). We further assume that the stochasticity is a white noise, and is uncorrelated with the density fluctuations i.e., $\langle \epsilon^2(\mathbf{k}) \rangle \equiv \epsilon_0^2$, $\langle \epsilon \delta_R \rangle = 0$. The coefficients of expansion, c_n ’s, encode the detailed formation history of galaxies, and may vary for different morphological types, colors, flux limits, etc.

By using a convolution theorem, we calculate the Fourier transform of the local bias expansion of equation (2.72)

$$\begin{aligned} \delta_g(\mathbf{k}) = & \epsilon(\mathbf{k}) + c_1 \delta_R(\mathbf{k}) + \frac{c_2}{2} \int \frac{d^3q_1}{(2\pi)^3} \int d^3q_2 \delta_R(\mathbf{q}_1) \delta_R(\mathbf{q}_2) \delta_D(\mathbf{k} - \mathbf{q}_{12}) \\ & + \frac{c_3}{6} \int \frac{d^3q_1}{(2\pi)^3} \int \frac{d^3q_2}{(2\pi)^3} \int d^3q_3 \delta_R(\mathbf{q}_1) \delta_R(\mathbf{q}_2) \delta_R(\mathbf{q}_3) \delta_D(\mathbf{k} - \mathbf{q}_{123}), \end{aligned} \quad (2.73)$$

¹¹For the notational simplicity, we shall drop the tilde, but it should be clear from the argument whether the filtering function is defined in real space or Fourier space.

in terms of the smoothed non-linear density field $\delta_R(\mathbf{k})$:

$$\delta_R(\mathbf{k}) \equiv \mathcal{W}_R(k) \left[\delta^{(1)}(\mathbf{k}) + \delta^{(2)}(\mathbf{k}) + \delta^{(3)}(\mathbf{k}) + \dots \right]. \quad (2.74)$$

Here, $\delta^{(n)}(\mathbf{k})$ denotes the n -th order perturbation theory solution in equation (2.22). We find the kernel for the real space galaxy density contrast by substituting equation (2.74) into equation (2.73).

$$\begin{aligned} K_1^{(s)}(\mathbf{q}_1) &= c_1 \mathcal{W}_R(q_1) \\ K_2^{(s)}(\mathbf{q}_1, \mathbf{q}_2) &= \frac{c_2}{2} \mathcal{W}_R(q_1) \mathcal{W}_R(q_2) + c_1 F_2^{(s)}(\mathbf{q}_1, \mathbf{q}_2) \mathcal{W}_R(q_{12}) \\ K_3^{(s)}(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) &= \frac{c_3}{6} \mathcal{W}_R(q_1) \mathcal{W}_R(q_2) \mathcal{W}_R(q_3) + c_1 F_3^{(s)}(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) \mathcal{W}_R(q_{123}) \\ &\quad + \frac{c_2}{3} \left[F_2^{(s)}(\mathbf{q}_1, \mathbf{q}_2) \mathcal{W}_R(q_3) \mathcal{W}_R(q_{12}) + (2 \text{ cyclic}) \right]. \end{aligned}$$

2.5.1 Gaussian case

As we assume that the stochastic parameter $\epsilon(\mathbf{k})$ is not correlated with the density field, we calculate the real space galaxy power spectrum in Gaussian case as

$$P_g(k, z) = \langle \epsilon^2 \rangle + D^2(z) P_{g,11}(k) + D^4(z) [P_{g,22}(k) + 2P_{g,13}(k)], \quad (2.75)$$

where

$$P_{g,11}(k) = c_1^2 \mathcal{W}_R^2(k) P_L(k) \quad (2.76)$$

is the linear bias term with linear matter power spectrum and $P_{g,22}$ and $P_{g,13}$ include the non-linear bias as well as the non-linear growth of the matter density field:

$$\begin{aligned} &P_{g,22}(k) \\ &= \frac{c_2^2}{2} \int \frac{d^3 q}{(2\pi)^3} \mathcal{W}_R^2(q) P_L(q) \mathcal{W}_R^2(|\mathbf{k} - \mathbf{q}|) P_L(|\mathbf{k} - \mathbf{q}|) \\ &\quad + 2c_1 c_2 \mathcal{W}_R(k) \int \frac{d^3 q}{(2\pi)^3} \mathcal{W}_R(q) P_L(q) \mathcal{W}_R(|\mathbf{k} - \mathbf{q}|) P_L(|\mathbf{k} - \mathbf{q}|) F_2^{(s)}(\mathbf{q}, \mathbf{k} - \mathbf{q}) \\ &\quad + 2c_1^2 \mathcal{W}_R^2(k) \int \frac{d^3 q}{(2\pi)^3} P_L(q) P_L(|\mathbf{k} - \mathbf{q}|) \left[F_2^{(s)}(\mathbf{q}, \mathbf{k} - \mathbf{q}) \right]^2 \end{aligned} \quad (2.77)$$

$$\begin{aligned} &P_{g,13}(k) \\ &= \frac{1}{2} c_1 c_3 \mathcal{W}_R^2(k) P_L(k) \sigma_R^2 + 3c_1^2 \mathcal{W}_R^2(k) P_L(k) \int \frac{d^3 q}{(2\pi)^3} P_L(q) F_3^{(s)}(\mathbf{q}, -\mathbf{q}, \mathbf{k}) \\ &\quad + 2c_1 c_2 \mathcal{W}_R(k) P_L(k) \int \frac{d^3 q}{(2\pi)^3} P_L(q) \mathcal{W}_R(q) \mathcal{W}_R(|\mathbf{k} - \mathbf{q}|) F_2(\mathbf{k}, -\mathbf{q}). \end{aligned} \quad (2.78)$$

Here,

$$\sigma_R^2 = \int \frac{d^3q}{(2\pi)^3} P_L(q) |\tilde{W}_R(q)|^2$$

is the root-mean-squared (r.m.s.) value of the smoothed linear density contrast at $z = 0$.

This equation is first derived in Smith et al. (2007) in the context of HaloPT, but they found the poor agreement between equation (2.75) and the halo power spectrum directly calculated from N-body simulation. However, it does not necessarily mean that the local bias ansatz of equation (2.72) is wrong. We rather attribute the failure of their modeling to the inaccurate modeling of the bias parameters (c_1 , c_2 and c_3) from the halo model. For example, the halo/galaxy power spectrum driven from local bias successfully models the halo/galaxy power spectrum from Millennium Simulation when fitting nonlinear bias parameters in Chapter 4.

Instead of using the bias parameters from the halo model, we shall treat the bias parameters as free parameters, and fit them to the observed galaxy power spectrum¹². In order to convert equation (2.75) into the practically useful form for fitting, we need to re-parametrize the bias parameters. It is because the theoretical template for fitting galaxy power spectrum shown in equation (2.75) has a few problems, as it was first pointed out by McDonald (2006). First, $P_{g,13}(k)$ in equation (2.75) contains σ_R^2 , which diverges, or is sensitive to the details of the small scale treatment, e.g. imposing a cut-off scale, choosing particular smoothing function, etc. Second, the first term in equation (2.77), one proportional to c_2^2 , approaches to a constant value on large scale limit i.e., $k \rightarrow 0$. The constant value can be large depending on the spectral index, or, again, sensitive to the small scale treatment.

In order to avoid these problems, we re-define the nonlinear bias parameters such that all terms sensitive to the small-scale treatment are absorbed into the parametrization. In other words, as we are interested in the power spectrum on sufficiently large scales, $k \ll 1/R$, we want to make the effect of small scale smoothing to be shown up only through the value of the bias parameters. On such large scales, we could approximate $W_R(k) = 1$, and the last term of $P_{g,13}(k)$ (Eq. [2.78]) is simply proportional to the linear power spectrum, and the proportionality constant depends only on the smoothing scale R . That is, if we

¹²For the goodness of the fitting method including the effect of fitting to extracting the cosmological parameters, see Chapter 4.

rewrite

$$\begin{aligned} & 2c_1c_2\mathcal{W}_R(k)P_L(k)\int\frac{d^3q}{(2\pi)^3}P_L(q)\mathcal{W}_R(q)\mathcal{W}_R(|\mathbf{k}-\mathbf{q}|)F_2(\mathbf{k},-\mathbf{q}) \\ & \equiv 4c_1c_2\sigma_R^2\mathcal{G}_R(k)\mathcal{W}_R^2(k)P_L(k), \end{aligned} \quad (2.79)$$

then, $\mathcal{G}_R(k)$ approaches to the constant value

$$\mathcal{G}_R(k) \rightarrow \frac{13}{84} + \frac{1}{4\sigma_R^2} \int \frac{d^3q}{(2\pi)^3} P_L(q) \mathcal{W}_R(q) \frac{\sin(qR)}{qR}, \quad (2.80)$$

as $k \rightarrow 0$ ¹³. See Figure 5.7 for the shape of $\mathcal{G}_R(k)$ for $R = 1, 2, 5$ and $10 \text{ Mpc}/h$. We also show the large scale asymptotic value of $\mathcal{G}_R(0)$ as a function of R in Figure 5.8. With the definition of $\mathcal{G}_R(k)$ in equation (2.79) and the non-linear matter power spectrum in equation (2.68), we rewrite the equation (2.75) as

$$\begin{aligned} P_g(k, z) = & \langle \epsilon^2 \rangle + c_1^2 \mathcal{W}_R^2(k) P_m(k, z) \\ & + D^2(z) [c_1c_3\sigma_R^2 + 8c_1c_2\sigma_R^2\mathcal{G}_R(k)] \mathcal{W}_R^2(k) P_L(k, z) \\ & + \frac{c_2^2}{2} \int \frac{d^3q}{(2\pi)^3} \mathcal{W}_R^2(q) P_L(q, z) \mathcal{W}_R^2(|\mathbf{k}-\mathbf{q}|) P_L(|\mathbf{k}-\mathbf{q}|, z) \\ & + 2c_1c_2\mathcal{W}_R(k) \int \frac{d^3q}{(2\pi)^3} \mathcal{W}_R(q) P_L(q, z) \\ & \times \mathcal{W}_R(|\mathbf{k}-\mathbf{q}|) P_L(|\mathbf{k}-\mathbf{q}|, z) F_2^{(s)}(\mathbf{q}, \mathbf{k}-\mathbf{q}). \end{aligned} \quad (2.82)$$

We re-parametrize the nonlinear bias parameters as,

$$P_0 = \langle \epsilon^2 \rangle + D^4(z) \frac{c_2^2}{2} \int \frac{dq}{2\pi^2} q^2 [P_L(q) \mathcal{W}_R^2(q)]^2 \quad (2.83)$$

$$b_1^2 = c_1^2 + D^2(z) [c_1c_3\sigma_R^2 + 8c_1c_2\mathcal{G}_R(k)\sigma_R^2] \quad (2.84)$$

$$b_2 = \frac{c_2}{b_1}, \quad (2.85)$$

then, the galaxy power spectrum becomes

$$P_g(k, z) = P_0 + b_1^2 [\mathcal{W}_R^2(k) P_m(k, z) + b_2 D^4(z) P_{b2}(k) + b_2^2 D^4(z) P_{b22}(k)], \quad (2.86)$$

¹³For general window function $\mathcal{W}_R(k)$, as $k \rightarrow 0$,

$$\begin{aligned} & \int \frac{d^3q}{(2\pi)^3} P_L(q) \mathcal{W}_R(q) \mathcal{W}_R(|\mathbf{k}-\mathbf{q}|) F_2^{(s)}(\mathbf{k}, -\mathbf{q}) \\ & \rightarrow \frac{17}{21} \sigma_R^2 + \frac{1}{6} \int \frac{d^3q}{(2\pi)^3} q P_L(q) \mathcal{W}_R(q) \frac{d\mathcal{W}_R(q)}{dq}. \end{aligned} \quad (2.81)$$

Therefore, if we do not employ the smoothing function, i.e. $\mathcal{W}_R(k) = 1$, the integration becomes $17/21\sigma_R^2$, and hence, the last term of $P_{g,13}(k)$ is simply $34/21c_1c_2\sigma^2P_L(k)$. This result coincides with McDonald (2006).

where $P_{b2}(k)$ and $P_{b22}(k)$ are given by

$$P_{b2}(k) = 2\mathcal{W}_R(k) \int \frac{d^3q}{(2\pi)^3} \mathcal{W}_R(q) P_L(q) \mathcal{W}_R(|\mathbf{k} - \mathbf{q}|) P_L(|\mathbf{k} - \mathbf{q}|) F_2^{(s)}(\mathbf{q}, \mathbf{k} - \mathbf{q}) \quad (2.87)$$

$$P_{b22}(k) = \frac{1}{2} \int \frac{d^3q}{(2\pi)^3} \mathcal{W}_R^2(q) P_L(q) [\mathcal{W}_R^2(|\mathbf{k} - \mathbf{q}|) P_L(|\mathbf{k} - \mathbf{q}|) - \mathcal{W}_R^2(q) P_L(q)]. \quad (2.88)$$

Note that equation (2.86) is the same as the original equation up to next-to-leading order, e.g. $\sigma_R^2 P_m \simeq \sigma_R^2 P_L + \mathcal{O}(P_L^3)$. As we have desired, the terms depending on the smoothing scale R are absorbed into the newly defined bias parameters P_0 , b_1 and b_2 , and $P_m(k)$, $P_{b2}(k)$, $P_{b22}(k)$ are independent of the smoothing scale on large scales, $k \gg 1/R$.

Note that b_1 we defined here reduces to the ‘effective bias’ of Heavens et al. (1998) in the $R \rightarrow 0$ limit, and in $k \rightarrow 0$ limit, equation (2.86) approaches to

$$P_g(k) \rightarrow P_0 + b_1^2 P_m(k),$$

the usual linear bias model plus a constant.

The ‘re-parametrized’ nonlinear bias parameters, P_0 , b_1 , b_2 , encode the complex galaxy formation processes, which will be very hard to model from the first principle (Smith et al., 2007). Nevertheless, the nonlinear galaxy power spectrum we calculate here has to be the ‘right’ prescription as long as the locality of bias assumption is correct in the quasi-nonlinear regime. In Chapter 4, we tested the nonlinear bias model in equation (2.86) against the halos/galaxies power spectrum of the Millennium Simulation (Springel et al., 2005). In order to test the prescription itself, we set P_0 , b_1 and b_2 as free parameters, and fit the measured power spectrum from Millennium simulation with equation (2.86). We found that nonlinear bias model provides a significantly better fitting than the linear bias model. In addition to that, we could reproduce the *correct* distance scales within a statistical error-bar, when marginalizing over three *free* nonlinear bias parameters.

2.5.2 Galaxy-matter cross power spectrum

As we shall marginalize over the bias parameters, the more do we add information about bias parameters, the better can we estimate the other cosmological parameters. The galaxy-matter cross power spectrum at high redshift can be a source of such an additional information, as it is proportional to the galaxy density contrast; thus, it also depends on the bias parameter. We can measure the galaxy-matter cross power spectrum from the galaxy-galaxy, and galaxy-CMB weak lensing measurements¹⁴.

¹⁴We study the galaxy-galaxy, and galaxy-CMB weak lensing on large scales in Chapter 6.

The fastest way of calculating the galaxy-matter cross power spectrum is using the calculation of galaxy-galaxy power spectrum. Let us abbreviate equation (2.73) as

$$\delta_g(\mathbf{k}) = \epsilon(\mathbf{k}) + c_1 \delta_R(\mathbf{k}) + c_2 \delta_g^{(2)}(\mathbf{k}) + c_3 \delta_g^{(3)}(\mathbf{k}). \quad (2.89)$$

Then, we can think of calculating the galaxy-galaxy power spectrum as

$$\begin{aligned} & \langle \delta_g(\mathbf{k}) \delta_g(\mathbf{k}') \rangle \\ &= \langle \epsilon^2 \rangle + c_1 \langle \delta_R(\mathbf{k}) \delta_g(\mathbf{k}') \rangle + \langle (c_2 \delta_g^{(2)}(\mathbf{k}) + c_3 \delta_g^{(3)}(\mathbf{k})) \delta_g(\mathbf{k}') \rangle \\ &= \langle \epsilon^2 \rangle + c_1 \langle \delta_R(\mathbf{k}) \delta_g(\mathbf{k}') \rangle + c_1 \langle [c_2 \delta_g^{(2)}(\mathbf{k}) + c_3 \delta_g^{(3)}(\mathbf{k})] \delta_R(\mathbf{k}') \rangle + \dots \\ &= \langle \epsilon^2 \rangle + c_1 [\langle \delta_R(\mathbf{k}) \delta_g(\mathbf{k}') \rangle + \langle \delta_R(\mathbf{k}') \delta_g(\mathbf{k}) \rangle - c_1 \langle \delta_R(\mathbf{k}) \delta_R(\mathbf{k}') \rangle] + \dots \end{aligned} \quad (2.90)$$

From equation (2.90), it is clear that adding up the terms proportional to c_1 in $P_g(\mathbf{k})$ are the same as $c_1 [2P_{gm}(\mathbf{k}) - c_1 P_m(k)]$. Therefore, the nonlinear galaxy-matter cross correlation function is

$$\begin{aligned} P_{gm}(k, z) &= c_1 \mathcal{W}_R^2(k) P_m(k, z) + D^2(z) \left[\frac{c_3}{2} \sigma_R^2 + 4c_2 \sigma_R^2 \mathcal{G}_R(k) \right] \mathcal{W}_R^2(k) P_L(k, z) \\ &+ c_2 \mathcal{W}_R(k) \int \frac{d^3 q}{(2\pi)^3} \mathcal{W}_R(q) P_L(q, z) \\ &\times \mathcal{W}_R(|\mathbf{k} - \mathbf{q}|) P_L(|\mathbf{k} - \mathbf{q}|, z) F_2^{(s)}(\mathbf{q}, \mathbf{k} - \mathbf{q}). \end{aligned} \quad (2.91)$$

We also re-parametrize the bias for this case,

$$\bar{b}_1 = c_1 + D^2(z) \left[\frac{c_3}{2} \sigma_R^2 + 4c_2 \sigma_R^2 \mathcal{G}_R(k) \right] \quad (2.92)$$

$$\bar{b}_2 = \frac{c_2}{\bar{b}_1}, \quad (2.93)$$

so that the galaxy-matter cross power becomes

$$P_{gm}(k, z) = \bar{b}_1 \left[\mathcal{W}_R^2(k) P_m(k, z) + \frac{\bar{b}_2}{2} D^4(z) P_{b2}(k) \right], \quad (2.94)$$

where $P_{b2}(k)$ is defined in equation (2.87). Note that when $\sigma_R \ll 1$, $\bar{b}_1 \sim b_1$ and $\bar{b}_2 \simeq b_2$.

2.5.3 non-Gaussian case

What about the non-Gaussian correction term? We calculate the non-Gaussian term by substituting the real space galaxy kernels into equation (2.66).

$$\begin{aligned}
& \Delta P_{g,nG}(k, z) \\
&= 2c_1 f_{\text{NL}} D^3(z) \mathcal{W}_R(k) \mathcal{M}(k) \int \frac{d^3 q}{(2\pi)^3} \mathcal{M}(q) \mathcal{M}(|\mathbf{k} - \mathbf{q}|) P_\phi(q) [2P_\phi(k) + P_\phi(|\mathbf{k} - \mathbf{q}|)] \\
&\quad \times \left[c_2 \mathcal{W}_R(q) \mathcal{W}_R(|\mathbf{k} - \mathbf{q}|) + 2c_1 \mathcal{W}_R(k) F_2^{(s)}(\mathbf{q}, \mathbf{k} - \mathbf{q}) \right] \\
&= c_1^2 \mathcal{W}_R^2(k) \Delta P_{m,ng}(k, z) + 4c_1 c_2 D^3(z) \sigma_R^2 f_{\text{NL}} \mathcal{W}_R(k) \mathcal{F}_R(k) \frac{P_L(k)}{M(k)} \tag{2.95}
\end{aligned}$$

Here, $\Delta P_{m,nG}$ is the non-Gaussian correction to the matter power spectrum, and

$$\mathcal{F}_R(k) \equiv \frac{1}{2\sigma_R^2} \int \frac{d^3 q}{(2\pi)^3} \mathcal{M}_R(q) \mathcal{M}_R(|\mathbf{k} - \mathbf{q}|) P_\phi(q) \left[\frac{P_\phi(|\mathbf{k} - \mathbf{q}|)}{P_\phi(k)} + 2 \right] \tag{2.96}$$

is a function which is unity on large scales ($k \ll 1/R$, See, e. g. Matarrese & Verde, 2008). See Figure 5.6 for the shape of $\mathcal{F}_R(k)$ for $R = 1, 2, 5$ and $10 \text{ Mpc}/h$.

The first term in equation (2.95) is simply the non-Gaussian matter power spectrum multiplied by the linear bias factor. The second term in Eq. (2.95) is the non-Gaussianity term generated by non-linear bias, and shows the same behavior as the scale dependent bias from local type primordial non-Gaussianity¹⁵. In fact, this term reduces to the result of MLB formula (Matarrese et al., 1986; Matarrese & Verde, 2008, see, Appendix K.2 for derivation)

$$\Delta P_{g,nG}(k, z) = 4c_1(c_1 - 1)\delta_c D(z) \alpha f_{\text{NL}} \mathcal{W}_R(k) \mathcal{F}_R(k) \frac{P_L(k)}{M(k)} \tag{2.97}$$

in the linear regime and for the high-peak limit of the halo model¹⁶. It is sufficient to show that $c_1 c_2 D^2(z) \sigma_R^2$ becomes $\alpha c_1(c_1 - 1)\delta_c$ in the high-peak limit. Consider the bias parameters from the halo model: (Scoccimarro et al., 2001b)

$$c_1 = 1 + \frac{\alpha\nu^2 - 1}{\delta_c} + \frac{2p/\delta_c}{1 + (\alpha\nu^2)^p} \tag{2.98}$$

$$c_2 = \frac{8}{21}(c_1 - 1) + \frac{\alpha\nu^2}{\delta_c^2} (\alpha\nu^2 - 3) + \frac{2p/\delta_c}{1 + (\alpha\nu^2)^p} \left(\frac{1 + 2p}{\delta_c} + 2 \frac{\alpha\nu^2 - 1}{\delta_c} \right), \tag{2.99}$$

where $\delta_c \simeq 1.686$ is the critical overdensity above which halo forms, and $\nu \equiv \delta_c/(D^2(z)\sigma_R^2)$. For Press-Schechter mass function (spherical collapse, Press & Schechter, 1974; Mo & White,

¹⁵For the scale dependent bias, see the introduction in Chapter 5, and Appendix I.3.

¹⁶ α here is the same as q in Carbone et al. (2008). We reserve q for the Fourier space measure.

1996), $\alpha = 1$, $p = 0$, and for Sheth-Tormen mass function (ellipsoidal collapse, Sheth & Tormen, 1999; Sheth et al., 2001), $\alpha = 0.75$, $p = 0.3$. And in the high-peak limit ($\nu \gg 1$), we approximate $c_1 - 1 \simeq \alpha\nu^2/\delta_c$, and $c_2 \simeq \alpha^2\nu^4/\delta_c^2$. Therefore,

$$c_1 c_2 D^2(z) \sigma_R^2 \simeq c_1 \alpha^2 \nu^4 \left(\frac{D^2(z) \sigma_R^2}{\delta_c^2} \right) = c_1 \alpha^2 \nu^2 \simeq \alpha c_1 (c_1 - 1) \delta_c. \quad (2.100)$$

This relation motivate us to define a new bias parameter $\tilde{b}_2 \equiv \sigma_R^2 D^2(z) c_2 / c_1$, which approaches $\tilde{b}_2 \rightarrow \alpha \delta_c$ for the high peak limit of the halo model.

By using a re-parametrized bias, b_1 and \tilde{b}_2 , the non-Gaussian correction term becomes

$$\begin{aligned} & \Delta P_{g,nG}(k, z) \\ &= b_1^2 \left[\mathcal{W}_R^2(k) \Delta P_{m,ng}(k, z) + 6\tilde{b}_2 f_{\text{NL}} D(z) \mathcal{W}_R(k) \mathcal{F}_R(k) \frac{H_0^2 \Omega_m P_L(k)}{k^2 T(k)} \right], \end{aligned} \quad (2.101)$$

and on large scales ($k \ll 1/R$), for high-peak, the formula reduces to the usual form in the literature (Dalal et al., 2008; Matarrese & Verde, 2008; Slosar et al., 2008; Afshordi & Tolley, 2008; Taruya et al., 2008; McDonald, 2008; Sefusatti, 2009):

$$\Delta P_{g,nG}(k, z) = b_1^2 \left[\Delta P_{m,ng}(k, z) + 6\alpha \delta_c f_{\text{NL}} D(z) \frac{H_0^2 \Omega_m P_L(k)}{k^2 T(k)} \right]. \quad (2.102)$$

Although equation (2.101) coincides with equation (2.102) for high peak limit, it may not be the dominant contribution of scale-dependent bias for intermediate size peaks where the nonlinear bias b_2 is actually small. Recent study based on Peak Background Split method (Giannantonio & Porciani, 2010) suggests that for non-Gaussian case, the local ansatz [Eq. (2.72)] has to be modified to include the effect of Gaussian piece of gravitational potential $\phi(\mathbf{x})$ directly as

$$\delta_g(\mathbf{x}) = b_{10} \delta(\mathbf{x}) + b_{01} \phi(\mathbf{x}) + \frac{1}{2!} (b_{20} \delta^2(\mathbf{x}) + 2b_{11} \delta(\mathbf{x}) \phi(\mathbf{x}) + b_{02} \phi^2(\mathbf{x})) + \dots, \quad (2.103)$$

where b_{ij} s are bias parameters. If this holds, the non-Gaussianity signal from the power spectrum of very massive clusters (where \tilde{b}_2 is indeed close to $\alpha \delta_c$) is expected to be twice as high as the scale dependent bias in equation (2.102).

2.6 Nonlinear mater power spectrum in redshift space

In the previous sections, we have calculated the matter power spectrum and the galaxy power spectrum in *real* space. By real space, we mean an idealistic universe where

we can observe the *true* distance of the galaxies (or matter particles) relative to us. With galaxy survey alone, however, we cannot measure the true distance to the galaxies, as we infer the distance to a galaxy from the galaxy's spectral line shift by assuming the Hubble law. The problem here is that the observed spectral shift depends not only on the position of the galaxy (as a result of the expansion of the Universe), but also on the peculiar velocity of the galaxy. As a result, the three-dimensional map of the galaxies generated from galaxy surveys is different from the real space galaxy distribution. In contrast to the real space, we call the *observed* coordinate of galaxies the *redshift space*, and the radial distortion in the redshift space due to the peculiar velocity is called *redshift space distortion*.

We formulate the redshift space position vector \mathbf{s} as follow:

$$\mathbf{s} = \mathbf{x} + (1+z) \frac{v_r(\mathbf{x})}{H(z)} \hat{\mathbf{r}}. \quad (2.104)$$

Here, \mathbf{x} denotes the real space comoving position vector, and z denotes the redshift of galaxy without peculiar velocity, $H(z)$ is the Hubble parameter at that redshift, and v_r denotes the line-of-sight directional peculiar velocity. As redshift space distortion is due to the peculiar velocity, we can model it by using the peculiar velocity solution $\theta_{\mathbf{k}}(\tau)$ (Eq. [2.23]) of perturbation theory. In this section and the next section, we calculate the matter power spectrum and the galaxy power spectrum in redshift space, respectively.

How does the real space power spectrum changed under the redshift space distortion? In order to simplify the analysis, we make the *plane parallel* approximation that the galaxies are so far away that the radial direction is parallel to the $\hat{\mathbf{z}}$ direction¹⁷. Also, we define the *reduced* velocity field $\mathbf{u} \equiv \mathbf{v}/(f\mathcal{H})$ so that equation (2.104) becomes

$$\mathbf{s} = \mathbf{x} + f u_z(\mathbf{x}) \hat{\mathbf{z}}. \quad (2.105)$$

As $\mathbf{u}(\mathbf{k}) \equiv \mathbf{v}(\mathbf{k})/(f\mathcal{H}) = -i\mathbf{k}\theta_{\mathbf{k}}(\tau)/(k^2 f\mathcal{H})$, the Fourier transform of $u_z(\mathbf{x})$ becomes

$$\begin{aligned} u_z(\mathbf{k}, \tau) &= \frac{i\mu}{k} \sum_{n=1}^{\infty} \int \frac{d^3 q_1}{(2\pi)^3} \cdots \int \frac{d^3 q_n - 1}{(2\pi)^3} \int d^3 q_n \delta^D(\mathbf{k} - \sum_{i=1}^n \mathbf{q}_i) \\ &\quad \times G_n^{(s)}(\mathbf{q}_1, \mathbf{q}_2, \cdots, \mathbf{q}_n) \delta_1(\mathbf{q}_1, \tau) \cdots \delta_1(\mathbf{q}_n, \tau) \equiv \frac{i\mu}{k} \eta(\mathbf{k}, \tau), \end{aligned} \quad (2.106)$$

where $\mu = \mathbf{k} \cdot \hat{\mathbf{z}}/k$ is the directional cosine between the wavenumber vector $\hat{\mathbf{k}}$ and the line of sight direction $\hat{\mathbf{z}}$. Note that the time evolution of the new variable $u_z(\mathbf{k})$ only comes from the linear density contrast.

¹⁷For a redshift space distortion including a light-cone effect, see, e.g. de Lai & Starkman (1998); Yamamoto et al. (1999); Nishioka & Yamamoto (2000); Wagner et al. (2008).

Let us denote the real space over-density as $\delta_r(\mathbf{x})$, and the redshift space over-density as $\delta_s(\mathbf{s})$. The mass conservation relates the measure in real space d^3r and that in redshift space d^3s as

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

By using this relation, we find the *exact* relation between two over-densities in Fourier space (Scoccimarro, 2004; Matsubara, 2008).

$$\begin{aligned} \delta_s(\mathbf{k}) &= \int d^3s [1 + \delta_s(\mathbf{s})] e^{-i\mathbf{k}\cdot\mathbf{s}} - \int d^3x e^{-i\mathbf{k}\cdot\mathbf{x}} \\ &= \int d^3x [1 + \delta_r(\mathbf{x})] e^{-i\mathbf{k}\cdot[\mathbf{x} + f u_z(\mathbf{x})\hat{\mathbf{z}}]} - \int d^3x e^{-i\mathbf{k}\cdot\mathbf{x}} \\ &= \delta_r(\mathbf{k}) + \int d^3x e^{-i\mathbf{k}\cdot\mathbf{x}} \left(e^{-ik_z f u_z(\mathbf{x})} - 1 \right) [1 + \delta_r(\mathbf{x})] \end{aligned} \quad (2.108)$$

In order to calculate the 3rd order power spectrum, we Taylor-expand the exponential function up to 3rd order:

$$\begin{aligned} \delta_s(\mathbf{k}) &= \delta_r(\mathbf{k}) + f\mu^2\eta(\mathbf{k}) - \int d^3x e^{-i\mathbf{k}\cdot\mathbf{x}} \\ &\quad \times \left[ik_z f u_z(\mathbf{x})\delta_r(\mathbf{x}) + \frac{1}{2}k_z^2 f^2 u_z^2(\mathbf{x}) + \frac{1}{2}k_z^2 f^2 u_z^2(\mathbf{x})\delta_r(\mathbf{x}) - \frac{i}{6}k_z^3 f^3 u_z^3(\mathbf{x}) \right]. \end{aligned} \quad (2.109)$$

We calculate the 3rd order nonlinear matter kernels in redshift space from equation (2.109) and using the convolution theorem:

$$K_1^{(s)}(\mathbf{k}) = 1 + f\mu^2 \quad (2.110)$$

$$\begin{aligned} K_2^{(s)}(\mathbf{q}_1, \mathbf{q}_2) &= F_2^{(s)}(\mathbf{q}_1, \mathbf{q}_2) + f\mu^2 G_2^{(s)}(\mathbf{q}_1, \mathbf{q}_2) \\ &\quad + \frac{fk\mu}{2} \left(\frac{q_{1z}}{q_1^2} + \frac{q_{2z}}{q_2^2} \right) + \frac{(fk\mu)^2}{2} \frac{q_{1z}q_{2z}}{q_1^2 q_2^2} \end{aligned} \quad (2.111)$$

$$\begin{aligned} K_3^{(s)}(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) &= F_3^{(s)}(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) + f\mu^2 G_3^{(s)}(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) \\ &\quad + \frac{(fk\mu)^2}{6} \left(\frac{q_{1z}q_{2z}}{q_1^2 q_2^2} + \frac{q_{2z}q_{3z}}{q_2^2 q_3^2} + \frac{q_{3z}q_{1z}}{q_3^2 q_1^2} \right) + \frac{(fk\mu)^3}{6} \frac{q_{1z}q_{2z}q_{3z}}{q_1^2 q_2^2 q_3^2} \\ &\quad + \frac{fk\mu}{3} \left\{ F_2^{(s)}(\mathbf{q}_1, \mathbf{q}_2) \frac{q_{3z}}{q_3^2} + (2 \text{ cyclic}) \right\} \\ &\quad + \frac{(fk\mu)^2}{3} \left\{ G_2^{(s)}(\mathbf{q}_1, \mathbf{q}_2) \frac{q_{3z}q_{(1+2)z}}{q_3^2 |\mathbf{q}_1 + \mathbf{q}_2|^2} + (2 \text{ cyclic}) \right\} \\ &\quad + \frac{fk\mu}{3} \left\{ G_2^{(s)}(\mathbf{q}_1, \mathbf{q}_2) \frac{q_{(1+2)z}}{|\mathbf{q}_1 + \mathbf{q}_2|^2} + (2 \text{ cyclic}) \right\}. \end{aligned} \quad (2.112)$$

Note that the kernels we present here coincide those in equation (13) of Heavens et al. (1998) when setting $b_1 = 1$ and $b_2 = 0$.

Before calculating the power spectrum, it is instructive to compare the result here with other formulas in the literature. In linear regime, equation (2.109) reduces to

$$\delta_s(\mathbf{k}) = (1 + f\mu^2)\delta^{(1)}(\mathbf{k}), \quad (2.113)$$

and the redshift space matter power spectrum becomes

$$P_s(k, \mu, z) = (1 + f\mu^2)^2 D^2(z) P_L(k), \quad (2.114)$$

as shown in Kaiser (1987). That is, as a result of non-linear mapping between real and redshift space, the redshift space matter power spectrum is no longer spherically symmetric. This is called ‘Kaiser effect’ in the literature. Note that the redshift space distortion effect is larger for the line of sight direction ($\mu = 1$), and it does not change the power spectrum along the direction perpendicular to the line of sight ($\mu = 0$).

If we pick up the linear terms in equation (2.109)

$$\delta_s(\mathbf{k}) = \delta_r(\mathbf{k}) + f\mu^2\eta(\mathbf{k}), \quad (2.115)$$

and use the third order solution of $\delta_r(\mathbf{k})$ and $\eta(\mathbf{k})$, the redshift space matter power spectrum reduces to the formula given in Scoccimarro (2004):

$$P_s(k, \mu, z) = P_{\delta\delta}(k, z) + 2f\mu^2 P_{\delta\theta}(k, z) + f^2\mu^4 P_{\theta\theta}(k, z). \quad (2.116)$$

Here, $P_{\delta\delta}(k, z)$ is the same as the non-linear matter power spectrum $P_m(k, z)$ in real space [Eq. (2.68)], $P_{\delta\theta}(k, z)$ is the non-linear density-velocity cross power spectrum

$$P_{\delta\theta}(k, z) = D^2(z)P_L(k) + D^4(z) [P_{\delta\theta,22}(k) + P_{\delta\delta,13}(k) + P_{\theta\theta,13}(k)], \quad (2.117)$$

and $P_{\theta\theta}(k, z)$ is the non-linear velocity power spectrum

$$P_{\theta\theta}(k, z) = D^2(z)P_L(k) + D^4(z) [P_{\theta\theta,22}(k) + 2P_{\theta\theta,13}(k)], \quad (2.118)$$

where $P_{\delta\delta,13}(k)$ is shown in equation (2.70), and

$$\begin{aligned}
P_{\delta\theta,22}(k) &= 2 \int \frac{d^3q}{(2\pi)^3} P_L(q) P_L(|\mathbf{k} - \mathbf{q}|) F_2^{(s)}(\mathbf{q}, \mathbf{k} - \mathbf{q}) G_2^{(s)}(\mathbf{q}, \mathbf{k} - \mathbf{q}), \\
&= \frac{1}{98} \frac{k^3}{(2\pi)^2} \int dr P_L(kr) \int_{-1}^1 dx P_L(k\sqrt{1+r^2-2rx}) \\
&\quad \times \frac{(7x+3r-10rx^2)(7x-r-6rx^2)}{(1+r^2-2rx)^2}, \tag{2.119}
\end{aligned}$$

$$\begin{aligned}
P_{\theta\theta,22}(k) &= 2 \int \frac{d^3q}{(2\pi)^3} P_L(q) P_L(|\mathbf{k} - \mathbf{q}|) \left[G_2^{(s)}(\mathbf{q}, \mathbf{k} - \mathbf{q}) \right]^2 \\
&= \frac{1}{98} \frac{k^3}{(2\pi)^2} \int dr P_L(kr) \int_{-1}^1 dx P_L(k\sqrt{1+r^2-2rx}) \left[\frac{7x-r-6rx^2}{1+r^2-2rx} \right]^2, \tag{2.120}
\end{aligned}$$

$$\begin{aligned}
P_{\theta\theta,13}(k) &= 3 P_L(k) \int \frac{d^3q}{(2\pi)^3} P_L(q) G_3^{(s)}(\mathbf{q}, \mathbf{k}, -\mathbf{q}) \\
&= \frac{1}{504} \frac{k^3}{(2\pi)^2} P_L(k) \int dr P_L(kr) \\
&\quad \times \left[\frac{36}{r^2} - 246 + 12r^2 - 18r^4 + \frac{9}{r^3} (r^2-1)^3 (r^2+2) \ln \left(\frac{1+r}{|1-r|} \right) \right]. \tag{2.121}
\end{aligned}$$

2.6.1 Gaussian case

Substituting these kernels into equation (2.58), we find the 3rd order nonlinear redshift space matter power spectrum. The integration becomes particularly easier when we align $\hat{\mathbf{q}}_z$ (or $\hat{\phi}_q$ in spherical polar coordinate) parallel to the direction of the wavenumber vector, $\hat{\mathbf{k}}$. With this coordinate choice, the line of sight component of \mathbf{q} can be written as

$$q_z = q(\mu x - \sqrt{1-\mu^2} \sqrt{1-x^2} \cos \phi_q),$$

where x denotes the directional cosine between \mathbf{q} and \mathbf{k} , and we can integrate some part of angular integration analytically (Matsubara, 2008). After the analytical angular integration, the 3rd order matter power spectrum in redshift space becomes

$$P_s(k, \mu, z) = D^2(z) P_{s,11}(k, \mu) + D^4(z) [P_{s,22}(k, \mu) + 2P_{s,13}(k, \mu)], \tag{2.122}$$

where

$$P_{s,11}(k, \mu) = (1 + f\mu^2)^2 P_L(k), \quad (2.123)$$

$$P_{s,22}(k, \mu) = \sum_{n,m} \mu^{2n} f^m \frac{k^3}{(2\pi)^2} \int_0^\infty dr P_L(kr) \int_{-1}^1 dx \\ \times P_L\left(k(1+r^2-2rx)^{1/2}\right) \frac{A_{nm}(r, x)}{(1+r^2-2rx)^2}, \quad (2.124)$$

$$P_{s,13}(k, \mu) = (1 + f\mu^2) P_L(k) \sum_{n,m} \mu^{2n} f^m \frac{k^3}{(2\pi)^2} \int dr P_L(kr) B_{nm}(r), \quad (2.125)$$

with

$$A_{00} = \frac{1}{98}(3r + 7x - 10rx^2)^2 \\ A_{11} = \frac{2}{49}(3r + 7x - 10rx^2)^2 \\ A_{12} = \frac{1}{28}(1 - x^2)(7 - 6r^2 - 42rx + 48r^2x^2) \\ A_{22} = \frac{1}{196}(-49 + 114r^2 + 714rx + 637x^2 - 942r^2x^2 - 1890rx^3 + 1416r^2x^4) \\ A_{23} = \frac{1}{14}(1 - x^2)(7 - 6r^2 - 42rx + 48r^2x^2) \\ A_{24} = \frac{3}{16}r^2(x - 1)^2(x + 1)^2 \\ A_{33} = \frac{1}{14}(-7 + 6r^2 + 54rx + 35x^2 - 66r^2x^2 - 110rx^3 + 88r^2x^4) \\ A_{34} = \frac{1}{8}(1 - x^2)(2 - 3r^2 - 12rx + 15r^2x^2) \\ A_{44} = \frac{1}{16}(-4 + 3r^2 + 24rx + 12x^2 - 30r^2x^2 - 40rx^3 + 35r^2x^4)$$

and

$$B_{00} = \frac{1}{504} \left[\frac{2}{r^2}(6 - 79r^2 + 50r^4 - 21r^6) + \frac{3}{r^3}(r^2 - 1)^3(2 + 7r^2) \ln \left(\frac{1+r}{|1-r|} \right) \right] \\ B_{11} = 3B_{00} \\ B_{12} = \frac{1}{336} \left[\frac{2}{r^2}(9 - 89r^2 - 33r^4 + 9r^6) - \frac{9}{r^3}(r^2 - 1)^4 \ln \left(\frac{1+r}{|1-r|} \right) \right] \\ B_{22} = \frac{1}{336} \left[\frac{2}{r^2}(9 - 109r^2 + 63r^4 - 27r^6) + \frac{9}{r^3}(r^2 - 1)^3(1 + 3r^2) \ln \left(\frac{1+r}{|1-r|} \right) \right] \\ B_{23} = -\frac{1}{3}.$$

This non-linear redshift space power spectrum shows more complicated angular dependence than Kaiser effect.

In Section 3.3, we compare the redshift power spectrum in equation (2.122) with redshift space matter power spectrum we measured from a series of N-body simulations.

2.6.2 non-Gaussian case

We calculate the non-Gaussian correction to the matter power spectrum in redshift space by substituting the kernels in equation (2.110) and equation (2.111) into equation (2.66):

$$\begin{aligned}
& \Delta P_s(k, \mu, z) \\
&= 4f_{\text{NL}} D^3(z) (1 + f\mu^2) \mathcal{M}(k) \int \frac{d^3 q}{(2\pi)^3} \mathcal{M}(q) \mathcal{M}(|\mathbf{k} - \mathbf{q}|) \\
&\quad \times P_\phi(q) [2P_\phi(k) + P_\phi(|\mathbf{k} - \mathbf{q}|)] \left[F_2^{(s)}(\mathbf{q}, \mathbf{k} - \mathbf{q}) + f\mu^2 G_2^{(s)}(\mathbf{q}, \mathbf{k} - \mathbf{q}) \right. \\
&\quad \left. + \frac{fk\mu}{2} \left\{ \frac{q_z}{q^2} + \frac{k\mu - q_z}{|\mathbf{k} - \mathbf{q}|^2} \right\} + \frac{(fk\mu)^2}{2} \frac{q_z(k\mu - q_z)}{q^2 |\mathbf{k} - \mathbf{q}|^2} \right]. \tag{2.126}
\end{aligned}$$

As is the case for the non-Gaussianity correction to the matter power spectrum in real space, equation (2.126) is parametrically smaller than the non-linear terms: $P_{s,22}(k, \mu)$ and $P_{s,13}(k, \mu)$.

2.7 Nonlinear galaxy power spectrum in redshift space

Finally, in this section, we combine all three non-linear effects on the galaxy power spectrum we shall measure from galaxy surveys: the galaxy power spectrum in redshift space. As real to redshift mapping is the same for the total matter and the galaxies, the redshift space density contrast is also given by equation (2.109), but changing the real space density contrast $\delta_r(\mathbf{k})$ to the real space galaxy density contrast $\delta_g(\mathbf{k})$:

$$\begin{aligned}
\delta_{gs}(\mathbf{k}) &= \delta_g(\mathbf{k}) + f\mu^2 \eta(\mathbf{k}) - \int d^3 x e^{-i\mathbf{k}\cdot\mathbf{x}} \\
&\quad \times \left[ik_z f u_z(\mathbf{x}) \delta_g(\mathbf{x}) + \frac{1}{2} k_z^2 f^2 u_z^2(\mathbf{x}) + \frac{1}{2} k_z^2 f^2 u_z^2(\mathbf{x}) \delta_g(\mathbf{x}) - \frac{i}{6} k_z^3 f^3 u_z^3(\mathbf{x}) \right], \tag{2.127}
\end{aligned}$$

again, $\mu \equiv k_z/k$ is the cosine between wave vector \mathbf{k} and the line of sight direction, and $\eta(\mathbf{k})$ is defined as $\tilde{u}_z(\mathbf{k}) \equiv i\mu\eta(\mathbf{k})/k$ (see, equation (2.106)). Note that the first two terms in equation (2.127) lead the linear (Kaiser, 1984) and nonlinear (Scoccimarro, 2004) redshift space power spectrum which has been studied before.

We find the 3rd order kernels of the galaxy density contrast in redshift space by substituting equation (2.73) into equation (2.127). The kernels are

$$K_1^{(s)}(\mathbf{k}) = c_1 \mathcal{W}_R(k) + f\mu^2 \quad (2.128)$$

$$\begin{aligned} K_2^{(s)}(\mathbf{q}_1, \mathbf{q}_2) = & \frac{c_2}{2} \mathcal{W}_R(q_1) \mathcal{W}_R(q_2) + c_1 \mathcal{W}_R(q_{12}) F_2^{(s)}(\mathbf{q}_1, \mathbf{q}_2) + f\mu^2 G_2^{(s)}(\mathbf{q}_1, \mathbf{q}_2) \\ & + c_1 \frac{fk\mu}{2} \left[\frac{q_{1z}}{q_1^2} \mathcal{W}_R(q_2) + \frac{q_{2z}}{q_2^2} \mathcal{W}_R(q_1) \right] + \frac{(fk\mu)^2}{2} \frac{q_{1z}q_{2z}}{q_1^2 q_2^2} \end{aligned} \quad (2.129)$$

$$\begin{aligned} K_3^{(s)}(\mathbf{q}, -\mathbf{q}, \mathbf{k}) = & \frac{c_3}{6} \mathcal{W}_R^2(q) \mathcal{W}_R(k) + c_1 \mathcal{W}_R(k) F_3^{(s)}(\mathbf{q}, -\mathbf{q}, \mathbf{k}) \\ & + \frac{c_2}{3} \mathcal{W}_R(q) \left[\mathcal{W}_R(|\mathbf{k} + \mathbf{q}|) F_2^{(s)}(\mathbf{q}, \mathbf{k}) + \mathcal{W}_R(|\mathbf{k} - \mathbf{q}|) F_2^{(s)}(-\mathbf{q}, \mathbf{k}) \right] \\ & + f\mu^2 \left[G_3^{(s)}(\mathbf{q}, -\mathbf{q}, \mathbf{k}) + \frac{c_2}{6} \mathcal{W}_R^2(q) \right] \\ & + c_1 \frac{fk\mu}{3} \left[\mathcal{W}_R(|\mathbf{k} - \mathbf{q}|) F_2^{(s)}(-\mathbf{q}, \mathbf{k}) \frac{q_z}{q^2} - \mathcal{W}_R(|\mathbf{k} + \mathbf{q}|) F_2^{(s)}(\mathbf{q}, \mathbf{k}) \frac{q_z}{q^2} \right. \\ & \quad \left. + \mathcal{W}_R(q) \left\{ G_2^{(s)}(-\mathbf{q}, \mathbf{k}) \frac{k_z - q_z}{|\mathbf{k} - \mathbf{q}|^2} + G_2^{(s)}(\mathbf{q}, \mathbf{k}) \frac{k_z + q_z}{|\mathbf{k} + \mathbf{q}|^2} \right\} \right] \\ & + \frac{(fk\mu)^2}{3} \left[-\frac{c_1}{2} \frac{q_z^2}{q^4} \mathcal{W}_R(k) + G_2^{(s)}(-\mathbf{q}, \mathbf{k}) \frac{q_z(k_z - q_z)}{q^2 |\mathbf{k} - \mathbf{q}|^2} \right. \\ & \quad \left. - G_2^{(s)}(\mathbf{q}, \mathbf{k}) \frac{q_z(k_z + q_z)}{q^2 |\mathbf{k} + \mathbf{q}|^2} \right] \\ & - \frac{(fk\mu)^3}{6} \frac{q_z^2 k_z}{q^4 k^2}. \end{aligned} \quad (2.130)$$

Note that we only show $K_3^{(s)}(\mathbf{q}, -\mathbf{q}, \mathbf{k})$, as it is what we need to calculate the third order power spectrum. The kernels in equation (2.128)~(2.130), when setting $\mathcal{W}_R \equiv 1$, coincide those shown in Heavens et al. (1998)¹⁸, and also reduces to the kernels for the redshift space matter density contrast, equation (2.110)~(2.112) when setting $c_1 = 1$ and $c_2 = c_3 = 0$. For $\mu = 0$, it reduces to the kernels for the galaxy density contrast in real space, as redshift space distortion does not affect the perpendicular directional wave modes.

As is the case for the mater power spectrum, we can reproduce the formulas widely used in the literature by taking the linear terms in equation (2.127):

$$\delta_{gs}(\mathbf{k}) = \delta_g(\mathbf{k}) + f\mu^2 \eta(\mathbf{k}). \quad (2.131)$$

¹⁸Except that the real space kernels, J_3 and K_3 , presented in Heavens et al. (1998) have to be replaced by $F_3^{(s)}$ s and $G_3^{(s)}$ s in equation (2.34) and equation (2.35), respectively.

By keeping only linear order of $\delta_g(\mathbf{k})$ and $\eta(\mathbf{k})$, and setting $\mathcal{W}_R(k) = 1$, we find the linear galaxy power spectrum in redshift space (Kaiser, 1987)

$$P_{gs}(k, \mu, z) = b^2(1 + \beta\mu^2)^2 D^2(z) P_L(k), \quad (2.132)$$

where b is the linear bias parameter and $\beta \equiv f/b$. Also, including third order density field and velocity field, while keeping bias linear, we find the formula used in Shoji et al. (2009):

$$P_{gs}(k, \mu, z) = b^2 [P_{\delta\delta}(k) + 2\beta\mu^2 P_{\delta\theta}(k) + \beta^2\mu^4 P_{\theta\theta}(k)], \quad (2.133)$$

where $P_{\delta\delta}(k)$, $P_{\delta\theta}(k)$ and $P_{\theta\theta}(k)$ are shown in equation (2.68), equation (2.117) and equation (2.118), respectively.

2.7.1 Gaussian case

Finally, we obtain the 3rd order galaxy power spectrum in redshift space by substituting these kernels into equation (2.58):

$$P_{gs}(k, z) = \langle \epsilon^2 \rangle + D^2(z) P_{gs,11}(k) + D^4(z) [P_{gs,22}(k) + 2P_{gs,13}(k)], \quad (2.134)$$

where

$$P_{gs,11}(k) = (c_1 \mathcal{W}_R(k) + f\mu^2)^2 P_L(k) \quad (2.135)$$

is the same as the linear galaxy power spectrum in the redshift space (linear redshift space distortion with linear bias, Kaiser, 1987), and non-linear terms are

$$\begin{aligned} P_{gs,22}(k) &= 2 \int \frac{d^3 q}{(2\pi)^3} P_L(q) P_L(|\mathbf{k} - \mathbf{q}|) \left[K_2^{(s)}(\mathbf{q}, \mathbf{k} - \mathbf{q}) \right]^2 \\ &= 2 \int \frac{d^3 q}{(2\pi)^3} P_L(q) P_L(|\mathbf{k} - \mathbf{q}|) \left[\left(f\mu^2 G_2^{(s)}(\mathbf{q}, \mathbf{k} - \mathbf{q}) \right)^2 \right. \\ &\quad + \left(\frac{c_2}{2} \mathcal{W}_R(q) \mathcal{W}_R(|\mathbf{k} - \mathbf{q}|) + c_1 F_2^{(s)}(\mathbf{q}, \mathbf{k} - \mathbf{q}) \mathcal{W}_R(k) \right)^2 \\ &\quad + 2c_1 f\mu^2 \mathcal{W}_R(k) F_2^{(s)}(\mathbf{q}, \mathbf{k} - \mathbf{q}) G_2^{(s)}(\mathbf{q}, \mathbf{k} - \mathbf{q}) \left. \right] \\ &\quad + P_{gs,22}^{(rest)}(k, \mu; c_1, c_2), \end{aligned} \quad (2.136)$$

$$\begin{aligned}
P_{gs,13}(k, \mu) &= 3(c_1 \mathcal{W}_R(k) + f\mu^2)P_L(k) \int \frac{d^3q}{(2\pi)^3} P_L(q) K_3^{(s)}(\mathbf{q}, -\mathbf{q}, \mathbf{k}) \\
&= 3(c_1 \mathcal{W}_R(k) + f\mu^2)P_L(k) \int \frac{d^3q}{(2\pi)^3} P_L(q) \\
&\quad \times \left[\frac{c_3}{6} \mathcal{W}_R^2(q) \mathcal{W}_R(k) + c_1 \mathcal{W}_R(k) F_3^{(s)}(\mathbf{q}, -\mathbf{q}, \mathbf{k}) \right. \\
&\quad \left. + \frac{c_2}{3} \mathcal{W}_R(q) \left\{ \mathcal{W}_R(|\mathbf{k} + \mathbf{q}|) F_2^{(s)}(\mathbf{q}, \mathbf{k}) + \mathcal{W}_R(|\mathbf{k} - \mathbf{q}|) F_2^{(s)}(-\mathbf{q}, \mathbf{k}) \right\} \right. \\
&\quad \left. + f\mu^2 \left\{ G_3^{(s)}(\mathbf{q}, -\mathbf{q}, \mathbf{k}) + \frac{c_2}{6} \mathcal{W}_R^2(q) \right\} \right] + P_{gs,13}^{(rest)}(k, \mu; c_1). \tag{2.137}
\end{aligned}$$

Here, we only show terms which we can be explicitly identified as terms from non-linear evolution of matter ($P_{m,22}(k)$ or $P_{m,13}(k)$), from non-linear bias ($P_{b2}(k)$ or $P_{b22}(k)$), or from non-linear redshift space distortion ($P_{\theta\theta}(k)$ or $P_{\delta\theta}(k)$). Terms which cannot fall into those categories are called $P_{gs,22}^{(rest)}(k, \mu; c_1, c_2)$ and $P_{gs,13}^{(rest)}(k, \mu; c_1)$.

Combining equation (2.135) to (2.137), we calculate the galaxy power spectrum in redshift space

$$\begin{aligned}
P_{gs}(k, \mu, z) &= \langle \epsilon^2 \rangle + c_1^2 \mathcal{W}_R^2(k) P_{\delta\delta}(k, z) + D^2(z) [c_1 c_3 \sigma_R^2 + 8c_1 c_2 \sigma_R^2 \mathcal{G}_R(k)] \mathcal{W}_R^2(k) P_L(k, z) \\
&\quad + c_2^2 D^4(z) P_{b22}(k, z) + D^4(z) \frac{c_2^2}{2} \int \frac{d^3q}{(2\pi)^3} \mathcal{W}_R^2(q) P_L^2(q) + c_1 c_2 D^4(z) P_{b2}(k) \\
&\quad + 2f\mu^2 \left[c_1 \mathcal{W}_R(k) P_{\delta\theta}(k, z) + \frac{1}{2} D^2(z) \{ c_3 \sigma_R^2 + 8c_2 \sigma_R^2 \mathcal{G}_R(k) \} \mathcal{W}_R(k) P_L(k, z) \right. \\
&\quad \left. + \frac{1}{2} D^2(z) c_1 c_2 \sigma_R^2 \mathcal{W}_R(k) P_L(k) \right] + f^2 \mu^4 [P_{\theta\theta}(k, z) + c_2 D^2(z) \sigma_R^2 P_L(k, z)] \\
&\quad + D^4(z) [P_{gs,22}^{(rest)}(k, \mu; c_1, c_2) + 2P_{gs,13}^{(rest)}(k, \mu; c_1)], \tag{2.138}
\end{aligned}$$

where, $P_{\delta\delta}(k)$, $P_{b2}(k)$, $P_{b22}(k)$, $P_{\delta\theta}(k)$, $P_{\theta\theta}(k)$ are defined in equation (2.68), equation (2.87), equation (2.88), equation (2.117), and equation (2.118), respectively.

Following the discussion in Section 2.5, we shall absorb the small scale dependent quantities in equation (2.138) by re-defining the bias parameters: P_0 , b_1 and b_2 as is in equation (2.83), equation (2.84) and equation (2.85), respectively. In addition, we also introduce the new bias parameter b_θ :

$$b_\theta^2 = 1 + c_2 D^2(z) \sigma_R^2, \tag{2.139}$$

so that additional σ_R^2 term can be absorbed. By using the re-defined bias parameters, the

galaxy power spectrum in redshift space becomes

$$\begin{aligned}
P_{gs}(k, \mu, z) = & P_0 + b_1^2 [\mathcal{W}_R^2(k) P_m(k, z) + b_2 D^4(z) P_{b2}(k) + b_2^2 D^4(z) P_{b22}(k)] \\
& + 2f\mu^2 b_1 b_\theta \mathcal{W}_R(k) P_{\delta\theta}(k, z) + f^2 \mu^4 b_\theta^2 P_{\theta\theta}(k, z) \\
& + D^4(z) [P_{gs,22}^{(rest)}(k, \mu; b_1, b_2) + 2P_{gs,13}^{(rest)}(k, \mu; b_1)].
\end{aligned} \tag{2.140}$$

In quasi-nonlinear regime, r.m.s. density fluctuation is small $\sigma_R^2 \ll 1$; thus we make following approximations

$$b_1 \simeq c_1 + \frac{1}{2} D^2(z) [c_3 \sigma_R^2 + 8c_2 \sigma_R^2 \mathcal{G}_R(k)] \tag{2.141}$$

$$b_\theta \simeq 1 + \frac{c_2}{2} D^2(z) \sigma_R^2, \tag{2.142}$$

and

$$\sigma_R^2 P_L \simeq \sigma_R^2 P_{\delta\delta}(k) \simeq \sigma_R^2 P_{\delta\theta}(k) \simeq \sigma_R^2 P_{\theta\theta}(k),$$

which are true up to the fourth order in the linear density contrast. We also replace c_1 in the rest terms, $P_{gs,22}^{(rest)}(k, \mu, f; c_1, c_2)$ and $P_{gs,13}^{(rest)}(k, \mu, f; c_1)$, to b_1 , as it is consistent up to the same order. However, we set c_2 to be free, as this may allow us to check the consistency: $c_2 = b_1 b_2$.

Note that the new bias parameter b_θ is multiplied to the velocity divergence field, $\theta_{\mathbf{k}}$, thus we call it *velocity bias*. However, it does not mean that the velocity field itself is biased. The velocity bias is the bias introduced when ignoring the coupling between the density field and the velocity field in equation (2.127). For example, if one uses equation (2.133) as an estimator for the velocity power spectrum, then the measured velocity will be biased by at least a factor of $1 + b_\theta$, and the bias factor will increase when we also take ‘rest’ terms into account.

The ‘rest’ terms $P_{gs,22}^{(rest)}(k, \mu, f; c_1, c_2)$ $P_{gs,13}^{(rest)}(k, \mu, f; c_1)$ are defined as the collection of terms which cannot be simplified as either previously known terms or cannot be absorbed in the re-defined bias parameters. We show the explicit functional formula for the rest terms in Appendix D.

2.7.2 non-Gaussian case

We calculate the non-Gaussian correction to the galaxy power spectrum in redshift space as

$$\begin{aligned}
& \Delta P_g^s(k, \mu, z) \\
&= 2f_{\text{NL}} D^3(z) (c_1 \mathcal{W}_R(k) + f\mu^2) \mathcal{M}(k) \\
& \times \int \frac{d^3 q}{(2\pi)^3} \mathcal{M}(q) \mathcal{M}(|\mathbf{k} - \mathbf{q}|) P_\phi(q) [2P_\phi(k) + P_\phi(|\mathbf{k} - \mathbf{q}|)] \left[c_2 \mathcal{W}_R(q) \mathcal{W}_R(|\mathbf{k} - \mathbf{q}|) \right. \\
& \quad + 2c_1 \mathcal{W}_R(k) F_2^{(s)}(\mathbf{q}, \mathbf{k} - \mathbf{q}) + 2f\mu^2 G_2^{(s)}(\mathbf{q}, \mathbf{k} - \mathbf{q}) + (fk\mu)^2 \frac{q_z(k\mu - q_z)}{q^2 |\mathbf{k} - \mathbf{q}|^2} \\
& \quad \left. + b_1(fk\mu) \left\{ \mathcal{W}_R(|\mathbf{k} - \mathbf{q}|) \frac{q_z}{q^2} + \mathcal{W}_R(q) \frac{k\mu - q_z}{|\mathbf{k} - \mathbf{q}|^2} \right\} \right]. \tag{2.143}
\end{aligned}$$

We, again, find that the non-Gaussianity term comes from the non-linear redshift space mapping is parametrically small, and the dominant term is

$$\begin{aligned}
\Delta P_g^s(k, \mu, z) &\simeq 2f_{\text{NL}} D^3(z) (c_1 \mathcal{W}_R(k) + f\mu^2) \mathcal{M}(k) \\
& \times c_2 \int \frac{d^3 q}{(2\pi)^3} \mathcal{M}_R(q) \mathcal{M}_R(|\mathbf{k} - \mathbf{q}|) P_\phi(q) [2P_\phi(k) + P_\phi(|\mathbf{k} - \mathbf{q}|)] \\
& = 4f_{\text{NL}} D^3(z) (c_1 \mathcal{W}_R(k) + f\mu^2) c_2 \sigma_R^2 \mathcal{F}_R(k) \frac{P_L(k)}{\mathcal{M}(k)}. \tag{2.144}
\end{aligned}$$

By using $\tilde{b}_2 \equiv \sigma_R^2 D^2(z) c_2 / c_1$ we define in Section 2.5.3, the correction term becomes

$$\Delta P_g^s(k, \mu, z) = 4f_{\text{NL}} D(z) c_1 (c_1 \mathcal{W}_R(k) + f\mu^2) \tilde{b}_2 \mathcal{F}_R(k) \frac{P_L(k)}{\mathcal{M}(k)}. \tag{2.145}$$

On large scales ($k \ll 1/R$), $\mathcal{F}_R(k) = \mathcal{W}_R(k) = 1$, and for the highly biased tracers, $\tilde{b}_2 \simeq \alpha \delta_c^{19}$, we can rewrite the non-Gaussian term as

$$\Delta P_g^s(k, \mu, z) = 6\alpha \delta_c f_{\text{NL}} D(z) c_1 (c_1 + f\mu^2) \frac{H_0^2 \Omega_m P_L(k)}{k^2 T(k)}. \tag{2.146}$$

Therefore, on large scale, the non-Gaussian galaxy power spectrum in redshift space is given by

$$P_{gs}(k, \mu, z) = [b_1 + f\mu^2]^2 P_L(k) + 6\alpha f_{\text{NL}} \delta_c b_1 (b_1 + f\mu^2) \frac{H_0^2 \Omega_m P_L(k)}{k^2 T(k)} + P_0. \tag{2.147}$$

Here, we approximate $b_1 \simeq c_1$ as the equation is written in the linear order.

¹⁹ $\alpha = 1$ for Press-Schechter mass function, and $\alpha = 0.75$ for Sheth-Tormen mass function. For more discussion, see Section 2.5.3.

2.8 Summary

In this chapter, we calculate the nonlinear galaxy power spectrum in redshift space within Eulerian perturbation theory framework. After defining three conditions in the quasi non-linear regime, 1) sub-Horizon, 2) pressureless, 3) curl-free, we find the perturbative solutions for the non-linear evolution of the density field and velocity field of cosmic matter field. By using the perturbation theory solution, we calculate the matter power spectrum and the galaxy power spectrum in both real and redshift space. For each case, we also calculate the leading order correction to the power spectrum from the local primordial non-Gaussianity.

2.8.1 On the smoothing of density field

When applying the locality of bias assumption, we assume that the galaxy density contrast is a local function of the smoothed matter density contrast. While it is not certain if locality of bias works for the *real* density contrast or the smoothed density contrast, we intentionally adopt the smoothing function as it facilitates to access the effect of the small scale smoothing on the large scale power spectrum.

As we are mainly interested in the power spectrum on sufficiently large scales, $k \ll 1/R$, we want to absorb the effect of small scale smoothing into the value of the *free* bias parameters; thus, our theoretical template of the galaxy power spectrum in redshift space should not depend on the smoothing scale R . We test that every term in equation (2.140) is indeed independent on the smoothing scale as we desire.

In short, in the light of the re-definition of bias parameters, the smoothing does not affect the power spectrum on large scales; smoothing scale (or halo mass) will only changes the exact value of the bias parameters. Therefore, we show the unsmoothed equations in the summary below:

2.8.2 Summary of equations

The galaxy power spectrum in redshift space on the quasi-nonlinear scales is given by

$$\begin{aligned}
 P_{gs}(k, \mu, z) = & P_0 + b_1^2 [P_m(k, z) + b_2 D^4(z) P_{b2}(k) + b_2^2 D^4(z) P_{b22}(k)] \\
 & + 2f\mu^2 b_1 b_\theta P_{\delta\theta}(k, z) + f^2 \mu^4 b_\theta^2 P_{\theta\theta}(k, z) \\
 & + D^4(z) [P_{gs,22}^{(rest)}(k, \mu; b_1, b_2) + 2P_{gs,13}^{(rest)}(k, \mu; b_1)], \quad (2.148)
 \end{aligned}$$

where P_0 , b_1 , b_2 , b_θ , c_2 are free bias parameters, and the components of nonlinear power spectrum are listed as follow.

$$P_m(k, z) = D^2(z) [P_L(k) + D^2(z) \{P_{m,22}(k) + 2P_{m,13}(k)\}] \quad (2.149)$$

$$P_{b2}(k) = 2 \int \frac{d^3q}{(2\pi)^3} P_L(q) P_L(|\mathbf{k} - \mathbf{q}|) F_2^{(s)}(\mathbf{q}, \mathbf{k} - \mathbf{q}) \quad (2.150)$$

$$P_{b22}(k) = \frac{1}{2} \int \frac{d^3q}{(2\pi)^3} P_L(q) [P_L(|\mathbf{k} - \mathbf{q}|) - P_L(q)] \quad (2.151)$$

$$P_{\delta\theta}(k, z) = D^2(z) [P_L(k) + D^2(z) \{P_{\delta\theta,22}(k) + P_{\delta\delta,13}(k) + P_{\theta\theta,13}(k)\}] \quad (2.152)$$

$$P_{\theta\theta}(k, z) = D^2(z) [P_L(k) + D^2(z) \{P_{\theta\theta,22}(k) + 2P_{\theta\theta,13}(k)\}] \quad (2.153)$$

The P_{22} and P_{13} terms in P_m , $P_{\delta\theta}$, $P_{\theta\theta}$ are

$$P_{m,22}(k) = \frac{1}{98} \frac{k^3}{(2\pi)^2} \int dr P_L(kr) \int_{-1}^1 dx P_L(k\sqrt{1+r^2-2rx}) \left[\frac{7x+3r-10rx^2}{1+r^2-2rx} \right]^2 \quad (2.154)$$

$$P_{m,13}(k) = \frac{1}{504} \frac{k^3}{(2\pi)^2} P_L(k) \int dr P_L(kr) \times \left[\frac{12}{r^2} - 158 + 100r^2 - 42r^4 + \frac{3}{r^3} (r^2 - 1)^3 (7r^2 + 2) \ln \left(\frac{r+1}{|r-1|} \right) \right] \quad (2.155)$$

$$P_{\delta\theta,22}(k) = \frac{1}{98} \frac{k^3}{(2\pi)^2} \int dr P_L(kr) \int_{-1}^1 dx P_L(k\sqrt{1+r^2-2rx}) \times \frac{(7x+3r-10rx^2)(7x-r-6rx^2)}{(1+r^2-2rx)^2} \quad (2.156)$$

$$P_{\theta\theta,22}(k) = \frac{1}{98} \frac{k^3}{(2\pi)^2} \int dr P_L(kr) \int_{-1}^1 dx P_L(k\sqrt{1+r^2-2rx}) \left[\frac{7x-r-6rx^2}{1+r^2-2rx} \right]^2 \quad (2.157)$$

$$P_{\theta\theta,13}(k) = \frac{1}{504} \frac{k^3}{(2\pi)^2} P_L(k) \int dr P_L(kr) \times \left[\frac{36}{r^2} - 246 + 12r^2 - 18r^4 + \frac{9}{r^3} (r^2 - 1)^3 (r^2 + 2) \ln \left(\frac{1+r}{|1-r|} \right) \right], \quad (2.158)$$

and the ‘rest’ terms $P_{gs,22}^{(rest)}(k, \mu, f; b_1, c_2)$ $P_{gs,13}^{(rest)}(k, \mu, f; b_1)$ are shown in Appendix D (except that now we set all smoothing function to be unity; $\mathcal{W}_R = 1$).

By setting appropriate bias parameters and an angle parameter μ , one can reproduce the result for different cases we study in this chapter. If setting $P_0 = 0$ $b_1 = 1$, $b_2 = 0$, $b_\theta = 0$, equation (2.148) reduces to the non-linear matter power spectrum in redshift space, whose $\mu = 0$ slice is the non-linear matter power spectrum in real space, and, for $b_\theta = 0$, $\mu = 0$, equation (2.148) reduces to the non-linear galaxy power spectrum in real space.

Finally, on large scales, the dominant correction of galaxy power spectrum due to primordial local non-Gaussianity is given by

$$\Delta P_g^s(k, \mu, z) = 6\alpha\delta_c f_{\text{NL}} D(z) b_1 (b_1 + f\mu^2) \frac{H_0^2 \Omega_m P_L(k)}{k^2 T(k)}. \quad (2.159)$$

Chapter 3

Perturbation Theory Reloaded: analytical calculation of the non-linear matter power spectrum in real and redshift space

How accurately can we model the non-linear evolution of the density and velocity fields of cosmic matter fluctuation by using Eulerian Perturbation Theory? How large is the quasi-nonlinear regime at high redshifts? Is Third order Perturbation Theory good enough to model the non-linear mapping between real space and redshift space? In this chapter¹, we shall address these questions by comparing the nonlinear matter power spectrum from Eulerian Perturbation Theory² against a series of N-body simulations, in both real and redshift space.

3.1 *N*-body simulations and analysis method

We use the TVD (Ryu et al., 1993) code to simulate the evolution of $\delta(\mathbf{x}, \tau)$. The TVD code uses the Particle-Mesh scheme for gravity, and the Total-Variation-Diminishing (TVD) scheme for hydrodynamics, although we do not use hydrodynamics in our calculations. To increase the dynamic range of the derived power spectrum and check for convergence of the results, we use four box sizes, $L_{\text{box}} = 512, 256, 128, \text{ and } 64 \ h^{-1} \text{ Mpc}$, with the same number of particles, $N = 256^3$. (We use 512^3 meshes for doing FFT.) We use the following cosmological parameters: $\Omega_{\text{m}} = 0.27$, $\Omega_{\text{b}} = 0.043$, $\Omega_{\Lambda} = 0.73$, $h = 0.7$, $\sigma_8 = 0.8$, and $n_{\text{s}} = 1$. We output the simulation data at $z = 6, 5, 4, 3, 2$ and 1 for $512, 256$ and $128 \ h^{-1} \text{ Mpc}$, while only at $z = 6, 5, 4$ and 3 for $64 \ h^{-1} \text{ Mpc}$.

We suppress sampling variance of the estimated $P(k, z)$ by averaging $P(k, z)$ from 60, 50, 20, and 15 independent realizations of $512, 256, 128, \text{ and } 64 \ h^{-1} \text{ Mpc}$ simulations, respectively. We calculate the density field on 512^3 mesh points from the particle distribution

¹Previous versions of Sections 3.1 and 3.2 of this chapter were published in Jeong, D. & Komatsu, E. 2006, *Astrophys. J.*, 651, 619.

²For the analytic calculation of the non-linear power spectrum, see Chapter 2.

by the Cloud-In-Cell (CIC) mass distribution scheme. We then Fourier transform the density field and average $|\delta_{\mathbf{k}}(\tau)|^2$ within $k - \Delta k/2 \leq |\mathbf{k}| < k + \Delta k/2$ over the angle to estimate $P(k, z)$. Here, $\Delta k = 2\pi/L_{\text{box}}$. Finally, we correct the estimated $P(k)$ for loss of power due to the CIC pixelization effect using the window function calculated from 100 realizations of random particle distributions.

We use the COSMICS package (Bertschinger, 1995) to calculate the linear transfer function (with `linger`) and generate the input linear matter power spectrum and initial conditions (with `grafic`). We have increased the number of sampling points for the transfer function in k space from the default value of COSMICS, as the default sampling rate is too low to sample the baryonic acoustic oscillations accurately. (The default rate resulted in an artificial numerical smoothing of the oscillations.) We locate initial particles on the regular grid (i.e., we do not randomize the initial particle distribution), and give each particle the initial velocity field using the Zel’dovich approximation. This procedure suppresses shot noise in the derived power spectrum, which arises from randomness of particle distribution. We have checked this by comparing $P(k, z)$ from the initial condition to the input linear spectrum. However, some shot noise would arise as density fluctuations grow over time. While it is difficult to calculate the magnitude of shot noise from the structure formation, we estimate it by comparing $P(k, z)$ from large-box simulations with that from small-box simulations. We do not find any evidence for shot noise at $z \geq 1$; thus, we do not subtract shot noise from the estimated $P(k, z)$. To be conservative, we use 512, 256, 128, and 64 h^{-1} Mpc simulations to obtain $P(k, z)$ at $k \leq 0.24 h \text{ Mpc}^{-1}$, $0.24 < k \leq 0.5 h \text{ Mpc}^{-1}$, $0.5 < k \leq 1.4 h \text{ Mpc}^{-1}$, and $1.4 < k \leq 5 h \text{ Mpc}^{-1}$, respectively, to avoid the residual CIC pixelization effect and potential contaminations from unaccounted shot noise terms as well as artificial “transients” from initial conditions generated by the Zel’dovich approximation (Scoccimarro, 1998; Crocce et al., 2006). The initial redshifts are $z_{\text{initial}} = 27, 34, 42$, and 50 for 512, 256, 128, and 64 h^{-1} Mpc simulations, respectively. In Section 3.2.3 we show more on the convergence test (see Fig. 3.6 and Fig. 3.7).

3.2 Non-linear Matter Power spectrum in Real Space

3.2.1 Two theories: Halo Fit vs. Perturbation Theory

The next-to-leading order correction to $P(k)$ we calculate in Chapter 2 is

$$P_m(k, z) = D^2(z)P_{m,11}(k) + D^4(z)[2P_{m,13}(k) + P_{m,22}(k)], \quad (3.1)$$

where

$$P_{m,22}(k) = 2 \int \frac{d^3 q}{(2\pi)^3} P_{11}(q) P_{11}(|\mathbf{k} - \mathbf{q}|) \left[F_2^{(s)}(\mathbf{q}, \mathbf{k} - \mathbf{q}) \right]^2, \quad (3.2)$$

$$F_2^{(s)}(\mathbf{k}_1, \mathbf{k}_2) = \frac{5}{7} + \frac{2}{7} \frac{(\mathbf{k}_1 \cdot \mathbf{k}_2)^2}{k_1^2 k_2^2} + \frac{\mathbf{k}_1 \cdot \mathbf{k}_2}{2} \left(\frac{1}{k_1^2} + \frac{1}{k_2^2} \right), \quad (3.3)$$

$$\begin{aligned} 2P_{m,13}(k) &= \frac{2\pi k^2}{252} P_{11}(k) \int_0^\infty \frac{dq}{(2\pi)^3} P_{11}(q) \\ &\times \left[100 \frac{q^2}{k^2} - 158 + 12 \frac{k^2}{q^2} - 42 \frac{q^4}{k^4} \right. \\ &\left. + \frac{3}{k^5 q^3} (q^2 - k^2)^3 (2k^2 + 7q^2) \ln \left(\frac{k+q}{|k-q|} \right) \right]. \end{aligned} \quad (3.4)$$

While $F_2^{(s)}(\mathbf{k}_1, \mathbf{k}_2)$ should be modified for different cosmological models, the difference vanishes when $\mathbf{k}_1 \parallel \mathbf{k}_2$. The biggest correction comes from the configurations with $\mathbf{k}_1 \perp \mathbf{k}_2$, for which $[F_2^{(s)}(\Lambda\text{CDM})/F_2^{(s)}(\text{EdS})]^2 \simeq 1.006$ and $\lesssim 1.001$ at $z = 0$ and $z \geq 1$, respectively. Here, $F_2^{(s)}(\text{EdS})$ is given by equation (3.3), while $F_2^{(s)}(\Lambda\text{CDM})$ contains corrections due to $\Omega_m \neq 1$ and $\Omega_\Lambda \neq 0$ (Matsubara, 1995; Scoccimarro et al., 1998), and we used $\Omega_m = 0.27$ and $\Omega_\Lambda = 0.73$ at present. The information about different background cosmology is thus almost entirely encoded in the linear growth factor.

In principle, modeling of matter power spectrum may be developed using N -body simulations. This approach has been widely used in the literature. One method builds on the so-called HKLM formalism (Hamilton et al., 1991), which interpolates between the linear regime on large scales and the stable clustering regime on small scales using a fitting function to N -body simulations. The HKLM method was further elaborated by Peacock & Dodds (1996). The other method builds on the so-called halo model (Scherrer & Bertschinger, 1991), which was further elaborated by e.g., Seljak (2000); Smith et al. (2003).

We also compare the N -body data with the latest fitting formula for the nonlinear matter power spectrum in Appendix C of Smith et al. (2003). Motivated from the halo model, they divide the power spectrum by two parts:

$$\Delta_{NL}^2(k) = \Delta_Q^2(k) + \Delta_H^2(k), \quad (3.5)$$

where

$$\Delta^2(k) \equiv \frac{k^3 P(k)}{2\pi^2} \quad (3.6)$$

is the dimensionless power spectrum. The one halo term, Δ_H^2 , and two halo term, $\Delta_Q^2(k)$, are given by fitting formula as follow. First, the fitting formula for two halo term is

$$\Delta_Q^2(k) = \Delta_L^2(k) \left\{ \frac{[1 + \Delta_L^2]^{\beta_n}}{1 + \alpha_n \Delta_L^2(k)} \right\} \exp \left[-\frac{y}{4} - \frac{y^2}{8} \right], \quad (3.7)$$

where $y \equiv k/k_\sigma$, and k_σ is the nonlinear scale defined by $\sigma(k_\sigma^{-1}) = 1$. Here, $\sigma(R)$ is the r.m.s. value of smoothed linear density fluctuation with smoothing radius R :

$$\sigma^2(R) \equiv \int d \ln k \Delta_L^2(k) \exp [-k^2 R^2].$$

The fitting formula for one halo term is

$$\Delta_H^2(k) = \frac{a_n y^{3f_1(\Omega_m)}}{(1 + \mu_n y^{-1} + \nu_n y^{-2})(1 + b_n y^{f_2(\Omega_m)} + [c_n f_3(\Omega_m) y]^{3-\gamma_n})}. \quad (3.8)$$

The coefficients $(a_n, b_n, c_n, \gamma_n, \alpha_n, \beta_n, \mu_n, \nu_n)$ depend on the spectral properties: effective spectral index

$$n_{\text{eff}} \equiv -3 - \left. \frac{d \ln \sigma^2(R)}{d \ln R} \right|_{\sigma=1}, \quad (3.9)$$

and the spectral curvature

$$C \equiv - \left. \frac{d \ln^2 \sigma^2(R)}{d \ln R^2} \right|_{\sigma=1}. \quad (3.10)$$

They are

$$\begin{aligned} \log_{10} a_n = & 1.4861 + 1.83693 n_{\text{eff}} + 1.67618 n_{\text{eff}}^2 + 0.7940 n_{\text{eff}}^3 \\ & + 0.1670756 n_{\text{eff}}^4 - 0.620695 C \end{aligned} \quad (3.11)$$

$$\log_{10} b_n = 0.9463 + 0.9466 n_{\text{eff}} + 0.3084 n_{\text{eff}}^2 - 0.940 C \quad (3.12)$$

$$\log_{10} c_n = -0.2807 + 0.6669 n_{\text{eff}} + 0.3214 n_{\text{eff}}^2 - 0.0793 C \quad (3.13)$$

$$\gamma_n = 0.86485 + 0.2989 n_{\text{eff}} + 0.1631 C \quad (3.14)$$

$$\alpha_n = 1.38848 + 0.3701 n_{\text{eff}} - 0.1452 n_{\text{eff}}^2 \quad (3.15)$$

$$\beta_n = 0.8291 + 0.9854 n_{\text{eff}} + 0.3400 n_{\text{eff}}^2 \quad (3.16)$$

$$\log_{10} \mu_n = -3.54419 + 0.19086 n_{\text{eff}} \quad (3.17)$$

$$\log_{10} \nu_n = 0.95897 + 1.2857 n_{\text{eff}}. \quad (3.18)$$

Finally, for Λ CDM model, $f_1(\Omega_m)$, $f_2(\Omega_m)$, $f_3(\Omega_m)$ are given by

$$f_1(\Omega_m) = f_m \Omega_m^{-0.0307} + (1 - f_m) \Omega_m^{-0.0732} \quad (3.19)$$

$$f_2(\Omega_m) = f_m \Omega_m^{-0.0585} + (1 - f_m) \Omega_m^{-0.1423} \quad (3.20)$$

$$f_3(\Omega_m) = f_m \Omega_m^{0.0743} + (1 - f_m) \Omega_m^{0.0725}, \quad (3.21)$$

where $f_m = \Omega_\Lambda / (1 - \Omega_m)$. Note that $f_m = 1$ for flat- Λ CMD universe.

3.2.2 Results

Figure 3.1 compares $P(k, z)$ at $z = 1, 2, 3, 4, 5$ and 6 (from top to bottom) from simulations (dashed lines), PT (solid lines), and linear theory (dot-dashed lines). The PT predictions agree with simulations so well that it is actually difficult to see the difference between PT and simulations in Figure 3.1. The simulations are significantly above the linear theory predictions at high k .

To facilitate the comparison better, we show $\Delta^2(k, z)$ [Eq. (3.6)] in Figure 3.2. We find that the PT predictions (thin solid lines) agree with simulations (thick solid lines) to better than 1% accuracy for $\Delta^2(k, z) \lesssim 0.4$. On the other hand, the latest predictions from halo approach (Smith et al., 2003) (dotted lines) perform significantly worse than PT. This result suggests that one must use PT to model non-linearity in the weakly non-linear regime.

The baryonic features in the matter power spectrum provide a powerful tool to constrain the equation of state of dark energy. This method uses the fact that the CMB angular power spectrum sets the physical acoustic scale, and thus the features in the matter power spectrum seen on the sky and in redshift space may be used as the standard ruler, giving us the angular diameter distance out to the galaxy distribution at a given survey redshift as well as $H(z)$ (Matsubara & Szalay, 2003; Hu & Haiman, 2003; Seo & Eisenstein, 2003; Blake & Glazebrook, 2003). In order for this method to be viable, however, it is crucial to understand distortion on the baryonic acoustic oscillations caused by non-linearity. This has been investigated so far mostly using direct numerical simulations (Meiksin et al., 1999; Springel et al., 2005; White, 2005; Seo & Eisenstein, 2005). Meiksin et al. (1999) also compared the PT prediction with their N -body simulations at $z = 0$, finding that PT was a poor fit. This is because non-linearity at $z = 0$ is too strong to model by PT. Figure 3.4 shows that PT provides an accurate *analytical* account of non-linear distortion at $z > 1$: even at $z = 1$, the third peak at $k \simeq 0.18 h \text{ Mpc}^{-1}$ is modeled at a few percent level. At $z > 2$, all the oscillatory features are modeled to better than 1% accuracy. A slight deficit in power from N -body simulations at $k \sim 0.2 h \text{ Mpc}^{-1}$ relative to the perturbation theory predictions at $z = 2$ may be due to artificial transient modes from the Zel'dovich approximation used to generate initial conditions. One may eliminate such an effect by either using a smaller box-size or a better initial condition from the second-order Lagrangian perturbation theory (Crocce et al., 2006). As the power spectrum at $k > 0.24 h \text{ Mpc}^{-1}$ from $256 h^{-1} \text{ Mpc}$ simulations at $z = 2$ agrees with the perturbation theory predictions very well, we conclude that this small deficit in power at $k \sim 0.2 h \text{ Mpc}^{-1}$ is a numerical effect, most likely the transients in low-resolution simulations.

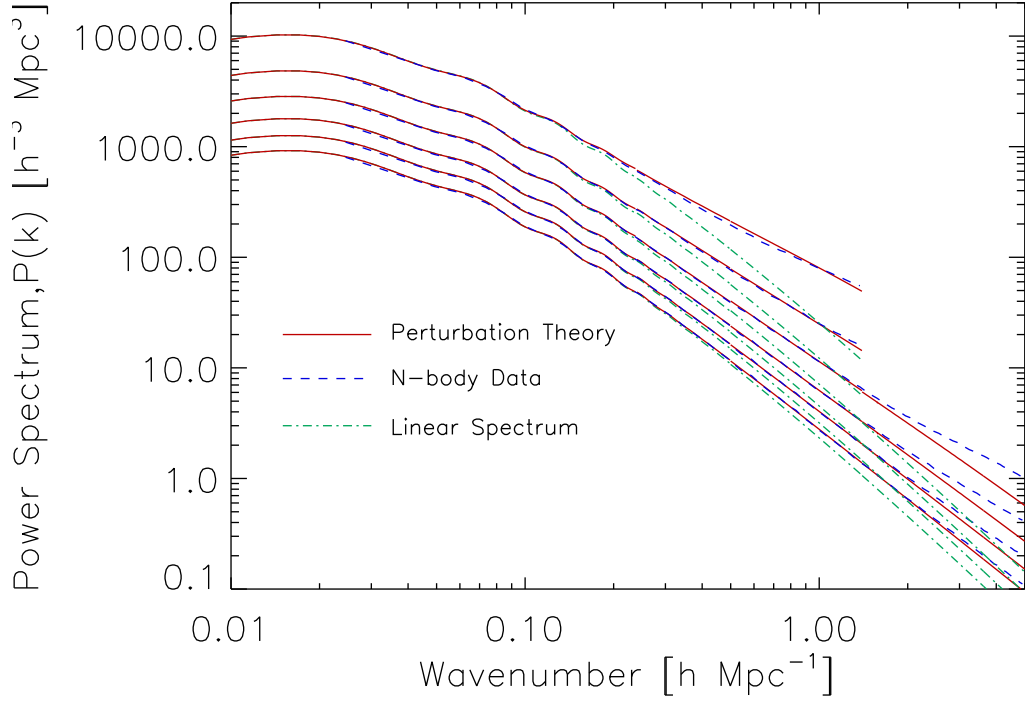


Figure 3.1: Power spectrum at $z = 1, 2, 3, 4, 5$ and 6 (from top to bottom), derived from N -body simulations (dashed lines), perturbation theory (solid lines), and linear theory (dot-dashed lines). We plot the simulation data from $512, 256, 128$, and $64 \ h^{-1} \text{ Mpc}$ simulations at $k \leq 0.24 \ h \text{ Mpc}^{-1}$, $0.24 < k \leq 0.5 \ h \text{ Mpc}^{-1}$, $0.5 < k \leq 1.4 \ h \text{ Mpc}^{-1}$, and $1.4 < k \leq 5 \ h \text{ Mpc}^{-1}$, respectively. Note that we did not run $64 \ h^{-1} \text{ Mpc}$ simulations at $z = 1$ or 2 .

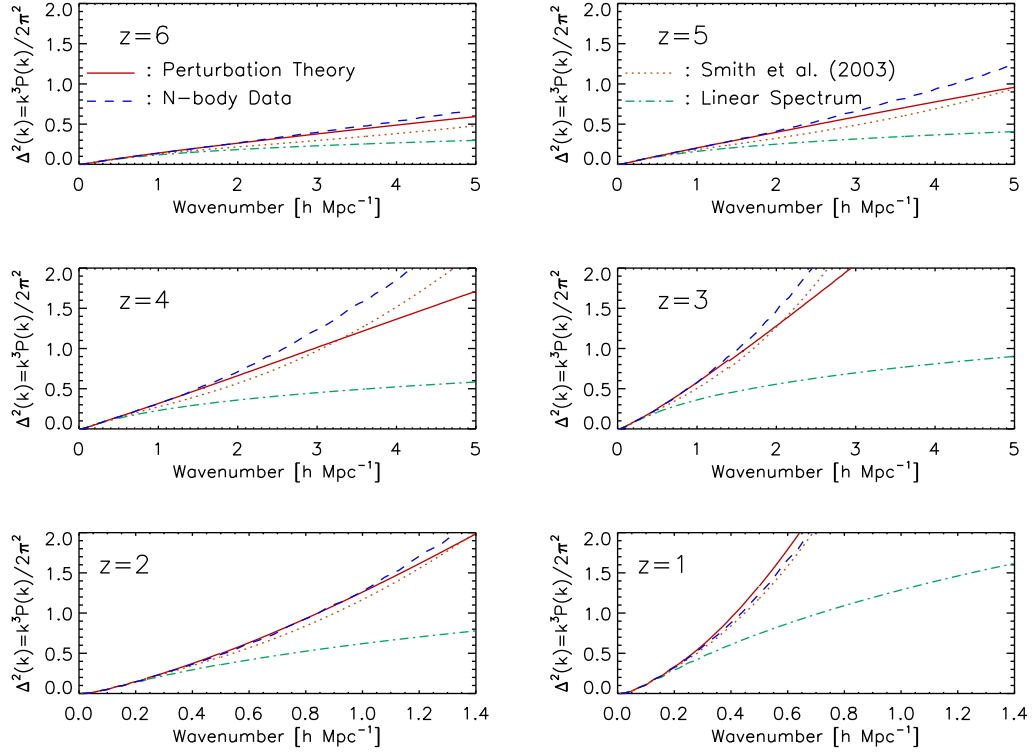


Figure 3.2: Dimensionless power spectrum, $\Delta^2(k)$. The solid and dashed lines show perturbation theory calculations and N -body simulations, respectively. The dotted lines show the predictions from halo approach (Smith et al., 2003). The dot-dashed lines show the linear power spectrum.

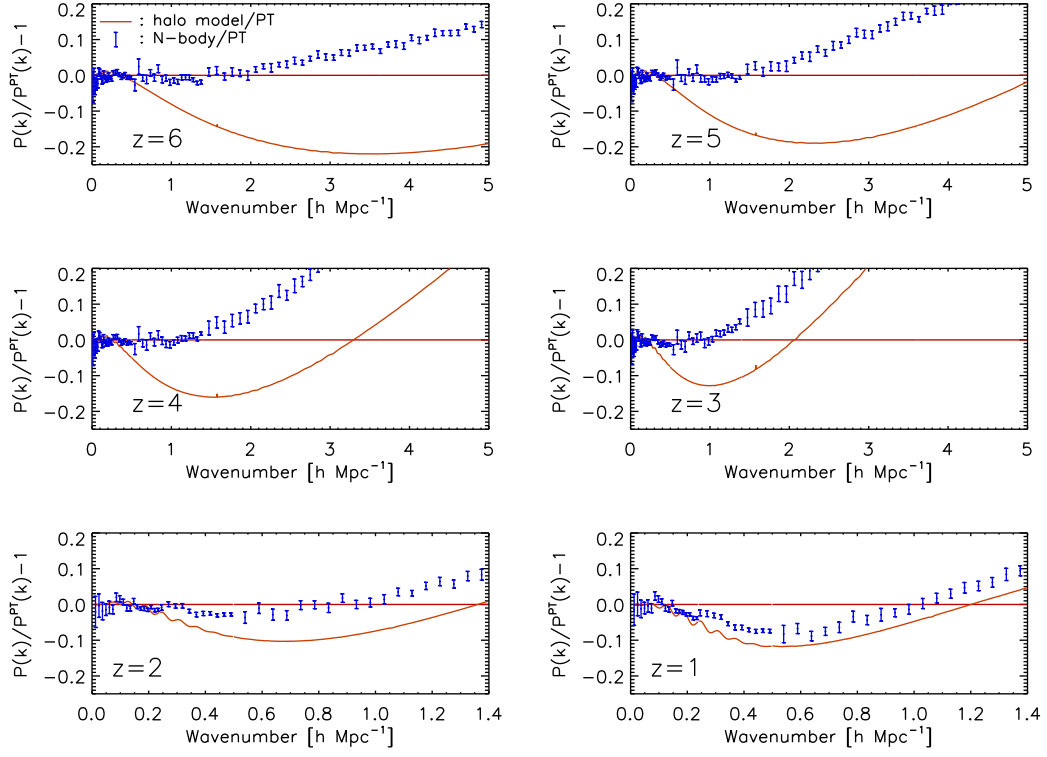


Figure 3.3: Residuals of Figure 3.2. The errorbars show the N -body data divided by the perturbation theory predictions minus one, while the solid curves show the halo model calculations given in Smith et al. (2003) divided by the perturbation theory predictions minus one. The perturbation theory predictions agree with simulations to better than 1% accuracy for $\Delta^2(k) \lesssim 0.4$.

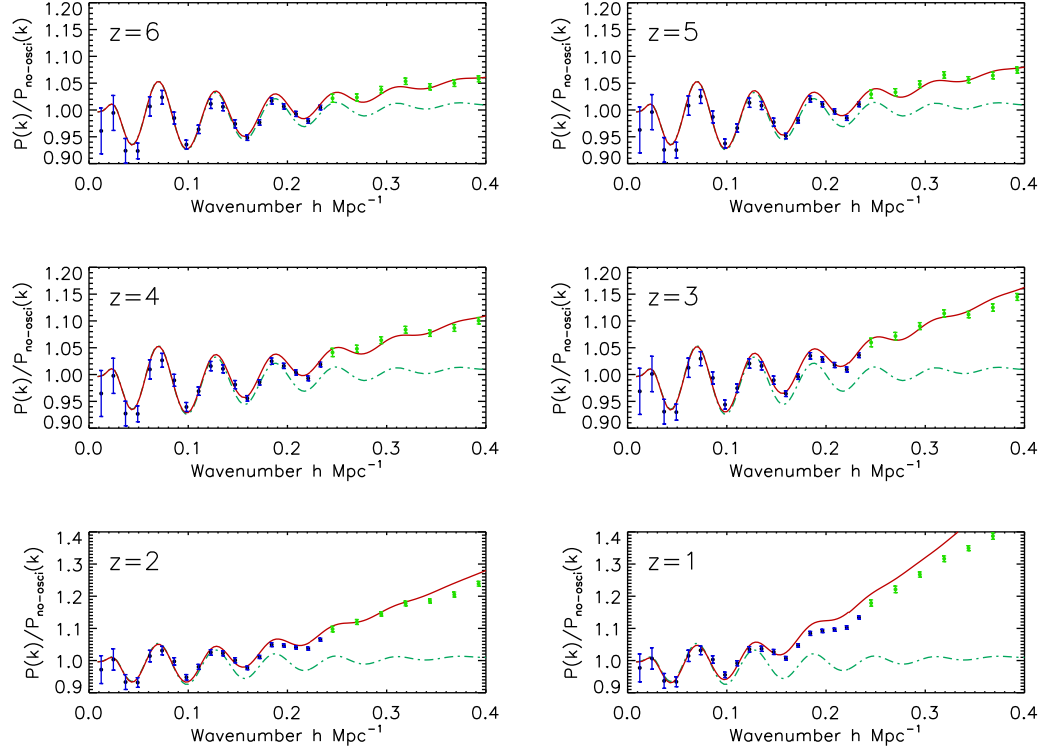


Figure 3.4: Non-linearity in baryonic acoustic oscillations. All of the power spectra have been divided by a smooth power spectrum without baryonic oscillations from equation (29) of Eisenstein & Hu (1998). The errorbars show N -body simulations, while the solid lines show perturbation theory calculations. The dot-dashed lines show the linear theory predictions. Perturbation theory describes non-linear distortion on baryonic oscillations very accurately at $z > 1$. Note that different redshift bins are not independent, as they have grown from the same initial conditions. The N -body data at $k < 0.24$ and $k > 0.24 \, h \, \text{Mpc}^{-1}$ are from 512 and 256 h^{-1} Mpc box simulations, respectively.

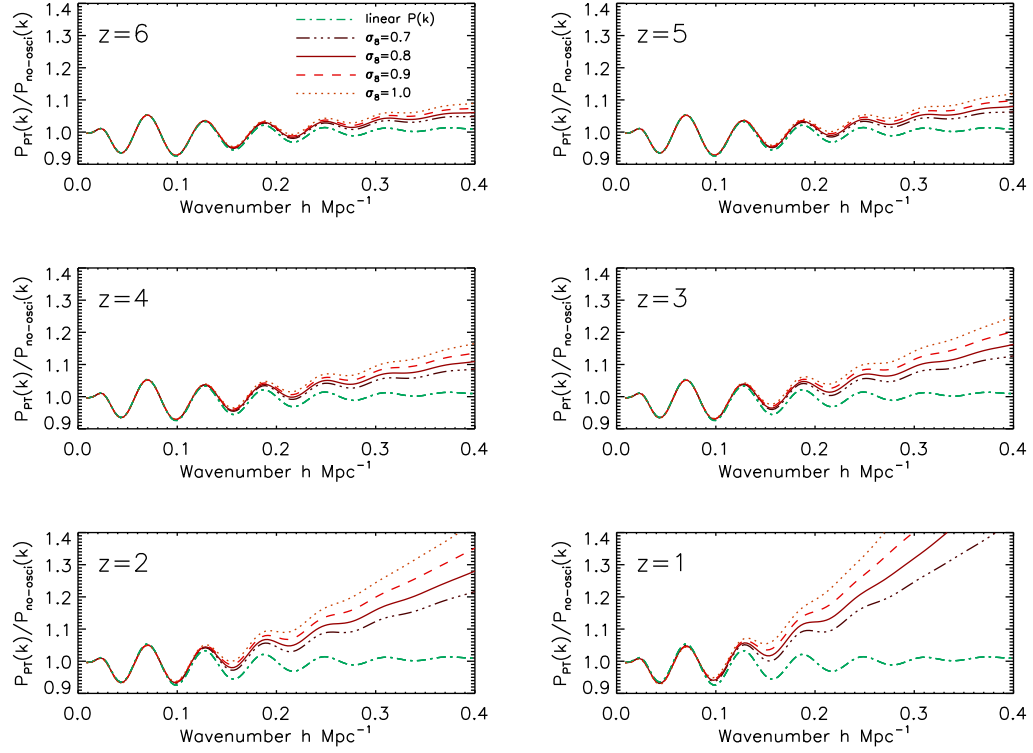


Figure 3.5: Non-linearity and the amplitude of matter fluctuations, σ_8 . In each panel the lines show the linear spectrum and non-linear spectrum with $\sigma_8 = 0.7, 0.8, 0.9$ and 1.0 from bottom to top.

How do the predicted non-linear power spectra depend on the amplitude of matter fluctuations? As the non-linear contributions to the power spectrum are given by the linear spectrum squared, a non-linear to linear ratio grows in proportion to σ_8^2 . In Fig 3.5 we show how the non-linear contributions increase as one increases σ_8 from 0.7 to 1.0 . This figure may be useful when one compares our results with the previous work that uses different values of σ_8 .

3.2.3 Convergence test

To test convergence of the power spectra derived from simulations and determine the valid range in wavenumber from each simulation box, we have run N -body simulations with four different box sizes, $L_{\text{box}} = 512, 256, 128$, and $64 h^{-1} \text{ Mpc}$, with the same number

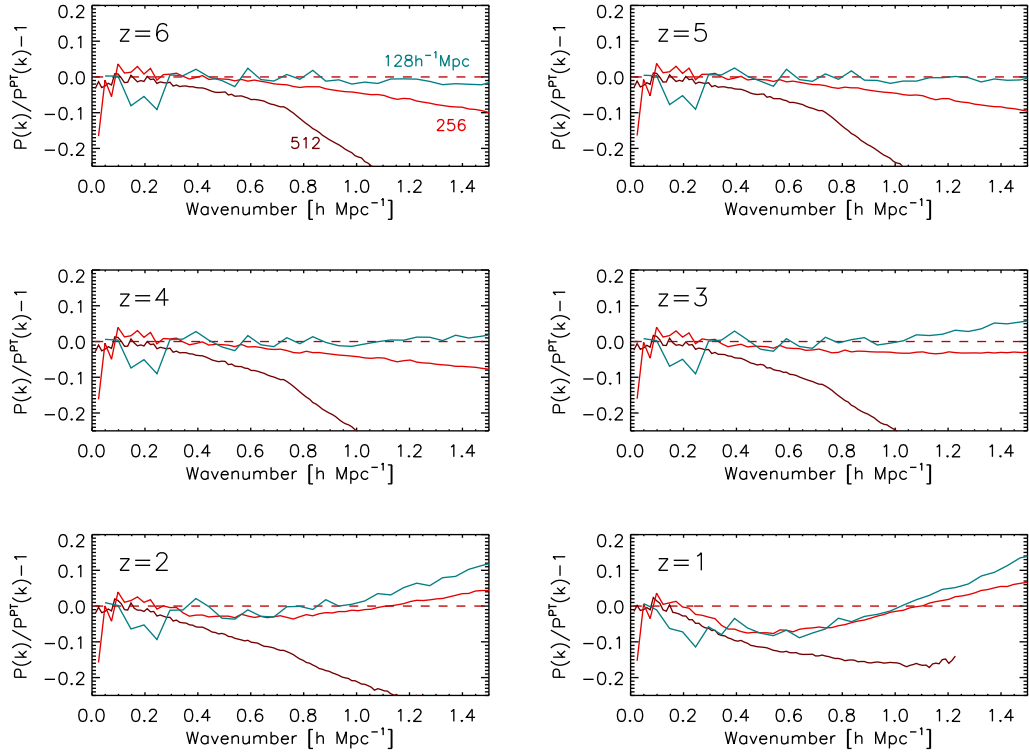


Figure 3.6: Convergence test. Fractional differences between the power spectra from N -body simulations in $L_{\text{box}} = 512, 256$, and $128 h^{-1} \text{ Mpc}$ box (from bottom to top lines) and the perturbation theory predictions in $k < 1.5 h \text{ Mpc}^{-1}$.

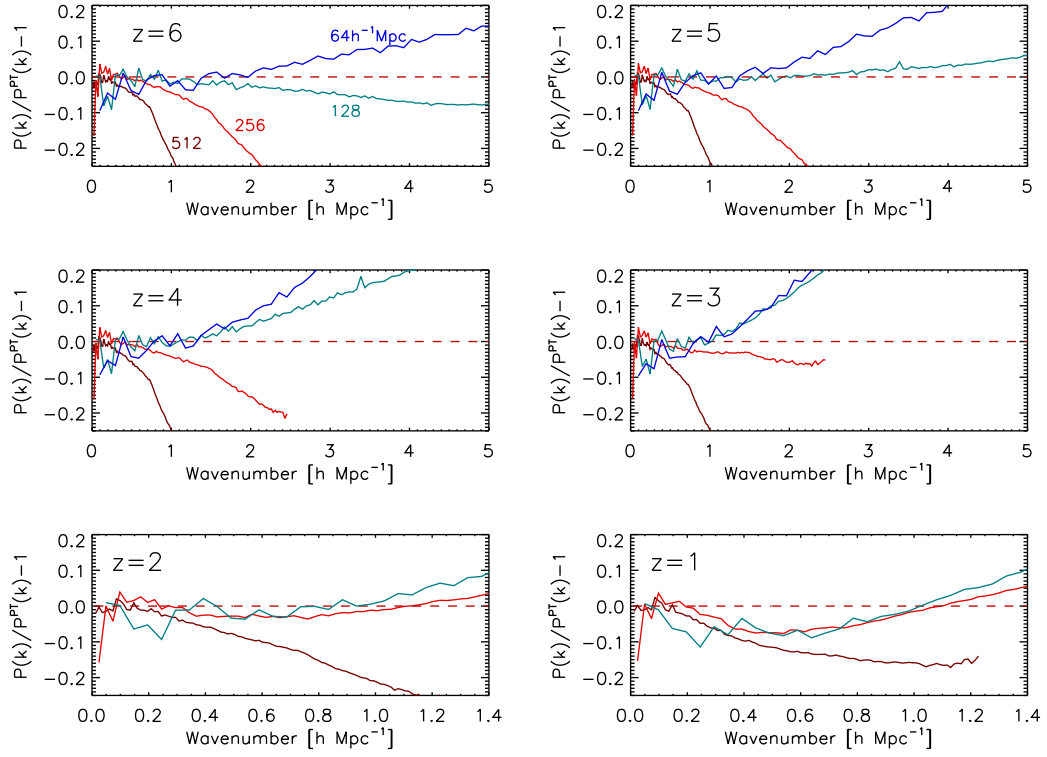


Figure 3.7: Convergence test. The same as left panel, but for simulations in $L_{\text{box}} = 512, 256, 128, \text{ and } 64 h^{-1} \text{ Mpc}$ box (from bottom to top lines) in the expanded range in wavenumber, $k < 5 h \text{ Mpc}^{-1}$.

of particles, $N = 256^3$. The initial redshifts are $z_{\text{initial}} = 27, 34, 42$, and 50 for 512, 256, 128, and 64 h^{-1} Mpc simulations, respectively.

Figure 3.6 and 3.7 shows that simulations with a larger box size lack power on larger scales due to the lack of resolution, as expected, while they have better statistics on large scales than those with a smaller box size. This figure helps us to determine the valid range in wavenumber from each simulation box. We find that one can use 512, 256, 128, and 64 h^{-1} Mpc simulations to calculate reliable estimates of the power spectrum in $k \leq 0.24 h \text{ Mpc}^{-1}$, $0.24 < k \leq 0.5 h \text{ Mpc}^{-1}$, $0.5 < k \leq 1.4 h \text{ Mpc}^{-1}$, and $1.4 < k \leq 5 h \text{ Mpc}^{-1}$, respectively.

3.3 Non-linear Matter Power spectrum in Redshift Space

In order to simulate the redshift space distortion from the N-body simulation, we displace the position of particles along the line of sight direction ($\hat{\mathbf{n}}$) according to the particle velocity of that direction ($\mathbf{v} \cdot \hat{\mathbf{n}}$) as

$$\mathbf{s} = \mathbf{x} + (1 + z) \frac{\mathbf{v} \cdot \hat{\mathbf{n}}}{H(z)} \hat{\mathbf{n}}, \quad (3.22)$$

where z and $H(z)$ are the redshift and the Hubble parameter at the snapshot of N-body simulation. For given simulation realization and redshift, we estimate three redshift space power spectra for $\hat{\mathbf{n}} = \hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}$ in order to increase the number of samples. Then, we follow the same procedure of estimating matter power spectrum described in Section 3.1. Note that the plane parallel approximation is *exact* in simulated redshift space defined by equation (3.22).

3.3.1 Three theories: Kaiser, Scoccimarro, and 3PT

We calculate the redshift space matter density contrast by expanding the exact relation [Eq. (2.108)]

$$\delta_s(\mathbf{k}, \tau) = \delta_r(\mathbf{k}, \tau) + \int d^3x e^{-i\mathbf{k} \cdot \mathbf{x}} \left(e^{-ik_z f u_z(\mathbf{x}, \tau)} - 1 \right) [1 + \delta_r(\mathbf{x}, \tau)] \quad (3.23)$$

up to third order in $\delta_1(\mathbf{k})$, [Eq. (2.109)]. Here, we rescale the velocity as $u(\mathbf{x}) = (1 + z)v(\mathbf{x})/(f(z)H(z))$. Then, the matter power spectrum in redshift space depends on both wavenumber k and angular cosine between wave vector and the line of sight direction μ , and is given by

$$P_s(k, \mu, z) = D^2(z)P_{s,11}(k, \mu) + D^4(z) [P_{s,22}(k, \mu) + 2P_{s,13}(k, \mu)], \quad (3.24)$$

where

$$P_{s,11}(k, \mu) = (1 + f\mu^2)^2 P_L(k), \quad (3.25)$$

$$P_{s,22}(k, \mu) = \sum_{n,m} \mu^{2n} f^m \frac{k^3}{(2\pi)^2} \int_0^\infty dr P_L(kr) \int_{-1}^1 dx \\ \times P_L \left(k(1 + r^2 - 2rx)^{1/2} \right) \frac{A_{nm}(r, x)}{(1 + r^2 - 2rx)^2}, \quad (3.26)$$

$$P_{s,13}(k, \mu) = (1 + f\mu^2) P_L(k) \sum_{n,m} \mu^{2n} f^m \frac{k^3}{(2\pi)^2} \int dr P_L(kr) B_{nm}(r), \quad (3.27)$$

where the functions $A_{nm}(r, x)$ and $B_{nm}(r)$ are given below equation (2.125). This formula is also given in Matsubara (2008), and is the same as $b_1 = 1$, $b_2 = 0$ case of Heavens et al. (1998).

In this section, we shall compare this *full* next-to-leading order power spectrum with two other formula which are widely used in the literature: linear redshift space power spectrum (Kaiser, 1987), and non-linear redshift space power spectrum from Scoccimarro (2004). Expanding equation (3.23) up to linear order leads

$$\delta_s(\mathbf{k}, \tau) = \delta(\mathbf{k}, \tau) + \mu^2 f^2 \eta(\mathbf{k}, \tau), \quad (3.28)$$

where $\delta(\mathbf{k})$ and $\eta(\mathbf{k})$ are defined as

$$\delta(\mathbf{k}, \tau) = \sum_{n=1}^\infty \int \frac{d^3 q_1}{(2\pi)^3} \cdots \int \frac{d^3 q_{n-1}}{(2\pi)^3} \int d^3 q_n \delta^D(\mathbf{k} - \sum_{i=1}^n \mathbf{q}_i) \\ \times F_n^{(s)}(\mathbf{q}_1, \cdots, \mathbf{q}_n) \delta_1(\mathbf{q}_1, \tau) \cdots \delta_1(\mathbf{q}_n, \tau), \quad (3.29)$$

$$\eta(\mathbf{k}, \tau) = \sum_{n=1}^\infty \int \frac{d^3 q_1}{(2\pi)^3} \cdots \int \frac{d^3 q_{n-1}}{(2\pi)^3} \int d^3 q_n \delta^D(\mathbf{k} - \sum_{i=1}^n \mathbf{q}_i) \\ \times G_n^{(s)}(\mathbf{q}_1, \cdots, \mathbf{q}_n) \delta_1(\mathbf{q}_1, \tau) \cdots \delta_1(\mathbf{q}_n, \tau). \quad (3.30)$$

By using linear $\delta(\mathbf{k}, \tau)$ and $\eta(\mathbf{k}, \tau)$, we calculate the linear redshift space power spectrum (Kaiser, 1987):

$$P_s^{\text{Kaiser}}(k, \mu, z) = (1 + f\mu^2)^2 D^2(z) P_L(k). \quad (3.31)$$

If we use $\delta(\mathbf{k}, \tau)$ and $\eta(\mathbf{k}, \tau)$ up to third order, we can also reproduce the nonlinear redshift space power spectrum of Scoccimarro (2004):

$$P_s^{\text{Scoccimarro}}(k, \mu, z) = P_{\delta\delta}(k, z) + 2f\mu^2 P_{\delta\theta}(k, z) + f^2\mu^4 P_{\theta\theta}(k, z), \quad (3.32)$$

where $P_{\delta\delta}(k, z)$ is the same as the matter power spectrum in real space $P_m(k, z)$, and

$$\begin{aligned}
& P_{\delta\theta}(k, \tau) \\
&= P_L(k, \tau) + 2 \int \frac{d^3 q}{(2\pi)^3} P_L(q, \tau) P_L(|\mathbf{k} - \mathbf{q}|, \tau) F_2^{(s)}(\mathbf{q}, \mathbf{k} - \mathbf{q}) G_2^{(s)}(\mathbf{q}, \mathbf{k} - \mathbf{q}) \\
&\quad + 3 P_L(k, \tau) \int \frac{d^3 q}{(2\pi)^3} P_L(q, \tau) \left[F_3^{(s)}(\mathbf{q}, -\mathbf{q}, \mathbf{k}) + G_3^{(s)}(\mathbf{q}, -\mathbf{q}, \mathbf{k}) \right], \tag{3.33}
\end{aligned}$$

$$\begin{aligned}
& P_{\theta\theta}(k, \tau) \\
&= P_L(k, \tau) + 2 \int \frac{d^3 q}{(2\pi)^3} P_L(q, \tau) P_L(|\mathbf{k} - \mathbf{q}|, \tau) \left[G_2^{(s)}(\mathbf{q}, \mathbf{k} - \mathbf{q}) \right]^2 \\
&\quad + 6 P_L(k, \tau) \int \frac{d^3 q}{(2\pi)^3} P_L(q, \tau) G_3^{(s)}(\mathbf{q}, -\mathbf{q}, \mathbf{k}), \tag{3.34}
\end{aligned}$$

are the density-velocity cross power spectrum, and velocity-velocity power spectrum, respectively. The explicit functional form of $P_{\delta\theta}$ and $P_{\theta\theta}$ are shown in equation (2.152) and equation (2.153), respectively.

3.3.2 Result

We first compare the angular averaged (monopole) redshift space power spectrum,

$$P_{red}(k) = \frac{1}{2} \int_{-1}^1 d\mu P_s(k, \mu), \tag{3.35}$$

from three theories (Kaiser, Scoccimarro, 3PT) to N-body simulation. Figure 3.8 shows the angular averaged redshift space power spectrum from N-body simulations (blue, dashed line) and three different theoretical power spectra at 6 different redshifts ($z = 1, 2, 3, 4, 5, 6$, top to bottom) : linear redshift space power spectrum (green, dot-dashed line, Kaiser, 1987), nonlinear redshift space power spectrum from (olive, dotted line Scoccimarro, 2004), and full third order nonlinear redshift space power spectrum in equation (3.24) (red, solid line).

On very large scales $k \lesssim 0.05 [h/\text{Mpc}]$, all power spectra agree with the prediction of Kaiser (1987): angular averaged redshift space power spectrum is boosted relative to the real space power spectrum by a factor of $(1 + 2f/3 + f^2/5) \simeq 28/15$. Linear theory power spectrum, however, deviate from the other curves on smaller scales, $k \gtrsim 0.2 [h/\text{Mpc}]$ even for $z = 6$ (bottom most) case. On smaller scales, two non-linear redshift power spectra are well above the linear theory prediction for all redshifts, and the Full third order redshift space power spectrum is larger than the non-linear power spectrum from Scoccimarro (2004).

To facilitate the comparison better, we show $\Delta_{red}^2(k, z)$, the dimensionless power spectrum [Eq. (3.6)], for each redshift in Figure 3.9. For higher redshifts, the full third

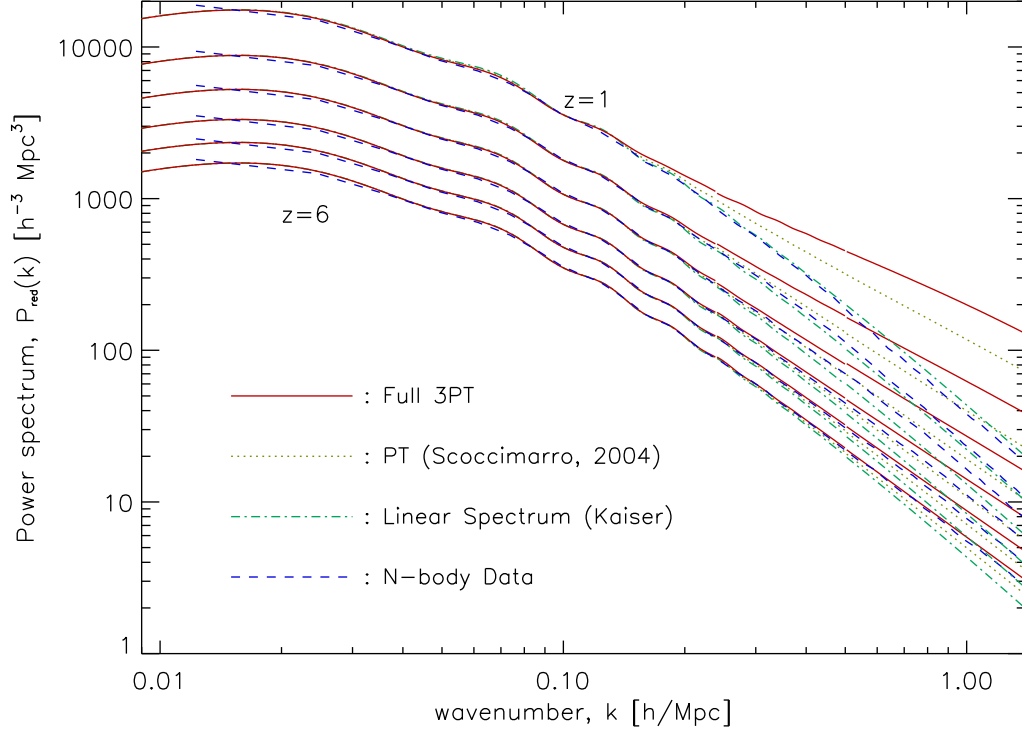


Figure 3.8: Comparing the angular averaged redshift space power spectrum for three different models at redshift $z = 1, 2, 3, 4, 5, 6$ (top to bottom). dashed line (*blue*) is the redshift space power spectrum calculated from the N-body simulations. Linear Kaiser $P(k)$ (dot dashed, *green* line), Scoccimarro's non-linear $P(k)$ (dotted, *olive* line), and full 3rd order non-linear redshift space $P(k)$ (solid, *red* line) are compared. Note that Finger of God effect is not yet included in the model and one can clearly see the suppression of the power on smaller scales. We plot the N-body power spectrum using the same data we used to draw Figure 3.1, so the k -range and the box size are all the same.

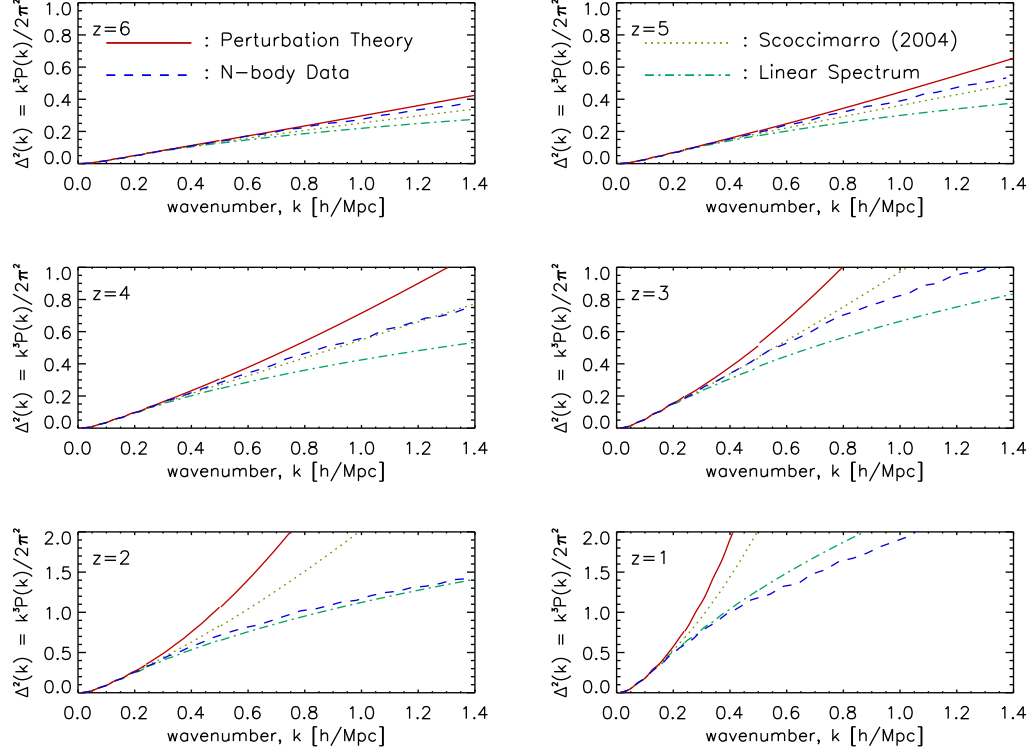


Figure 3.9: Dimensionless power spectrum, $\Delta_{red}^2(k)$, of redshift space for six different redshifts. The solid and dashed lines show full non-linear PT calculations and N-body simulations, respectively. The dotted line shows the Scoccimarro's formula and the dot-dashed line represents the linear kaiser spectrum. PT prediction agree with simulations very well at redshift $z = 6$ up to $k \simeq 0.8 h \text{ Mpc}^{-1}$, and at redshift $z = 5$ up to $k \simeq 0.5 h \text{ Mpc}^{-1}$. On low redshifts, however, N-body result shows power suppression relative to the PT power spectrum. It may be the Finger of God suppression. Note that the Scoccimarro's formula and linear kaiser spectrum are lower than the N-body result in the intermediate scale.

order perturbation theory models the non-linear redshift space distortion very accurately: $k \lesssim 0.8 h \text{ Mpc}^{-1}$, $k \lesssim 0.5 h \text{ Mpc}^{-1}$, for $z = 6$, $z = 5$, respectively. However, for redshift $z < 4$, it cannot model the N-body power spectrum at all for the scale smaller than where linear theory works. Unlike the case for the real space matter power spectrum (Figure 3.2), where PT underpredicts the non-linear power spectrum, the redshift space N-body power spectrum always shows the power suppression relative to the perturbation theory calculation. In particular, at $z = 1$, the power suppression occurs even on the very larger scales $k \sim 0.1 h \text{ Mpc}^{-1}$. We attribute this suppression to the small scale velocity dispersion effect, which is called ‘Finger of God’ effect (Jackson, 1972) in the literature. We shall model this effect in the next section. On the other hand, the linear redshift space power spectrum and the non-linear power spectrum of Scoccimarro (2004) are below the N-body redshift space power spectrum for $z > 2$ and $z > 4$, respectively.

We compare the Baryonic Acoustic Oscillations for each redshift in Figure 3.10. As we have already mentioned in section 4.1., we can measure the angular diameter distance and Hubble parameter by using BAO. Especially BAO provide an essential geometrical test of the expansion history of the universe; thus, we can constrain the dark energy equation of state by accurately measuring them.

For redshift higher than 3, we can actually model the baryonic oscillation up to third peak using PT. However, we can only model up to second peak for $z = 2$ due to severe power loss in the larger scale. PT modeling of the baryonic oscillations in $z = 1$ is even worse, and we can only model the first peak and trough. Again, the problem is the small scale power loss.

3.3.3 Modeling Finger of God effect

In the previous section, we show that third order perturbation theory fails to model the non-linear redshift space power spectrum even in the region it models the non-linear real space power spectrum [Fig. (3.8)-(3.10)]. When it fails, we clearly see the power loss of the N-body power spectrum compare to the perturbation theory prediction on the smaller scale at all redshifts.

Such a small scale suppression in redshift space is often called ‘Finger of God’ effect (Jackson, 1972). A folk-lore is that Finger of God only affect the power spectrum on very small scales as the effect is caused by a velocity dispersion inside of virialized halos. However, by comparing against the N-body simulation, we find that the suppression of power happens on fairly large scales in the previous section.

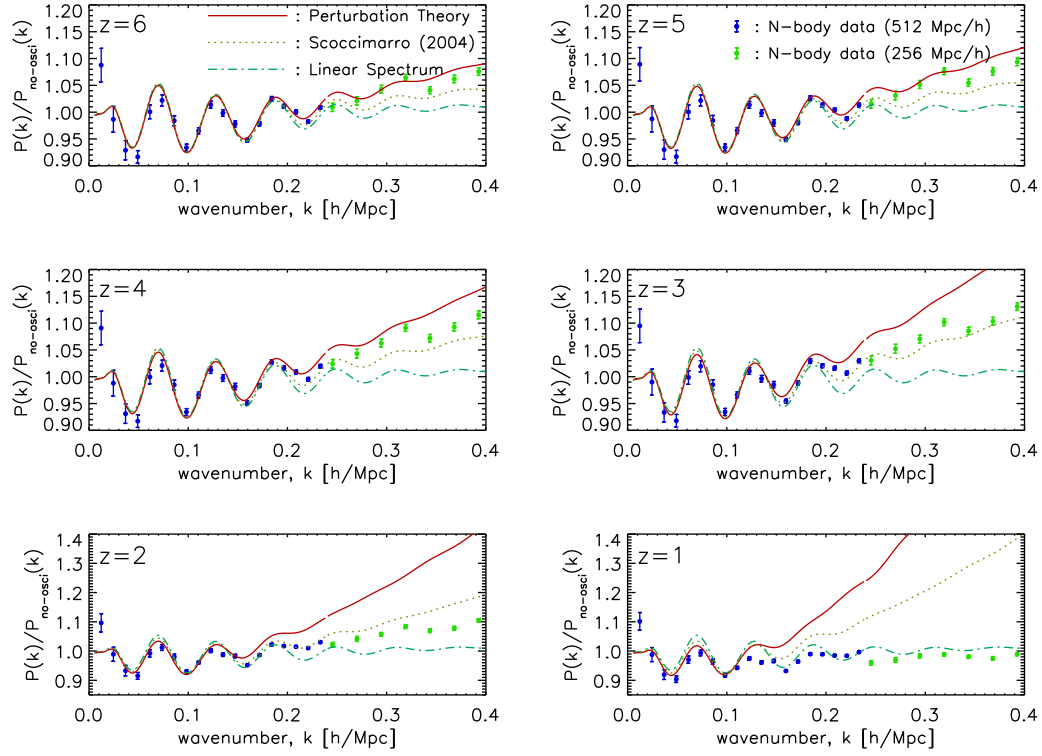


Figure 3.10: Nonlinearity in baryonic acoustic oscillations due to the redshift space distortion. Note that we don't include the Finger of God effect in the modeling power spectrum yet. All the power spectra have been divided by a smooth power spectrum without baryonic oscillations from equation (29) of Eisenstein & Hu (1998). The errorbars show N-body simulations, while solid line show PT calculations. The dot-dashed line show the linear theory predictions, and dotted line show the power spectrum from Scoccimarro's formula. PT describes nonlinear distortion on BAO in the redshift space very accurately at $z > 3$, up to third peak, and at $z = 2$ up to second peak. Although peak position is preserved, at $z = 1$, we lose almost all the information on the amplitude. It clearly shows that we need to model the power suppression. The N-body data at $k < 0.24$ and $> 0.24 \text{ h Mpc}^{-1}$ are from 512 and 256 $h \text{ Mpc}^{-1}$ box simulations, respectively.

What causes the suppression of the power on large scale? As Scoccimarro (2004) points out, it may be related to the velocity dispersion on large scale. As cosmological density field and velocity field are generated from inflation, they are naturally a ‘random’ field. Therefore, even on large scales, velocity field is not solely ‘coherent’, but includes a certain dispersion. In this section, we study the large scale suppression of redshift space power spectrum relating to such velocity dispersions.

Let us first consider the two-point correlation function in redshift space. If the velocity field is perfectly coherent, then the correlation function in real space will be re-mapped onto the corresponding redshift space position; thus, resulting a ‘squashed’ redshift space correlation function (Hamilton, 1998). Now, let us add the dispersion to the velocity field. The main effect of velocity dispersion is simply blurring the real-to-redshift space mapping along the radial direction. If the velocity distribution does not depend on scales, we can simply model it by the convolution of the line of sight directional pairwise velocity distribution function (PVDF) to the ‘coherent’ redshift space correlation function (Ballinger et al., 1996). Schematically,

$$\xi_s^{\text{total}}(s_{\perp}, s'_{\parallel}) = \xi_s^{\text{coherent}}(s_{\perp}, s_{\parallel}) \star \mathcal{P}_v(s'_{\parallel} - s_{\parallel}), \quad (3.36)$$

where s_{\perp} , s_{\parallel} are redshift space position parallel and perpendicular to the line of sight direction, respectively. We use \star symbol for convolution.

In Fourier space, by convolution theorem, the effect of velocity dispersion becomes a multiplication of the Fourier transform of PVDF:

$$P_s^{\text{total}}(k_{\perp}, k_{\parallel}) = P_s^{\text{coherent}}(k_{\perp}, k_{\parallel}) \mathcal{P}_v(k_{\parallel}), \quad (3.37)$$

where k_{\perp} , k_{\parallel} are wavenumber perpendicular and parallel to the line of sight direction. It is $\mathcal{P}_v(k_{\parallel})$ which suppress the power spectrum along the line of sight direction; thus, responsible for the Finger of God effect.

We shall model the Finger of God effect by using an ansatz based on equation (3.37). For a coherent power spectrum, we use the non-linear perturbation theory calculation of the redshift space power spectrum [Eq. (3.24)] as it successfully model the non-linear evolution at high redshifts ($z > 4$). Note that the linear kaiser spectrum and Scoccimarro’s formula are both smaller than the N-body power spectrum; thus multiplying the suppression factor will only make the disagreement larger.

What about $\mathcal{P}_v(k_{\parallel})$? For \mathcal{P}_v , we have to model a Pairwise Velocity Distribution Function (PVDF). There are two widely used models of PVDF in the literature. If dark

matter particles within a halo is thermalized, then PVDF may be Gaussian (Maxwell-Boltzmann); thus the Finger of God suppression factor is given by,

$$\mathcal{P}_v(k_{\parallel}) = \exp(-f^2 k_{\parallel}^2 \sigma_p^2). \quad (3.38)$$

If PVDF follow the exponential distribution, the factor becomes Lorentzian (Fourier transform of exponential) in k-space.

$$\mathcal{P}_v(k_{\parallel}) = \frac{1}{1 + f^2 k_{\parallel}^2 \sigma_p^2}. \quad (3.39)$$

Here, σ_p^2 is the line of sight pairwise velocity dispersion. In the literature, $\sigma_p = \sigma_v/\sqrt{2}$ is widely used where σ_v^2 is the linear velocity dispersion:

$$\sigma_v^2 = \frac{1}{3} \int \frac{d^3 k}{(2\pi)^3} \frac{P(k)}{k^2}. \quad (3.40)$$

Note that these two formulas are the same in the limit where $f^2 k_z^2 \sigma_v^2$ is small. The measured PVDF from simulations, however, turns out to be well approximated by the exponential distribution (e.g. see, Figure 1 of Scoccimarro (2004)). Therefore, we use equation (3.40) for modeling Finger of God effect.

We first use the conventional value of $\sigma_p = \sigma_v/\sqrt{2}$, and compare with N-body power spectrum, but it suppress the power spectrum too much, especially for higher redshifts. By noticing that σ_p is not a value derived from the first principle, but a fitted value to the simulation, we fit σ_p^2 value to the N-body power spectrum in two dimensional Fourier-space (k_{\perp}, k_{\parallel}):

$$P_s^{\text{total}}(k_{\perp}, k_{\parallel}) = P_s^{\text{coherent}}(k_{\perp}, k_{\parallel}) \frac{1}{1 + f^2 k_{\parallel}^2 \sigma_p^2}. \quad (3.41)$$

In Table 3.3.3, we summarize the resulting best-fitting values of σ_p^2 and its 1- σ interval (or 68.27% confidence interval) for six different redshifts ($z = 6, 5, 4, 3, 2$, and 1), and for three different box sizes ($L_{\text{box}} = 512, 256$, and 128 [Mpc/h]). Note that in most of cases, k_{max} 's are set by the resolution of each simulation boxes: $k_{\text{max}} = 0.24, 0.5, 1.4 h \text{ Mpc}^{-1}$ for 512, 256, 128 $h^{-1} \text{ Mpc}$, respectively³. However, when k_{max} set by the simulation box resolution is higher than the k_{max} set by PT (the maximum wavenumber below which PT is proven to be valid in real space (Section 3.2.2)), we choose the latter. We also fit for the highly non-linear regime ($k < 1.4$ for $z = 2$, and $k < 0.5, k < 1.4$ for $z = 1$) in order to show the Finger of God ansatz does not work when non-linearity is too strong,

³For the discussion on the box resolution, see Section 3.1

redshift	k-range [h/Mpc]	$[\sigma_v^2]$ [Mpc/h] ²	$[\sigma_p^2]$ [Mpc/h] ²	ratio $[\sigma_p^2]/[\sigma_v^2]$	χ_{red}^2	d.o.f.
6	$k < 0.24$	1.1530	0.4964 ± 0.1151	0.431	1.102	318
	$k < 0.5$	1.1686	0.1769 ± 0.0279	0.151	1.152	345
	$k < 1.4$	1.1574	0.1009 ± 0.0034	0.0871	1.580	667
5	$k < 0.24$	1.5778	0.6096 ± 0.1156	0.386	1.091	318
	$k < 0.5$	1.5989	0.3013 ± 0.0284	0.188	1.149	345
	$k < 1.4$	1.5832	0.2166 ± 0.0039	0.136	1.502	667
4	$k < 0.24$	2.2427	0.8306 ± 0.1171	0.370	1.086	318
	$k < 0.5$	2.2707	0.5895 ± 0.0294	0.260	1.144	345
	$k < 1.4$	2.2506	0.5155 ± 0.0049	0.229	1.411	667
3	$k < 0.24$	3.5667	1.3945 ± 0.1205	0.391	1.079	318
	$k < 0.5$	3.4785	1.4445 ± 0.0333	0.415	1.155	345
	$k < 1.2$	3.5427	1.5606 ± 0.0118	0.441	1.442	494
2	$k < 0.24$	6.0760	3.4408 ± 0.1338	0.566	1.144	318
	$k < 0.33$	6.1519	4.2194 ± 0.1553	0.686	1.053	154
	$k < 1.4$	6.0887	5.0000 ± 0.0167	0.821	2.431	667
1	$k < 0.15$	12.8654	10.2650 ± 0.8443	0.798	1.149	131
	$k < 0.5$	12.6851	19.8754 ± 0.0975	1.567	2.292	345
	$k < 1.4$	12.6543	23.8262 ± 0.0598	1.883	10.335	667

Table 3.1: The best-fit $[\sigma_p^2]$ parameter for 6 different redshifts and 3 different k_{max} 's for each redshift. Note that in most of cases, k_{max} 's are set by the resolution of each simulation boxes : $k_{max} = 0.24, 0.5, 1.4 \text{ h Mpc}^{-1}$ for 512, 256, 128 $h^{-1} \text{ Mpc}$, respectively. However, when k_{max} set by the resolution is higher than the k_{max} we get from the valid k region of the nonlinear PT in real space, we choose the later. The higher values of χ_{red}^2 in the case of $k < 1.4$ at $z = 2$, and $k < 0.5, k < 1.4$ at $z = 1$ show the failure of the fitting in the highly nonlinear region. Note that $[\sigma_v^2]$ from equation (3.40) is always larger than the best-fit value when perturbation theory ansatz give rise to the small χ_{red}^2 .

We notice that the measured σ_p^2 value on largest scales, $k < 0.24 [h/\text{Mpc}]$, are systematically larger than the σ_p^2 from smaller scale for $z > 4$, while this trend reverses for lower redshifts. The latter cases (for lower redshifts $z \leq 4$) are reasonable, as we resolve more particle per clump in the high resolution simulation. On the other hand, we think the former cases (for higher redshifts $z > 4$) is an artifact, and it is simply a result from that changing power spectrum on large scale requires larger σ_p^2 . For example, we can see how small the Finger of God effect at $z = 6$ on large scale (left top panel of Figure 3.11).

We test the goodness of fitting by calculating the reduced- χ^2 (χ_{red}^2). The reduced χ^2 values are close to 1 in the k-region where PT works well, but they become larger in the strongly nonlinear region at lower redshifts. Thus, the FoG prescription based on the exponential Pairwise Velocity Distribution Function can be an effective modeling of the Finger of God effect. We also find that the conventional velocity dispersion values $\sigma_p^2 = \sigma_v^2/2$ are always larger than the best-fitting σ_p^2 values when χ_{red}^2 is close to unity; the conventional Finger of God ansatz over-suppress the power.

From Figure 3.11 to Figure 3.13, we show the contour of three 2-dimensional power spectra: power spectrum from N-body simulation (blue, solid line), power spectrum from non-linear PT (red, dot-dashed line), and power spectrum from non-linear PT multiplied by Finger of God ansatz (red, solid line). Figure 3.11 shows the result from the $512 h^{-1}$ Mpc box simulation, and we fit only for $|k| < 0.24 h \text{ Mpc}^{-1}$, except for $z = 1$, where we use $k_{max} = 0.15 h \text{ Mpc}^{-1}$ which is the maximum wavenumber determined from real space power spectrum. On such large scales, Finger of God does not affect at high redshifts ($z > 3$), and we start to see the difference from $z = 3$. Figure 3.12 shows the result from the $256 h^{-1}$ Mpc box simulation, and fitting is done for $|k| < 0.5 h \text{ Mpc}^{-1}$, except for $z = 2$ ($k_{max} = 0.33 h \text{ Mpc}^{-1}$). The Finger of God effect (difference between dot-dashed line and blue solid line) is large for low redshifts ($z \lesssim 2$) that we start to see the contour being ‘squashed’. Figure 3.13 shows the result from the $128 h^{-1}$ Mpc box simulation, and fitting is done for $|k| < 1.4 h \text{ Mpc}^{-1}$, except for $z = 3$ ($k_{max} = 1.2 h \text{ Mpc}^{-1}$). For $z = 1$, the Finger of God effect is so strong that its suppression dominates over the non-linear power spectrum from PT.

How does angular averaged power spectrum look now? In Figure 3.14, we compare four angular averaged power spectra: the N-body power spectrum (blue, dashed line), the linear redshift space power spectrum (green, dot-dashed line), the non-linear power spectrum from Full third order theory (red, dots-dashed line), and the non-linear PT + Lorentzian Finger of God model. We find that the agreement between the non-linear power spectrum

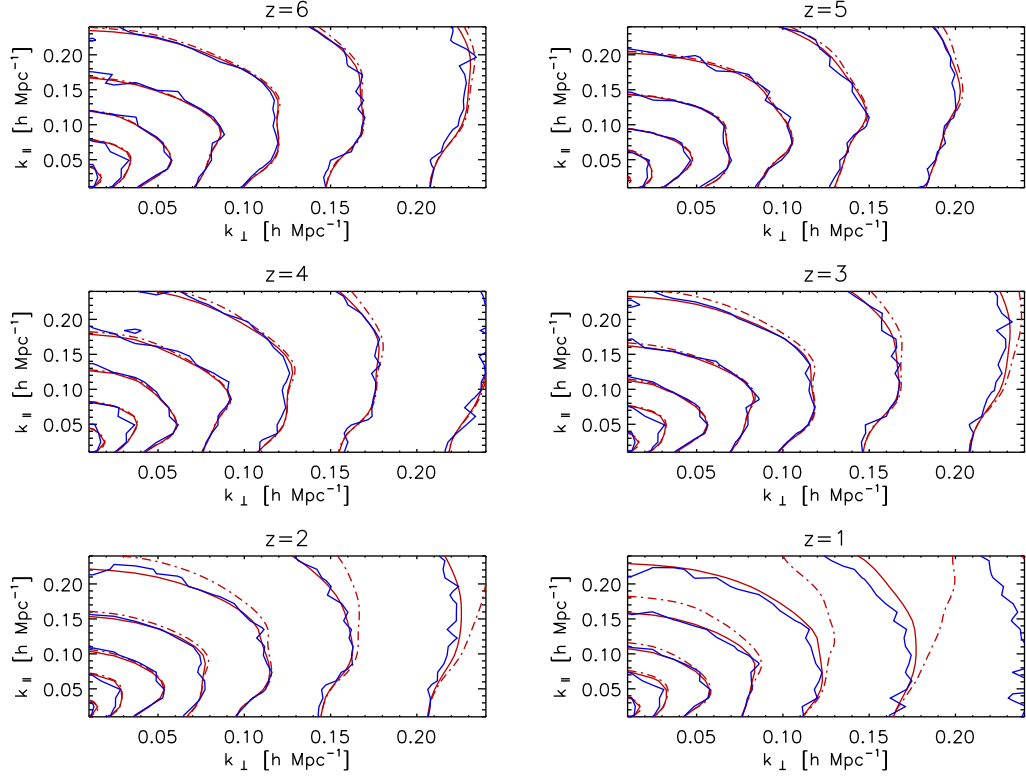


Figure 3.11: Result of 2-dimensional fitting of σ_p^2 using $512 h^{-1}\text{Mpc}$ box. N-body power spectrum (blue line) in redshift space, PT power spectrum + Finger of God ansatz (red line), and PT power spectrum (red dot-dashed line) is shown. Note that we fit only for $|k| < 0.24 h \text{ Mpc}^{-1}$, except for $z = 1$, where we use $k_{max} = 0.15 h \text{ Mpc}^{-1}$ which is the maximum wavenumber below which PT works well [Section 3.2.2]

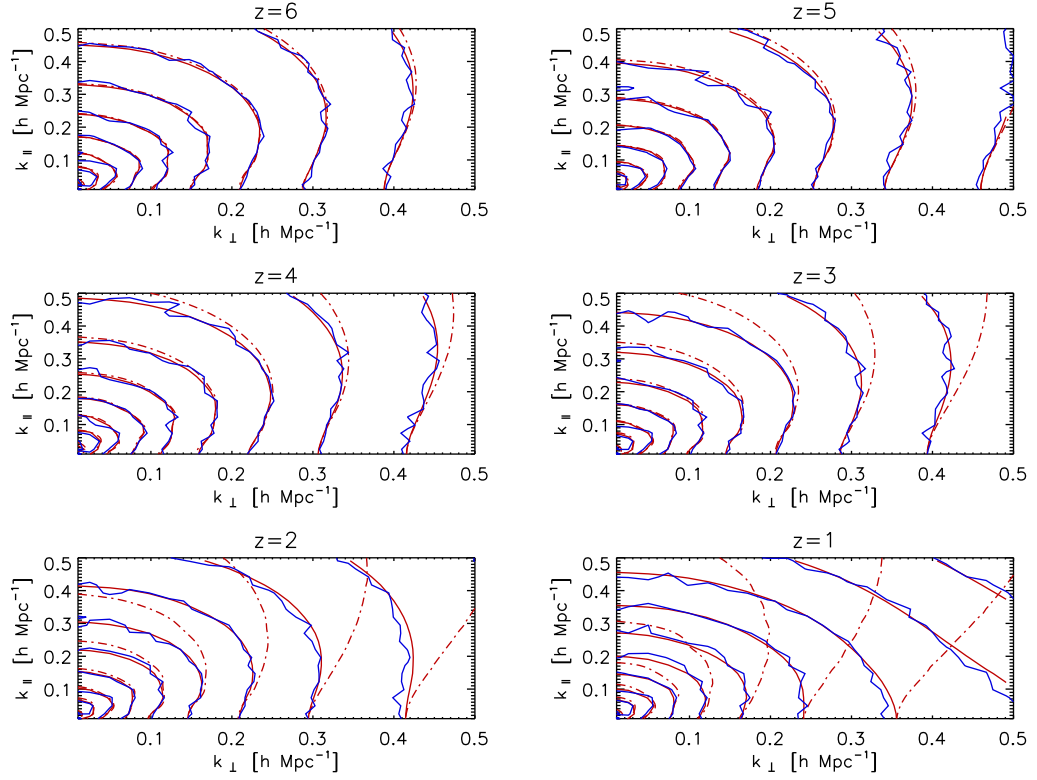


Figure 3.12: Same as Figure 3.11, but for 256 $h^{-1}\text{Mpc}$ box. Note that we fit only for $|k| < 0.5 h \text{ Mpc}^{-1}$, except for $z = 2$ ($k_{max} = 0.33 h \text{ Mpc}^{-1}$).

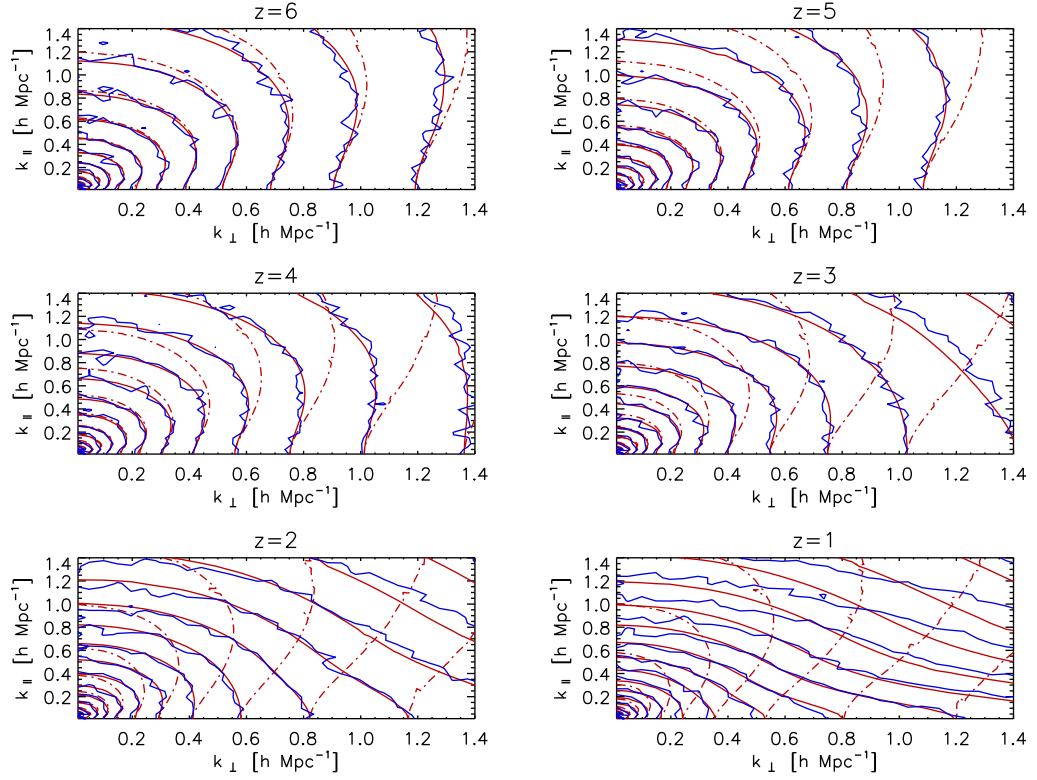


Figure 3.13: Same as Figure 3.11, but for 128 $h^{-1}\text{Mpc}$ box. Note that we fit only for $|k| < 1.4 h \text{ Mpc}^{-1}$, except for $z = 3$ ($k_{max} = 1.2 h \text{ Mpc}^{-1}$).

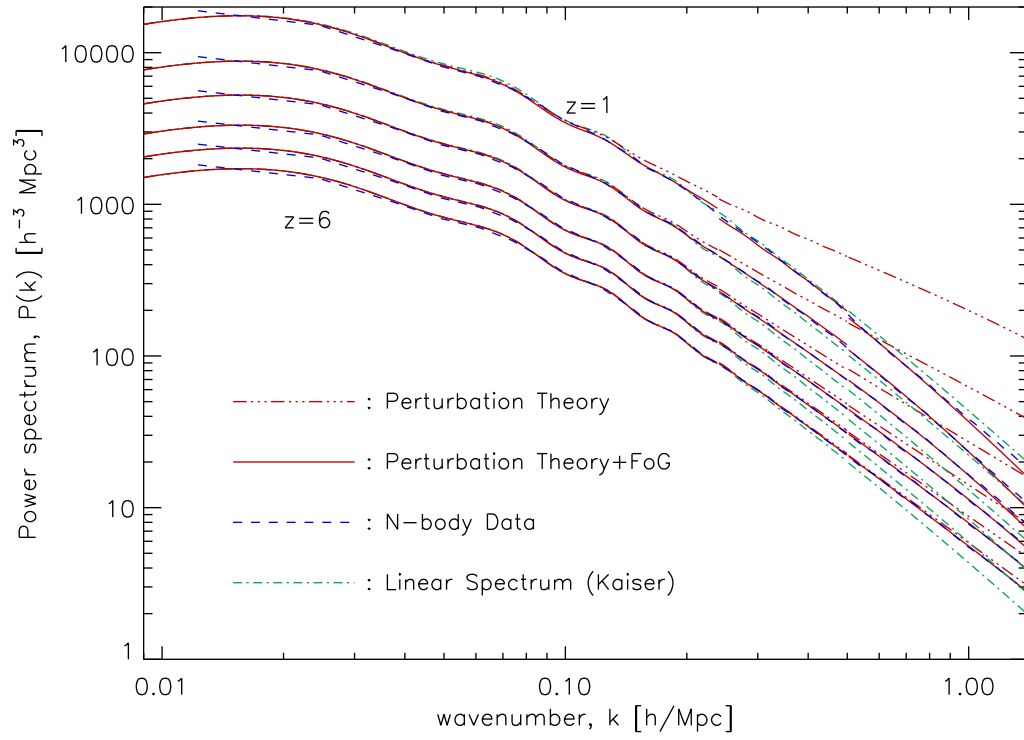


Figure 3.14: Same as Figure 3.8, but with the PT+Finger of God fitting result (red, solid line). Finger of God modeling improves the agreement between PT and N-body power spectrum dramatically.

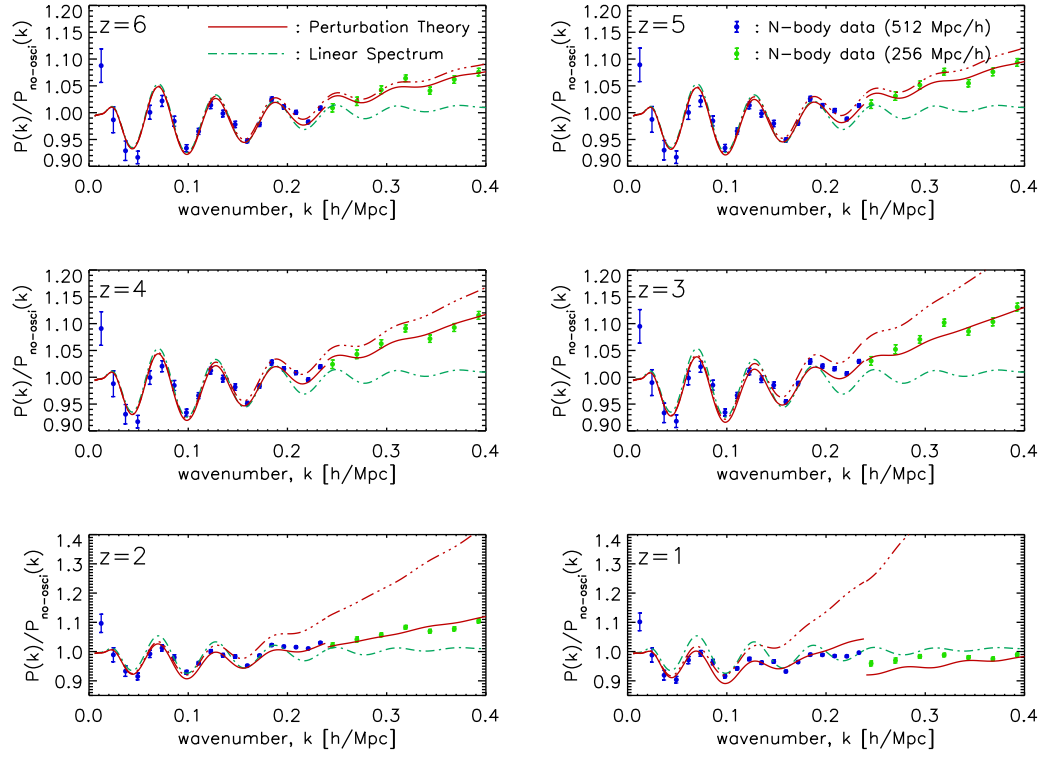


Figure 3.15: FoG effect on the angular averaged baryonic oscillations. This figure is the same as Figure 3.10, except for the PT+FoG plot (solid line). Note that the overall suppression of baryonic oscillation at $z = 1$ is modeled quite correctly up to the second peak.

and the N-body power spectrum is improved dramatically by including the Finger of God ansatz.

We also show the angular averaged baryonic oscillations in Figure 3.15. This figure is the same as Figure 3.10, except for one more line: the non-linear perturbation theory + Finger of God ansatz. By comparing PT+FoG with N-body spectrum, we see that the overall suppression of baryonic oscillation at $z = 1$ is modeled quite correctly up to the second peak. Note that, for low redshifts, there is a discontinuity around $k \sim 0.24 h \text{ Mpc}^{-1}$. This may suggest that the Finger of God effect also depends on the resolution of the simulation in a way that the suppression is larger for the higher resolution simulation.

Although the Finger of God modeling of equation (3.41) fits the observed power spectrum, there is a few caveats. First, the decomposition of ‘coherent’ and ‘dispersion’ part is not clear in reality. As a non-linear perturbative solution of the exact relation [Eq. 3.23], the non-linear redshift space power spectrum from PT may also include the term which is responsible for the velocity dispersion. Second, we assume that the Pairwise Velocity Distribution Function (PVDF) is constant over the scales in which we are interested. However, the measured PVDF from simulation does depend on scales, and in fact, the shape is quite different for different scales. For the scale dependent PVDF, the simple convolution does not work. Instead, we may have to model the *Full* PVDF, and use so called *Streaming moel* (Scoccimarro, 2004), which relate the correlation function in real space and redshift space by

$$1 + \xi_s(s_\perp, s_\parallel) = \int_{-\infty}^{\infty} dr_\perp [1 + \xi(r)] \mathcal{P}(r_\perp - s_\perp, \mathbf{r}), \quad (3.42)$$

where $\mathcal{P}(r_\perp - s_\perp, \mathbf{r})$ is the scale (which is parametrized by \mathbf{r}) dependent PVDF.

3.4 Discussion and Conclusions

The next-to-leading order correction to the matter power spectrum calculated analytically from 3rd-order PT provides an almost exact description of the matter power spectrum in real space in the weakly non-linear regime, where $\Delta^2(k) \lesssim 0.4$ (Fig. 3.2). The next-to-leading order correction to the matter power spectrum in redshift space is also analytic, but models only limited range in k -space at very high redshifts $z \gtrsim 5$. However, by using an empirically motivated, exponential Pairwise Velocity Distribution Function and fitting to the effective velocity dispersion parameter σ_p^2 , we find that the theory provides a reasonably good fitting to N-body data for the wavenumber ranges where PT models well for the real space matter power spectrum.

The most important implications of our results for the planned high- z galaxy surveys are that we can use PT to calculate (a) non-linearity in the baryonic acoustic oscillations (Fig. 3.4), which should reduce systematics in constraining dark energy properties, and (b) the matter power spectrum up to much higher k than that was accessible before, which should vastly increase our ability to measure the shape of the primordial power spectrum as well as the mass of neutrinos (Takada et al., 2006).

Of course, these surveys measure the *galaxy* power spectrum in *redshift* space; thus, the future work should include PT calculations of non-linearity in redshift space distortion, and halo biasing (Fry & Gaztanaga, 1993; Heavens et al., 1998) at the same time, as well as an extensive comparison with numerical simulations. PT also allows one to calculate the higher-order statistics such as the bispectrum, which has been shown to be a powerful tool to check for systematics in our understanding of non-linear galaxy bias (Matarrese et al., 1997; Verde et al., 1998). We should therefore “reload” cosmological perturbation theory and make a serious assessment of its validity in light of the planned high- z galaxy surveys constraining properties of dark energy, inflation, and neutrinos.

Chapter 4

Perturbation Theory Reloaded II: Non-linear Bias, Baryon Acoustic Oscillations and Millennium Simulation In Real Space

In the previous chapter, we show that non-linear Eulerian perturbation theory can model the non-linear matter power spectrum in both real and redshift space. In this section, we shall extend the analysis into the galaxy bias, and show that non-linear perturbation theory combined with the locality of bias ansatz can model the non-linear power spectrum of biased tracers¹.

4.1 Non-linear galaxy power spectrum from perturbation theory

4.1.1 Locality Assumption

Galaxies are biased tracers of the underlying density field (Kaiser, 1984), which implies that the distribution of galaxies depends on the underlying matter density fluctuations in a complex way. This relation must depend upon the detailed galaxy formation processes, which are not yet understood completely.

However, on large enough scales, one may approximate this function as a local function of the underlying density fluctuations, i.e., the number density of galaxies at a given position in the universe is given solely by the underlying matter density at the same position. With this approximation, one may expand the density fluctuations of galaxies, δ_g , in terms of the underlying matter density fluctuations, as (Fry & Gaztanaga, 1993; McDonald, 2006)

$$\delta_g(\mathbf{x}) = \epsilon + c_1\delta(\mathbf{x}) + \frac{1}{2}c_2\delta^2(\mathbf{x}) + \frac{1}{6}c_3\delta^3(\mathbf{x}) + \dots, \quad (4.1)$$

¹Previous version of this chapter was published in Jeong, D. & Komatsu, E. 2009 *Astrophys. J.*, 691, 569.

where c_n are the galaxy bias parameters, and ϵ is a random variable that represents the “stochasticity” of the galaxy bias, i.e., the relation between $\delta_g(\mathbf{x})$ and $\delta(\mathbf{x})$ is not deterministic, but contains some noise (e.g., Yoshikawa et al., 2001, and references therein). We assume that the stochasticity is white noise, and is uncorrelated with the density fluctuations, i.e., $\langle \epsilon \delta \rangle = 0$. While both of these assumptions should be violated at some small scales, we assume that these are valid assumptions on the scales that we are interested in – namely, on the scales where the 3rd-order PT describes the non-linear matter power spectrum with 1% accuracy. Since both bias parameters and stochasticity evolve in time (Fry, 1996; Tegmark & Peebles, 1998), we allow them to depend on redshifts.

One obtains the traditional “linear bias model” when the Taylor series expansion given in Eq. (4.1) is truncated at the first order and the stochasticity is ignored.

The precise values of the galaxy bias parameters depend on the galaxy formation processes, and different types of galaxies have different galaxy bias parameters. However, we are *not* interested in the precise values of the galaxy bias parameters, but only interested in extracting cosmological parameters from the observed galaxy power spectra with *all the bias parameters marginalized over*.

4.1.2 3rd-order PT galaxy power spectrum

The analysis in this paper adopts the framework of McDonald (2006), and we briefly summarize the result for clarity. We shall use the 3rd-order PT; thus, we shall keep the terms up to the 3rd order in δ . The resulting power spectrum can be written in terms of the linear matter power spectrum, $P_L(k)$, and the 3rd order matter power spectrum, $P_{\delta\delta}(k)$, as

$$P_g(k) = P_0 + b_1^2 \left[P_{\delta\delta}(k) + b_2 P_{b_2}(k) + b_2^2 P_{b_{22}}(k) \right], \quad (4.2)$$

where P_{b_2} and $P_{b_{22}}$ are given by

$$P_{b_2} = 2 \int \frac{d^3 \mathbf{q}}{(2\pi)^3} P_L(q) P_L(|\mathbf{k} - \mathbf{q}|) F_2^{(s)}(\mathbf{q}, \mathbf{k} - \mathbf{q}),$$

and

$$P_{b_{22}} = \frac{1}{2} \int \frac{d^3 \mathbf{q}}{(2\pi)^3} P_L(q) \left[P_L(|\mathbf{k} - \mathbf{q}|) - P(q) \right],$$

respectively, with $F_2^{(2)}$ given by

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

We use the standard formula for $P_{\delta\delta}$ (see Eq. (14) of Paper I and references therein). Here, b_1 , b_2 , and P_0 are the non-linear bias parameters², which are given in terms of the original coefficients for the Taylor expansion as

$$\begin{aligned} b_1^2 &= c_1^2 + c_1 c_3 \sigma^2 + \frac{68}{21} c_1 c_2 \sigma^2, \\ b_2 &= \frac{c_2}{b_1}, \\ P_0 &= \langle \epsilon^2 \rangle + \frac{1}{2} c_2^2 \int \frac{k^2 dk}{2\pi^2} P_L^2(k), \end{aligned} \tag{4.3}$$

where σ is the r.m.s. of density fluctuations.

We will never have to deal with the original coefficients, c_1 , c_2 , c_3 , or ϵ .³ Instead, we will only use the re-parametrized bias parameters, b_1 , b_2 , and P_0 , as these are related more directly to the observables. As shown by McDonald (2006), in the large-scale limit, $k \rightarrow 0$, one finds

$$P_g(k) \rightarrow P_0 + b_1^2 P_L(k). \tag{4.4}$$

Therefore, in the large-scale limit one recovers the traditional linear bias model plus the constant term. Note that b_1 is the same as what is called the “effective bias” in Heavens et al. (1998).

Throughout this paper we shall use Eq. (4.2) for calculating the non-linear galaxy power spectra.

4.1.3 Why we do not care about the precise values of bias parameters

The precise values of the galaxy bias parameters depend on the details of the galaxy formation and evolution, as well as on galaxy types, luminosities, and so on.

However, our goal is to extract the cosmological information from the observed galaxy power spectra, without having to worry about which galaxies we are using as tracers of the underlying density field.

Therefore, we will marginalize the likelihood function over the bias parameters, without ever paying attention to their precise values. Is this approach sensible?

²These parameters correspond to b_1 , b_2 , and N in the original paper by McDonald (2006).

³For the expression of $P_g(k)$ with the original coefficients, see equation (2.134). It is also shown in literature such as Heavens et al. (1998); Smith et al. (2007).

One might hope that one should be able to calculate the bias parameters for given properties of galaxies from the first principles using, e.g., sophisticated numerical simulations.

Less numerically expensive way of doing the same thing would be to use the semi-analytical halo model approach, calibrated with a smaller set of numerical simulations (see Cooray & Sheth, 2002, for a review). Using the peak-background split method (Sheth & Tormen, 1999) based upon the excursion set approach (Bond et al., 1991), one can calculate b_1 , b_2 , b_3 , etc., the coefficients of the Taylor series expansion given in Eq. (4.1), for the density of dark matter halos. Once the bias parameters for dark matter halos are specified, the galaxy bias parameters may be calculated using the so-called Halo Occupation Distribution (HOD) (Seljak, 2000).

Smith et al. (2007) have attempted this approach, and shown that it is difficult to calculate even the power spectrum of dark matter halos that matches N -body simulations. The halo-model predictions for bias parameters are not yet accurate enough, and we do not yet have a correct model for P_0 .

The situation would be even worse for the galaxy power spectrum, as we would have to model the HOD in addition to the halo bias. At the moment the form of HOD is basically a free empirical function. We therefore feel that it is dangerous to rely on our limited understanding of these complications for computing the bias parameters.

This is the reason why we have decided to give up predicting the precise values of bias parameters entirely. Instead, we shall treat 3 bias parameters, b_1 , b_2 , and P_0 , as free parameters, and fit them to the observed galaxy power spectra simultaneously with the cosmological parameters.

The most important question that we must ask is the following, “using the 3rd-order PT with 3 bias parameters, can we extract the correct cosmological parameters from the galaxy power spectra?” If the answer is yes, we will not have to worry about the precise values of bias parameters anymore.

4.2 Dark Matter Power spectrum from Millennium Simulation

In this section we show that the matter power spectrum computed from the 3rd-order PT agrees with that estimated from the Millennium Simulation (Springel et al., 2005). This result confirms our previous finding (Paper I).

Using the result obtained in this section we define the maximum wavenumber, k_{max} , below which the 3rd-order PT may be trusted. The matter power spectrum gives an unambiguous definition of k_{max} , which will then be used thereafter when we analyze power spectra of halos and galaxies in § 4.3.

4.2.1 Millennium Simulation

The Millennium simulation (Springel et al., 2005) is a large N -body simulation with the box size of $(500 \text{ Mpc}/h)^3$ and 2160^3 dark matter particles. The cosmological parameters used in the simulation are $(\Omega_{dm}, \Omega_b, \Omega_\Lambda, h) = (0.205, 0.045, 0.75, 0.73)$.

The primordial power spectrum used in the simulation is the scale-invariant Peebles-Harrison-Zel'dovich spectrum, $n_s = 1.0$, and the linear r.m.s. density fluctuation smoothed with a top-hat filter of radius $8 h^{-1} \text{Mpc}$ is $\sigma_8 = 0.9$. Note that these values are significantly larger than the latest values found from the WMAP 5-year data, $\sigma_8 \simeq 0.8$ and $n_s \simeq 0.96$ (Dunkley et al., 2009; Komatsu et al., 2010), which implies that non-linearities in the Millennium Simulation should be stronger than those in our Universe.

The Millennium Simulation was carried out using the **GADGET** code (Springel et al., 2001; Springel, 2005). The **GADGET** uses the tree Particle Mesh (tree-PM) gravity solver, which tends to have a larger dynamic range than the traditional PM solver for the same box size and the same number of particles (and meshes) (Heitmann et al., 2007). Therefore, the matter power spectrum from the Millennium Simulation does not suffer from an artificial suppression of power as much as those from the PM codes.

The initial particle distribution was generated at the initial redshift of $z_{ini} = 127$ using the standard Zel'dovich approximation. While the initial conditions generated from the standard Zel'dovich approximation tend to produce an artificial suppression of power at later times, and the higher-order scheme such as the second-order Lagrangian perturbation theory usually produces better results (Scoccimarro, 1998; Crocce et al., 2006), the initial redshift of the Millennium Simulation, $z_{ini} = 127$, is reasonably high for the resulting power spectra to have converged in the weakly non-linear regime.

The mass of each dark matter particle in the simulation is $M_{dm} = 8.6 \times 10^8 M_\odot/h$. They require at least 20 particles per halo for their halo finder, and thus the minimum mass resolution of halos is given by $M_{halo} \geq 20 M_{dm} \simeq 1.7 \times 10^{10} M_\odot/h$. Therefore, the Millennium Simulation covers the mass range that is relevant to real galaxy surveys that would detect galaxies with masses in the range of $M \simeq 10^{11} - 10^{12} M_\odot$. This property

distinguishes our study from the previous studies on non-linear distortion of BAOs due to galaxy bias (e.g., Smith et al., 2007; Huff et al., 2007), whose mass resolution was greater than $\sim 10^{12} M_{\odot}$.

In addition to the dark matter halos, the Millennium database⁴ also provides galaxy catalogues from two different semi-analytic galaxy formation models (De Lucia & Blaizot, 2007; Croton et al., 2006; Bower et al., 2006; Benson et al., 2003; Cole et al., 2000). These catalogues give us an excellent opportunity for testing validity of the non-linear galaxy power spectrum model based upon the 3rd-order PT with the unprecedented precision.

4.2.2 3rd-order PT versus Millennium Simulation: Dark Matter Power Spectrum

First, we compare the matter power spectrum from the Millennium simulation with the 3rd-order PT calculation. The matter power spectrum we use here was measured directly from the Millennium simulation on the fly.⁵

Table 4.1: Maximum wavenumbers, k_{max} , for the Millennium Simulation

z	k_{max} [h/Mpc]	\tilde{k}_{max} [h/Mpc]
6	1.5	1.99
5	1.3	1.37
4	1.2	1.02
3	1.0	0.60
2	0.25	0.35
1	0.15	0.20

z : redshift

k_{max} : the maximum wavenumber for the simulated $P_m(k)$ to agree with the PT calculation at 2% accuracy within the statistical error of the Millennium Simulation

\tilde{k}_{max} : \tilde{k}_{max} is defined by $\Delta_m^2(\tilde{k}_{max}) = 0.4$ which is the criteria recommended in Paper I.

Figure 4.1 shows the matter power spectrum from the Millennium simulation (dashed lines), the 3rd-order PT calculation (solid lines), and the linear PT (dot-dashed lines) for seven different redshifts, $z = 0, 1, 2, 3, 4, 5$, and 6. The analytical calculation of the 3rd-order PT reproduces the non-linear matter power spectrum from the Millennium Simulation accurately at high redshifts, i.e., $z > 1$, up to certain maximum wavenumbers, k_{max} ,

⁴<http://www.g-vo.org/MyMillennium2/>

⁵We thank Volker Springel for providing us with the matter power spectrum data.

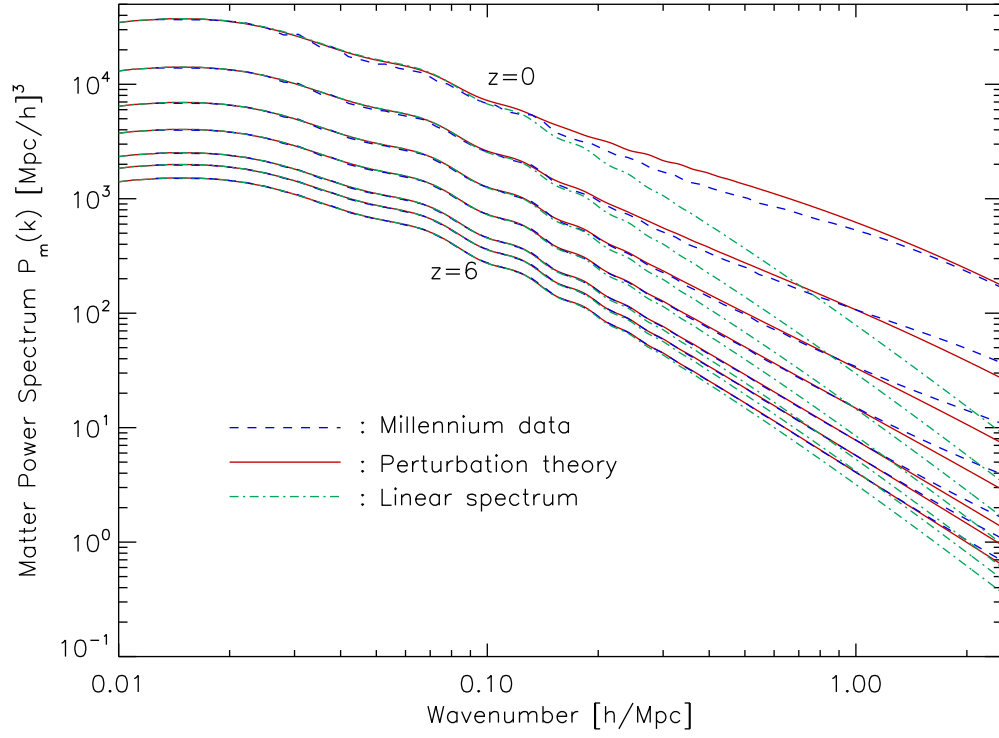


Figure 4.1: Matter power spectrum at $z = 0, 1, 2, 3, 4, 5$ and 6 (*from top to bottom*) derived from the Millennium Simulation (*dashed lines*), the 3rd-order PT (*solid lines*), and the linear PT (*dot-dashed lines*).

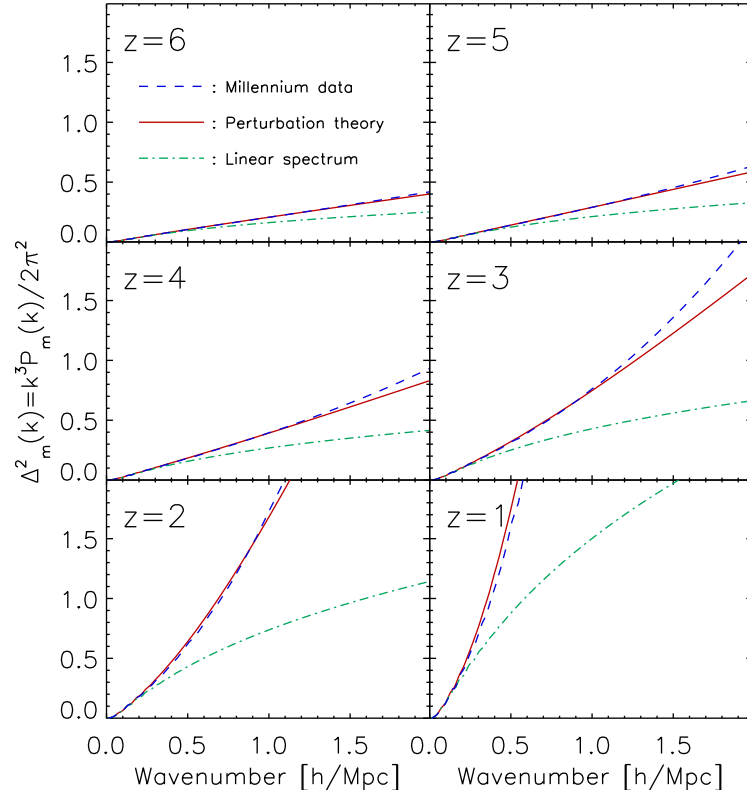


Figure 4.2: Dimensionless matter power spectrum, $\Delta^2(k)$, at $z = 1, 2, 3, 4, 5$, and 6 . The dashed and solid lines show the Millennium Simulation data and the 3rd-order PT calculation, respectively. The dot-dashed lines show the linear power spectrum.

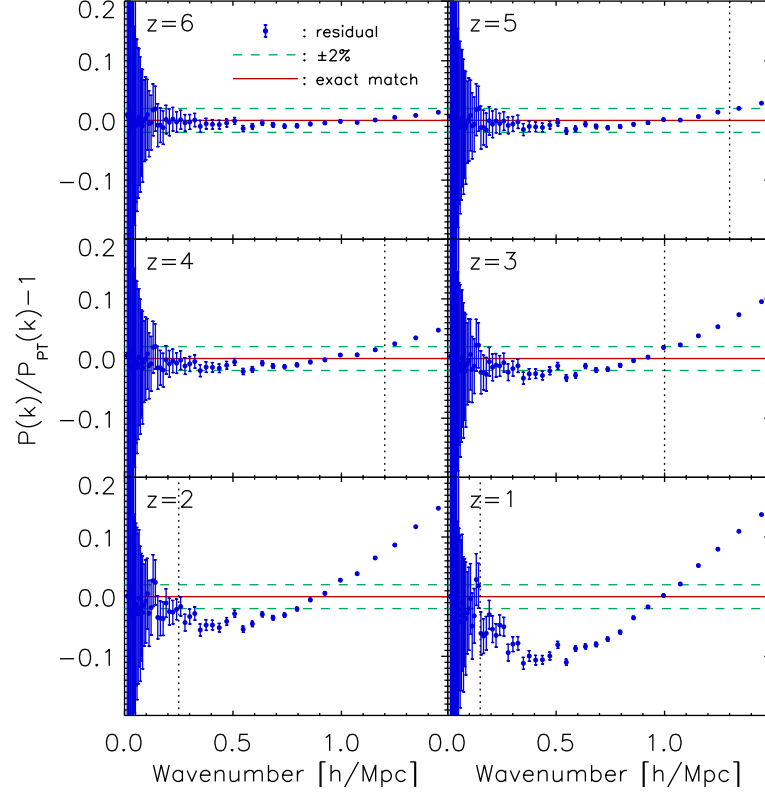


Figure 4.3: Fractional difference between the matter power spectra from the 3rd-order PT and that from the Millennium Simulation, $P_m^{sim}(k)/P_m^{PT} - 1$ (dots with errorbars). The solid lines show the perfect match, while the dashed lines show $\pm 2\%$ accuracy. We also show $k_{max}(z)$, below which we trust the prediction from the 3rd-order PT, as a vertical dotted line.

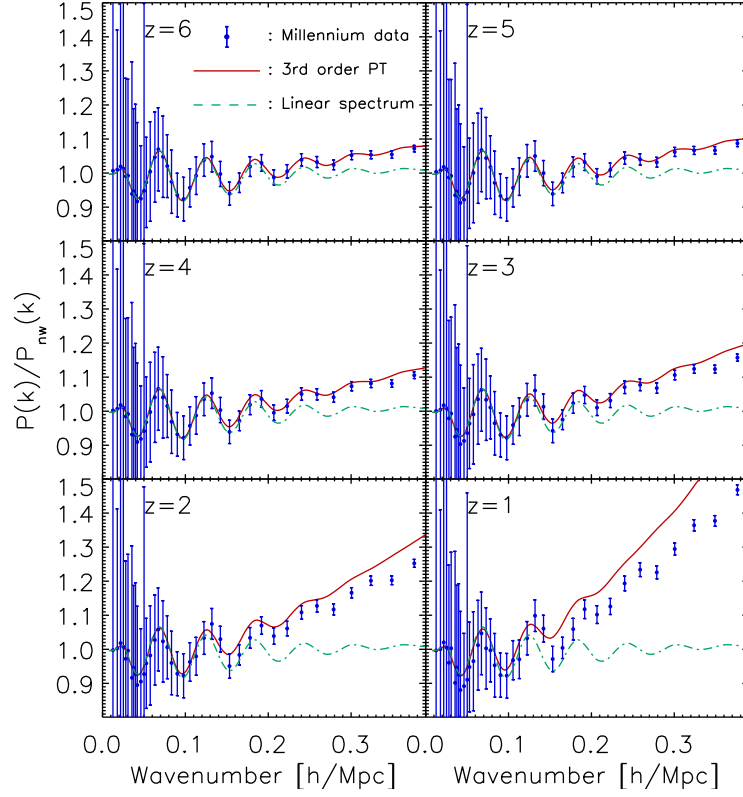


Figure 4.4: Distortion of BAOs due to non-linear matter clustering. All of the power spectra have been divided by a smooth power spectrum without baryonic oscillations from eq. (29) of Eisenstein & Hu (1998). The error bars show the simulation data, while the solid lines show the PT calculations. The dot-dashed lines show the linear theory calculations. The power spectrum data shown here have been taken from Figure 6 of Springel et al. (2005).

that will be specified below. To facilitate the comparison better, we show the dimensionless matter power spectrum, $\Delta_m^2(k) \equiv k^3 P_m(k)/2\pi^2$, in Figure 4.2.

We find the maximum wavenumber, $k_{max}(z)$, below which we trust the prediction from the 3rd-order PT, by comparing the matter power spectrum from PT and the Millennium Simulation. The values of k_{max} found here will be used later when we analyze the halo/galaxy power spectra.

In Paper I we have defined k_{max} such that the fractional difference between PT and the average of ~ 100 simulations is 1%. Here, we have only one realization, and thus the results are subject to statistical fluctuations that might be peculiar to this particular realization. Therefore, we relax our criteria for k_{max} : we define k_{max} such that the fractional difference between PT and the Millennium Simulation is 2%.

Figure 4.3 shows the fractional differences at $z = 1, 2, 3, 4, 5$, and 6. Since we have only one realization, we cannot compute statistical errors from the standard deviation of multiple realizations. Therefore, we derive errors from the leading-order 4-point function assuming Gaussianity of the underlying density fluctuations (see Appendix F.1), $\sigma_{P(k)} = P(k)/\sqrt{N_k}$, where N_k is the number of independent Fourier modes per bin at a given k shown in Figure 4.3.

We give the values of k_{max} in Table 4.1. We shall use these values when we fit the halo/galaxy power spectrum in the next section. Note that k_{max} decreases rapidly below $z = 2$. It is because $P(k)/P_{PT}(k) - 1$ is not a monotonic function of k . The dip in $P(k)/P_{PT}(k) - 1$ is larger than 2% at lower redshift, $z < 2$, while it is inside of the 2% range at $z \geq 3$. Therefore, our criteria of 2% make that sudden change. This feature is due to the limitation of the standard 3rd order PT. However, we can remove this feature by using the improved perturbation theory, e.g. using renormalization group techniques. (See, Figure 9 of Matarrese & Pietroni (2007).)

We also give the values of \tilde{k}_{max} , for which $\Delta_m^2(\tilde{k}_{max}) = 0.4$ (criteria recommended in Paper I). The difference between k_{max} and \tilde{k}_{max} is probably due to the fact that we have only one realization of the Millennium Simulation, and thus estimation of k_{max} is noisier. Note that the values of \tilde{k}_{max} given in Table 4.1 are smaller than those given in Paper I. This is simply because σ_8 of the Millennium Simulation ($\sigma_8 = 0.9$) is larger than that of Paper I ($\sigma_8 = 0.8$).

In Figure 4.4 we show the matter power spectra divided a smooth spectra without BAOs (Eq. (29) of Eisenstein & Hu, 1998). The results are consistent with what we have

found in Paper I: although BAOs in the matter power spectrum are distorted heavily by non-linear evolution of matter fluctuations, the analytical predictions from the 3rd-order PT capture the distortions very well at high redshifts, $z > 2$.

At lower redshifts, $z \sim 1$, the 3rd-order PT is clearly insufficient, and one needs to go beyond the standard PT. This is a subject of recent studies (Crocce & Scoccimarro, 2008; Matarrese & Pietroni, 2007; Taruya & Hiramatsu, 2008; Valageas, 2007; Matsubara, 2008; McDonald, 2007).

4.3 HALO/GALAXY POWER SPECTRUM AND THE NON-LINEAR BIAS MODEL

In this section we compare the 3rd-order PT galaxy power spectrum with the power spectra of dark matter halos and galaxies estimated from the Millennium Simulation. After briefly describing the analysis method in § 4.3.1, we analyze the halo bias and galaxy bias in § 4.3.2 and § 4.3.3, respectively. We then study the dependence of bias parameters on halo/galaxy mass in § 4.3.4.

4.3.1 Analysis method

Table 4.2: Summary of six snapshots from the Millennium Simulation

z	z_{show}	N_h	$1/n_h$ ([Mpc/h] ³)	N_{Mg}	$1/n_{Mg}$ ([Mpc/h] ³)	N_{Dg}	$1/n_{Dg}$ ([Mpc/h] ³)
5.724	6	5,741,720	21.770	6,267,471	19.944	4,562,368	27.398
4.888	5	8,599,981	14.535	9,724,669	12.854	7,604,063	16.439
4.179	4	11,338,698	11.024	13,272,933	9.418	10,960,404	11.405
3.060	3	15,449,221	8.091	19,325,842	6.468	17,238,935	7.251
2.070	2	17,930,143	6.972	23,885,840	5.233	22,962,129	5.444
1.078	1	18,580,497	6.727	26,359,329	4.742	27,615,058	4.527

z : the exact redshift of each snapshot

z_{show} : the redshift we quote in this paper

N_h : the number of MPA halos in each snapshot; $1/n_h$: the corresponding Poisson shot noise

N_{Mg} : the number of MPA galaxies in each snapshot; $1/n_{Mg}$: the corresponding Poisson shot noise

N_{Dg} : the number of Durham galaxies in each snapshot; $1/n_{Dg}$: the corresponding Poisson shot noise

We choose six redshifts between $1 \leq z \leq 6$ from 63 snapshots of the Millennium Simulation, and use all the available catalog of halos (MPA Halo (MHalo), hereafter ‘halo’) and two galaxy catalogues (MPA Galaxies, hereafter ‘Mgalaxy’; Durham Galaxies, hereafter

‘Dgalaxy’) at each redshift. The exact values of redshifts and the other relevant information of chosen snapshots are summarized in Table 4.2.

Halos are the groups of matter particles found directly from the Millennium Simulation. First, the dark matter groups (called FOF group) are identified by using Friends-of-Friends (FoF) algorithm with a linking length equal to 0.2 of the mean particle separation. Then, each FoF group is divided into the gravitationally bound local overdense regions, which we call halos here.

Mgalaxies and Dgalaxies are the galaxies assigned to the halos using two different semi-analytic galaxy formation codes: L-Galaxies (Mgalaxies, De Lucia & Blaizot, 2007; Croton et al., 2006) and GALFORM (Dgalaxies, Bower et al., 2006; Benson et al., 2003; Cole et al., 2000).

While both models successfully explain a number of observational properties of galaxies like the break shape of the galaxy luminosity function, star formation rate, etc, they differ in detailed implementation. For example, while the L-Galaxies code uses the halo merger tree constructed by MHalos, the GALFORM code uses different criteria for identifying subhalos inside the FOF group, and thus uses a different merger tree. Also, two models use different gas cooling prescriptions and different initial mass functions (IMF) of star formation: L-Galaxies and GALFORM define the cooling radius, within which gas has a sufficient time to cool, by comparing the cooling time with halo dynamical time and the age of the halo, respectively. Cold gas turns into stars with two different IMFs: the L-Galaxies code uses IMF from Chabrier (2003) and the GALFORM code uses Kennicutt (1983). In addition to that, they treat AGN (Active Galactic Nucleus) feedback differently: the L-Galaxies code introduces a parametric model of AGN feedback depending on the black hole mass and the virial velocity of halo, and the GALFORM code imposes the condition that cooling flow is quenched when the energy released by radiative cooling (cooling luminosity) is less than some fraction (which is modeled by a parameter, ϵ_{SMBH}) of Eddington luminosity of the black hole. For more detailed comparison of the two model, we refer readers to the original papers cited above.

We compute the halo/galaxy power spectra from the Millennium Simulation as follows:

- (1) Use the Cloud-In-Cell (CIC) mass distribution scheme to calculate the density field on 1024^3 regular grid points from each catalog.

- (2) Fourier-transform the discretized density field using FFTW⁶.
- (3) Deconvolve the effect of the CIC pixelization and aliasing effect. We divide $P(\mathbf{k}, z) \equiv |\delta(\mathbf{k}, z)|^2$ at each cell by the following window function (Jing, 2005):

$$W(\mathbf{k}) = \prod_{i=1}^3 \left[1 - \frac{2}{3} \sin^2 \left(\frac{\pi k_i}{2k_N} \right) \right], \quad (4.5)$$

where $\mathbf{k} = (k_1, k_2, k_3)$, and $k_N \equiv \pi/H$ is the Nyquist frequency, (H is the physical size of the grid).⁷

- (4) Compute $P(k, z)$ by taking the angular average of CIC-corrected $P(\mathbf{k}, z) \equiv |\delta(\mathbf{k}, z)|^2$ within a spherical shell defined by $k - \Delta k/2 < |\mathbf{k}| < k + \Delta k/2$. Here, $\Delta k = 2\pi/500$ [h/Mpc] is the fundamental frequency that corresponds to the box size of the Millennium Simulation.

From the measured power spectra we find the maximum likelihood values of the bias parameters using the likelihood function approximated as a Gaussian:

$$\mathcal{L}(b_1, b_2, P_0) = \prod_{k_i < k_{max}} \frac{1}{\sqrt{2\pi\sigma_{P_i}^2}} \exp \left[-\frac{(P_{obs,i} - P_{g,i})^2}{2\sigma_{P_i}^2} \right], \quad (4.6)$$

where k_i 's are integer multiples of the fundamental frequency Δk , $P_{obs,i}$ is the measured power spectrum at $k = k_i$, $P_{g,i}$ is the theoretical model given by Eq. (4.2), and σ_{P_i} is the statistical error in the measured power spectrum.

We estimate σ_{P_i} in the same way as in § 4.2 (see also Appendix F.1). However, the power spectrum of the point-like particles like halos and galaxies includes the Poisson shot noise, $1/n$, where n is the number density of objects, on top of the power spectrum due to clustering. Therefore, σ_{P_i} must also include the shot-noise contribution. We use

$$\sigma_{P_i} = \sigma_P(k_i) = \sqrt{\frac{1}{N_{ki}}} \left[P_g(k_i) + \frac{1}{n} \right], \quad (4.7)$$

where

$$N_{ki} = 2\pi \left(\frac{k}{\Delta k} \right)^2 \quad (4.8)$$

⁶<http://www.fftw.org>

⁷Note that Eq. (4.5) is strictly valid for the flat (white noise) power spectrum, $P(k) = \text{constant}$. Nevertheless, it is still accurate for our purposes because, on small scales, both the halo and galaxy power spectra are dominated by the shot noise, which is also given by $P(k) = \text{constant}$.

is the number of independent Fourier modes used for estimating the power spectrum and $P_g(k_i)$ is the halo/galaxy power spectrum at $k = k_i$. Here, $\Delta k = 2\pi/(500 \ h^{-1} \text{ Mpc})$ is the fundamental wavenumber of the Millennium Simulation. Note that we subtract the Poisson shot noise contribution, $P_{shot} = 1/n$, from the observed power spectrum before the likelihood analysis.

Eq. (4.7) shows that the error on $P_{obs}(k)$ depends upon the underlying $P_g(k)$. For the actual data analysis one should vary $P_g(k)$ in the numerator of Eq. (4.6) as well as that in σ_{Pi} , simultaneously. However, to simplify the analysis, we evaluate the likelihood function in an iterative way: we first find the best-fitting $P_g(k)$ using σ_{Pi} with $P_g(k)$ in Eq. (4.7) replaced by $P_{obs}(k)$. Let us call this $\tilde{P}_g(k)$. We then use $\tilde{P}_g(k)$ in Eq. (4.7) for finding the best-fitting $P_g(k)$ that we shall report in this paper. Note that we iterate this procedure only once for current study.

Finally, we compute the 1-d marginalized 1- σ interval (or the marginalized 68.27% confidence interval) of each bias parameter by integrating the likelihood function (Eq. (4.6)), assuming a flat prior on the bias parameters (see also Appendix G).

We first analyze the power spectrum of halos (in § 4.3.2) as well as that of galaxies (in § 4.3.3) using all the halos and all the galaxies in the Millennium halo/galaxy catalogues. We then study the mass dependence of bias parameters in § 4.3.4.

In order to show that the non-linear bias model (Eq. 4.2) provides a much better fit than the linear bias model, we also fit the measured power spectra with two linear bias models: (i) linear bias with the linear matter power spectrum, and (ii) linear bias with the non-linear matter power spectrum from the 3rd-order PT. When fitting with the linear model, we use $k_{max} = 0.15 \ [h/\text{Mpc}]$ for all redshift bins.

4.3.2 Halo power spectra

4.3.2.1 Measuring non-linear halo bias parameters

Figure 4.5 shows the best-fitting non-linear (*solid lines*) and linear bias models (*dashed and dot-dashed lines*), compared with the halo spectra estimated from the Millennium Simulation (*points with errorbars*). The smaller panels show the residuals of fits. The maximum wavenumber used in the fits, $k_{max}(z)$, are also marked with the arrows (bigger panels), and the vertical lines (smaller panels). We find that the non-linear bias model provides substantially better fits than the linear bias models.

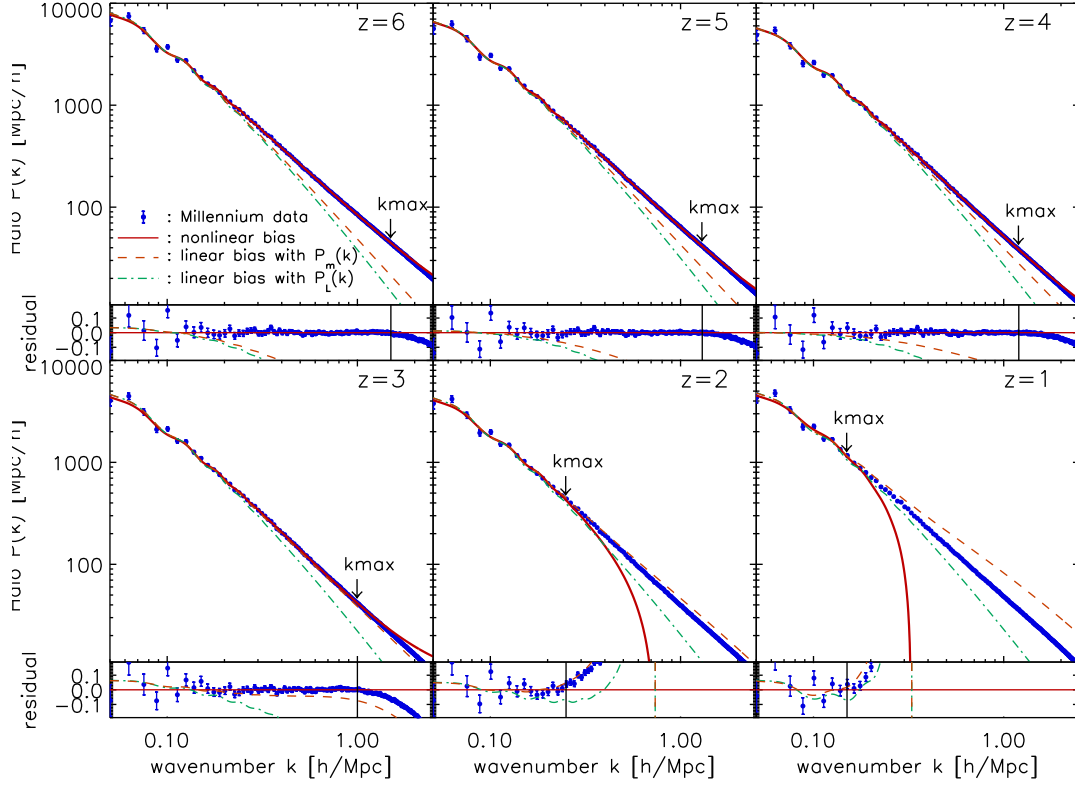


Figure 4.5: Halo power spectra from the Millennium Simulation at $z = 1, 2, 3, 4, 5$, and 6 . Also shown in smaller panels are the residual of fits. The points with errorbars show the measured halo power spectra, while the solid, dashed, and dot-dashed lines show the best-fitting non-linear bias model (Eq. (4.2)), the best-fitting linear bias with the non-linear matter power spectrum, and the best-fitting linear bias with the linear matter power spectrum, respectively. Both linear models have been fit for $k_{max,linear} = 0.15 [h \text{ Mpc}^{-1}]$, whereas $k_{max}(z)$ given in Table 4.1 (also marked in each panel) have been used for the non-linear bias model.

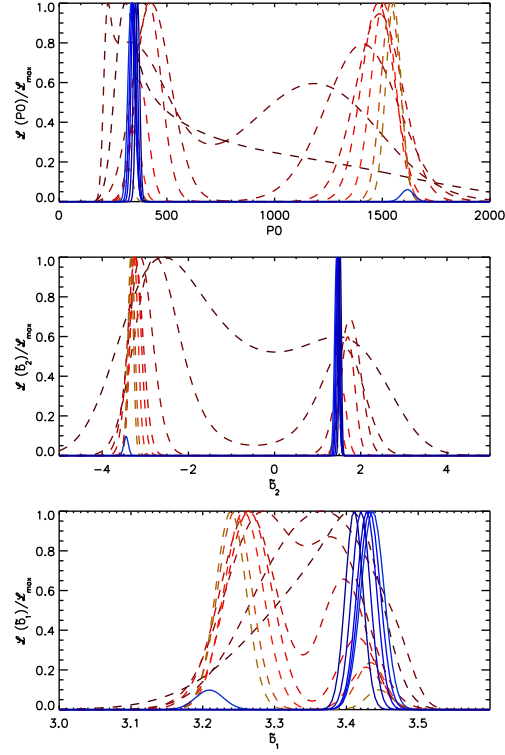


Figure 4.6: One-dimensional marginalized distribution of non-linear bias parameters at $z = 6$: from top to bottom panels, P_0 , b_2 , and b_1 . Different lines show the different values of k_{max} used for the fits. The dashed and solid lines correspond to $0.3 \leq k_{max}/[h \text{ Mpc}^{-1}] \leq 1.0$ and $1.0 < k_{max}/[h \text{ Mpc}^{-1}] \leq 1.5$, respectively. The double-peak structure disappears for higher k_{max} .

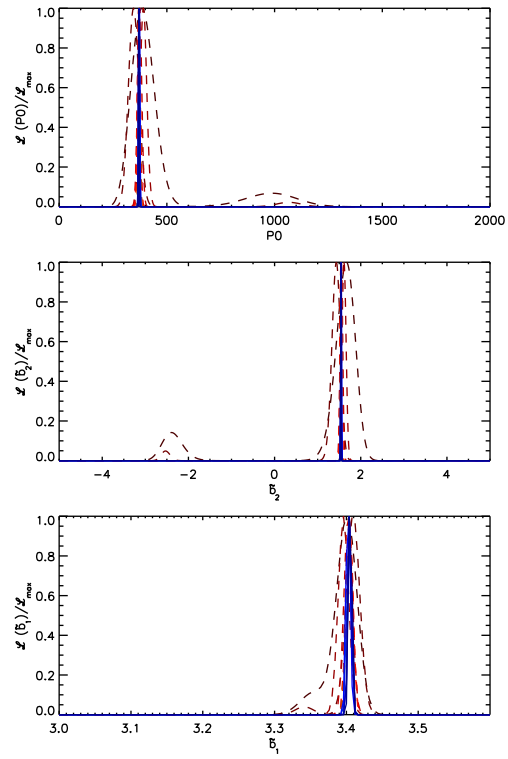


Figure 4.7: Same as Figure 4.6, but for a Monte Carlo simulation of a galaxy survey with a bigger box size, $L_{box} = 1.5 \text{ Gpc}/h$.

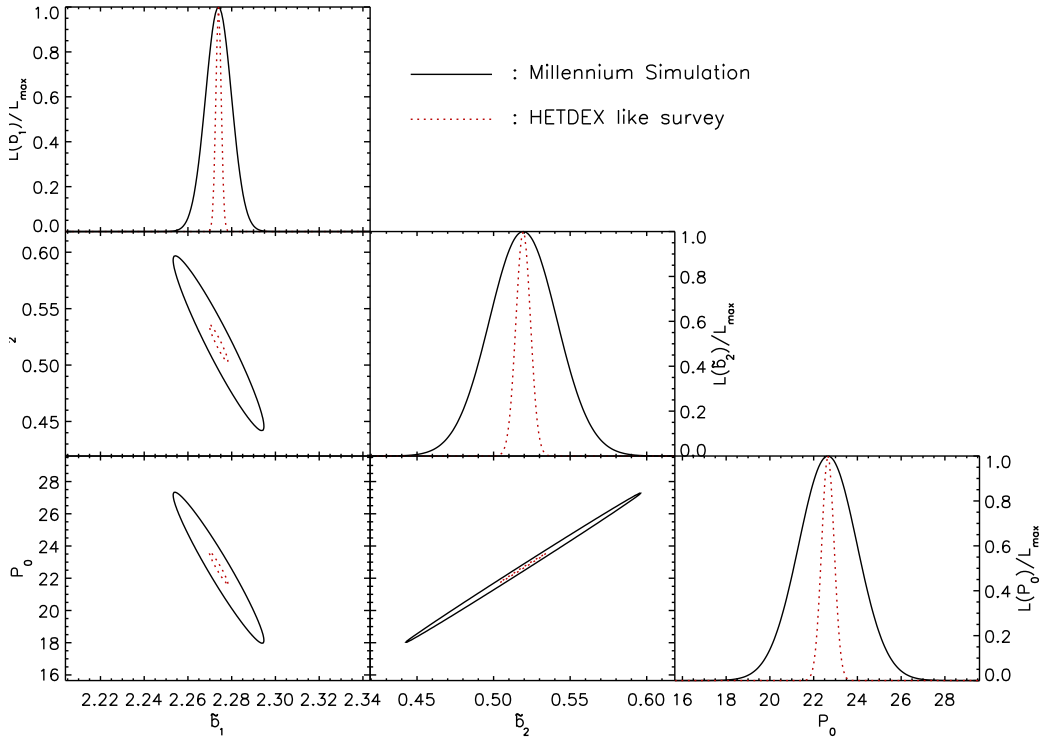


Figure 4.8: One-dimensional marginalized constraints and two-dimensional joint marginalized constraint of $2\text{-}\sigma$ (95.45% CL) range for bias parameters (b_1, b_2, P_0). Covariance matrices are calculated from the Fisher information matrix (Eq. (4.9)) with the best-fitting bias parameters for halo at $z = 4$.

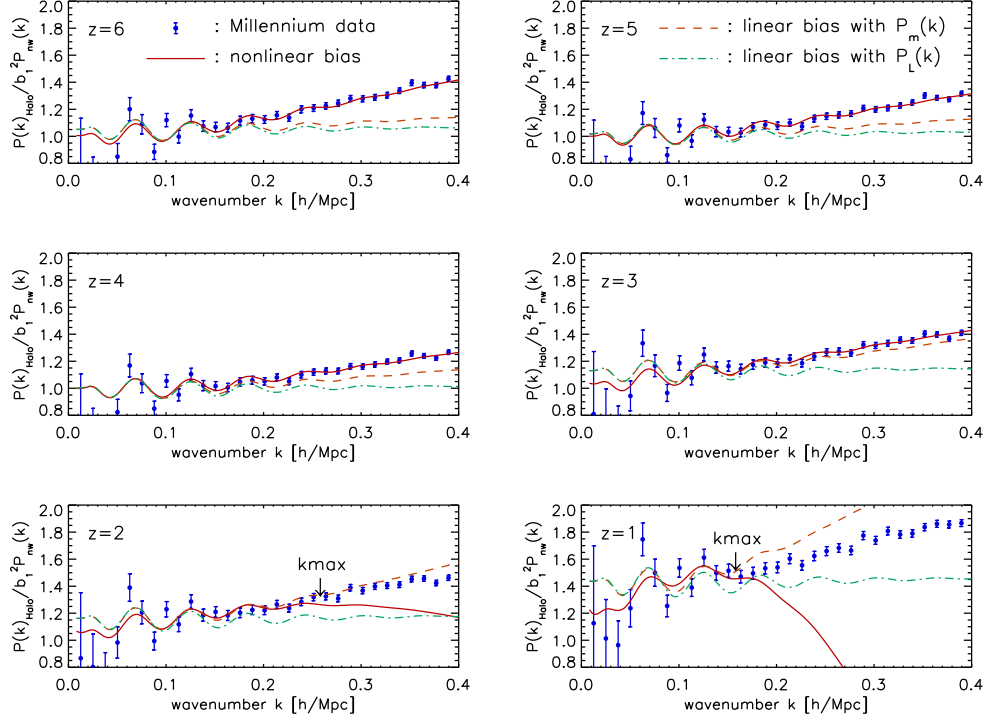


Figure 4.9: Distortion of BAOs due to non-linear matter clustering and non-linear halo bias. All of the power spectra have been divided by a smooth power spectrum without baryonic oscillations from equation (29) of Eisenstein & Hu (1998). The errorbars show the Millennium Simulation, while the solid lines show the PT calculations. The dashed lines show the linear bias model with the non-linear matter power spectrum, and the dot-dashed lines show the linear bias model with the linear matter power spectrum. Therefore, the difference between the solid lines and the dashed lines shows the distortion solely due to non-linear halo bias.

Table 4.3: Non-linear halo bias parameters and the corresponding 68% interval estimated from the MPA halo power spectra

z	b_1	b_2	P_0 ([Mpc/h] ³)	b_1^L	b_1^{LL}	b_1^{ST}	b_2^{ST}
6	3.41±0.01	1.52±0.03	141.86±3.73	3.50±0.03	3.51±0.03	3.69	2.10
5	2.76±0.01	0.91±0.03	57.77±2.84	2.79±0.03	2.80±0.03	3.16	1.70
4	2.27±0.01	0.52±0.03	22.65±1.88	2.28±0.02	2.29±0.02	2.77	1.40
3	1.52±0.01	-1.94±0.05	329.42±10.6	1.62±0.01	1.63±0.01	2.23	1.07
2	1.10±0.06	-2.12±0.65	507.25±214.7	1.19±0.01	1.20±0.01	1.84	0.76
1	0.74±0.09	-3.05±1.49	1511.46±526.7	0.88±0.01	0.90±0.01	1.54	0.58

z : redshift

b_1, b_2, P_0 : non-linear bias parameters

b_1^L : linear bias parameter for the linear bias model with the 3rd-order matter power spectrum

b_1^{LL} : linear bias parameter for the linear bias model with the linear power spectrum

b_1^{ST}, b_2^{ST} : non-linear bias parameters calculated from the Sheth-Tormen model, $b_2^{ST} = b_2^{ST}/b_1$

Caution: We estimate 1- σ ranges for the low redshift ($z \leq 3$) only for the peak which involves the maximum likelihood value. If two peaks in marginalized likelihood function are blended, we use only unblended side of the peak to estimate the 1- σ range.

We find that all of non-linear bias parameters, b_1 , b_2 , and P_0 , are strongly degenerate, when the maximum wavenumbers used in the fits, k_{max} , are small. In Figure 4.6 we show the one-dimensional marginalized distribution of bias parameters at $z = 6$, as a function of k_{max} . For lower k_{max} , $0.3 \leq k_{max}/[h \text{ Mpc}^{-1}] \leq 1.0$, the marginalized distribution has two peaks (*dashed lines*), indicating strong degeneracy with the other parameters. The double-peak structure disappears for $1.0 < k_{max}/[h \text{ Mpc}^{-1}] \leq 1.5$ (*solid lines*).

We find that the origin of degeneracy is simply due to the small box size of the Millennium Simulation, i.e., the lack of statistics, or too a large sampling variance. To show this, we have generated a mock Monte Carlo realization of halo power spectra, assuming a much bigger box size, $L_{box} = 1.5 h^{-1} \text{ Gpc}$, which gives the fundamental frequency of $\Delta k = 5.0 \times 10^{-4} h \text{ Mpc}^{-1}$. Note that this volume roughly corresponds to that would be surveyed by the HETDEX survey (Hill et al., 2004). We have used the same non-linear matter power spectrum and the best-fitting bias parameters from the Millennium Simulation (MPA halos) when creating Monte Carlo realizations. The resulting marginalized likelihood function at $z = 6$ is shown in Figure 4.7. The double-peak structure has disappeared even for low k_{max} , $k_{max} = 0.3 h \text{ Mpc}^{-1}$. Therefore, we conclude that the double-peak problem can be resolved simply by increasing the survey volume.

The best-fitting non-linear halo bias parameters and the corresponding 1- σ intervals

are summarized in Table 4.3. Since we know that the double-peak structure is spurious, we pick one peak that corresponds to the maximum likelihood value, and quote the $1\text{-}\sigma$ interval. At $z \leq 2$, the bias parameters are not constrained very well because of lower k_{max} and the limited statistics of the Millennium Simulation, and hence the two peaks are blended; thus, we estimate $1\text{-}\sigma$ range only from the unblended side of the marginalized likelihood function. Two linear bias parameters, one with the linear matter power spectrum and another with the non-linear PT matter power spectrum, are also presented with their $1\text{-}\sigma$ intervals.

4.3.2.2 Degeneracy of bias parameters

In order to see how strongly degenerate bias parameters are, we calculate the covariance matrix of each pair of bias parameters. We calculate the covariance matrix of each pair of bias parameters by using the Fisher information matrix, which is the inverse of the covariance matrix. The Fisher information matrix for the galaxy power spectrum can be approximated as (Tegmark, 1997) ⁸

$$F_{ij} = \sum_n \frac{1}{\sigma_P^2(k_n)} \frac{\partial P(k_n, \theta)}{\partial \theta_i} \frac{\partial P(k_n, \theta)}{\partial \theta_j} \quad (4.9)$$

where θ is a vector in the parameter space, $\theta_i = b_1, b_2, P_0$, for $i = 1, 2, 3$, respectively. We calculate the marginalized errors on the bias parameters as following. We first calculate the full Fisher matrix and invert it to estimate the covariance matrix. Then, we get the the covariance matrices of any pairs of bias parameters by taking the 2 by 2 sub-matrix of the full covariance matrix. Figure 4.8 shows the resulting $2\text{-}\sigma$ (95.45% interval) contour for the bias parameters at $z = 4$. We find the strong degeneracy between \tilde{P}_0 and b_2 .

⁸Eq. (4.9) is equivalent to Eq. (6) in Tegmark (1997). The number of k mode in real space power spectrum from a survey of volume V is (See Appendix F.1 for notations.)

$$N_{k_n} = \frac{4\pi k_n^2 \delta k_n}{2(\delta k_n)^3} = \frac{V k_n^2 \delta k_n}{4\pi^2}.$$

Then, the variance of power spectrum (Eq. (4.7)) becomes

$$\sigma_P^2(k_n) = \frac{4\pi^2}{V k_n^2 \delta k_n} \left[P(k_n) + \frac{1}{n} \right]^2 = \frac{4\pi^2 P(k_n)^2}{k_n^2 \delta k_n} \frac{1}{V_{\text{eff}}(k_n)},$$

where V_{eff} is the constant density version of Eq. (5) of Tegmark (1997). Finally, the elements of Fisher matrix are given by

$$F_{ij} = \sum_n \frac{1}{\sigma_P^2(k_n)} \frac{\partial P(k_n, \theta)}{\partial \theta_i} \frac{\partial P(k_n, \theta)}{\partial \theta_j} = \frac{1}{4\pi^2} \sum_n \frac{\partial P(k_n, \theta)}{\partial \theta_i} \frac{\partial P(k_n, \theta)}{\partial \theta_j} \frac{V_{\text{eff}}(k_n) k_n^2 \delta k_n}{P(k_n)^2}$$

which is the same as Eq. (6) in Tegmark (1997).

We also find that b_1 is degenerate with the other two parameters. On top of the error contours for the Millennium Simulation, we show the expected contour from the HETDEX like survey (1.5 Gpc/h). Since the volume of HETDEX like survey is 27 times bigger, the likelihood functions and the error-contours are about a factor of 5 smaller than those from the Millennium Simulation. Other than that, two contours follow the same trend. Results are the same for the other redshifts.

4.3.2.3 Comparison with the halo model predictions

The effective linear bias, b_1 , is larger at higher redshifts. This is the expected result, as halos of mass greater than $\sim 10^{10} M_\odot$ were rarer in the earlier time, resulting in the larger bias.

From the same reason, we expect that the non-linear bias parameters, b_2 and P_0 , are also larger at higher z . While we observe the expected trend at $z \geq 4$, the results from $z \leq 3$ are somewhat peculiar. This is probably due to the large sampling variance making the fits unstable: for $z \leq 3$ the maximum wavenumbers inferred from the matter power spectra are less than $1.0 h \text{ Mpc}^{-1}$ (see Table 4.1), which makes the likelihood function double-peaked and leaves the bias parameters poorly constrained.

How do these bias parameters compare with the expected values? We use the halo model for computing the mass-averaged bias parameters, b_1^{ST} and b_2^{ST} , assuming that the minimum mass is given by the minimum mass of the MPA halo catalog, $M_{min} = 1.72 \times 10^{10} M_\odot/h$:

$$b_i^{ST} = \frac{\int_{M_{min}}^{M_{max}} \frac{dn}{dM} M b_i(M) dM}{\int_{M_{min}}^{M_{max}} \frac{dn}{dM} M dM}, \quad (4.10)$$

where dn/dM is the Sheth-Tormen mass function and $b_i(M)$ is the i -th order bias parameter from Scoccimarro et al. (2001a).

There is one subtlety. The halo model predicts the coefficients of the Taylor series (Eq. (4.1)), whereas what we have measured are the re-parametrized bias parameters given by Eq. (4.3). However, the formula for b_1 includes the mass variance, σ^2 , which depends on our choice of a smoothing scale that is not well defined. This shows how difficult it is to actually compute the halo power spectrum from the halo model. While the measured values of b_1 and the predicted b_1^{ST} compare reasonably well, it is clear that we cannot use the predicted bias values for doing cosmology.

For b_2 , we compute $b_2^{ST} = b_2^{ST}/b_1$ where b_1 is the best-fitting value from the Millennium Simulation. This would give us a semi apple-to-apple comparison. Nevertheless, while the agreement is reasonable at $z \geq 4$, the halo model predictions should not be used for predicting b_2 either.

4.3.2.4 Comments on the bispectrum

While the degeneracy between bias parameters may appear to be a serious issue, there is actually a powerful way of breaking degeneracy: the bispectrum, the Fourier transform of the 3-point correlation function (Matarrese et al., 1997). The reduced bispectrum, which is the bispectrum normalized properly by the power spectrum, depends primarily on two bias parameters, b_1 and b_2 , nearly independent of the cosmological parameters (Sefusatti et al., 2006). Therefore, one can use this property to fix the bias parameters, and use the power spectrum for determining the cosmological parameters and the remaining bias parameter, P_0 . Sefusatti & Komatsu (2007) have shown that the planned high- z galaxy surveys would be able to determine b_1 and b_2 with a few percent accuracy.

We have begun studying the bispectrum of the Millennium Simulation. Our preliminary results show that we can indeed obtain better constraints on b_1 and b_2 from the bispectrum than from the power spectrum, provided that we use the same k_{max} for both the bispectrum and power spectrum analysis. Therefore, even when the non-linear bias parameters are poorly constrained by the power spectrum alone, or have the double-peak likelihood function from the power spectrum for lower k_{max} , we can still find tight constraints on b_1 and b_2 from the bispectrum. These results will be reported elsewhere.

4.3.2.5 Effects on BAOs

In Figure 4.9 we show the distortion of BAO features due to non-linear matter clustering and non-linear bias. To show only the distortions of BAOs at each redshift, we have divided the halo power spectra by smooth power spectra without baryonic oscillations from equation (29) of Eisenstein & Hu (1998) with b_1^2 multiplied. Three theoretical models are shown: the non-linear bias model (*solid line*), a linear bias model with the 3rd-order matter power spectrum (*dashed line*), and a linear bias model with the linear matter power spectrum (*dot-dashed line*). Therefore, the difference between the solid lines and the dashed lines is solely due to non-linear halo bias.

The importance of non-linear bias affecting BAOs grows with z ; however, as the matter clustering is weaker at higher z , the 3rd-order PT still performs better than at lower

z . In other words, the higher bias at higher z does not mean that surveys at higher z are worse at measuring BAOs; on the contrary, it is still easier to model the halo power spectrum at higher z than at lower z . For $z \geq 3$, where k_{max} is larger than the BAO scale, the distortion of BAOs is modeled very well by the non-linear bias model, while the linear bias models fail badly.

The sampling variance of the Millennium Simulation at $k \lesssim 0.15 h \text{ Mpc}^{-1}$ is too large for us to study the distortion on the first two BAO peaks. Since the PT performs well at higher k , we expect that the PT describes the first two peaks even better. However, to show this explicitly one would need to run a bigger simulation with a bigger volume with the same mass resolution as the Millennium Simulation, which should be entirely doable with the existing computing resources.

4.3.3 Galaxy power spectra

4.3.3.1 Measuring non-linear galaxy bias parameters

Table 4.4: Non-linear halo bias parameters and the corresponding 68% interval estimated from the MPA galaxy power spectra

z	b_1	b_2	P_0 ($[h/\text{Mpc}]^3$)	b_1^L	b_1^{LL}	b_1^{ST}	b_2^{ST}
6	3.55 ± 0.01	1.70 ± 0.03	194.23 ± 4.45	3.67 ± 0.03	3.68 ± 0.03	3.10	1.03
5	2.93 ± 0.01	1.08 ± 0.03	94.08 ± 3.71	2.97 ± 0.03	2.98 ± 0.03	2.55	0.59
4	2.46 ± 0.01	0.68 ± 0.03	47.79 ± 2.84	2.47 ± 0.02	2.48 ± 0.02	2.13	0.28
3	1.69 ± 0.01	-2.12 ± 0.04	486.69 ± 12.7	1.83 ± 0.02	1.83 ± 0.02	1.58	-0.12
2	1.28 ± 0.08	-2.16 ± 0.64	738.22 ± 291.3	1.40 ± 0.01	1.40 ± 0.01	1.19	-0.34
1	0.89 ± 0.11	-2.97 ± 1.60	2248.35 ± 786.13	1.09 ± 0.01	1.10 ± 0.01	0.91	-0.45

z : redshift

b_1, b_2, P_0 : non-linear bias parameters

b_1^L : linear bias parameter for the linear bias model with the 3rd-order matter power spectrum

b_1^{LL} : linear bias parameter for the linear bias model with the linear power spectrum

b_1^{ST}, b_2^{ST} : non-linear bias parameters calculated from the Sheth-Tormen model, $b_2^{ST} = b_2^{ST}/b_1$

Caution: We estimate 1- σ ranges for the low redshift ($z \leq 3$) only for the peak which involves the maximum likelihood value. If two peaks in maginalized likelihood function are blended, we use only unblended side of the peak to estimate the 1- σ range.

Figures 4.10 and 4.11 show the galaxy power spectra estimated from the MPA (Mgalaxy) and Durham (Dgalaxy) galaxy catalogues, respectively. Here, we basically find the same story as we have found for the halo power spectra (§ 4.3.2): for $k < k_{max}$ the non-linear bias model fits both galaxy power spectra (Mgalaxy and Dgalaxy), whereas the

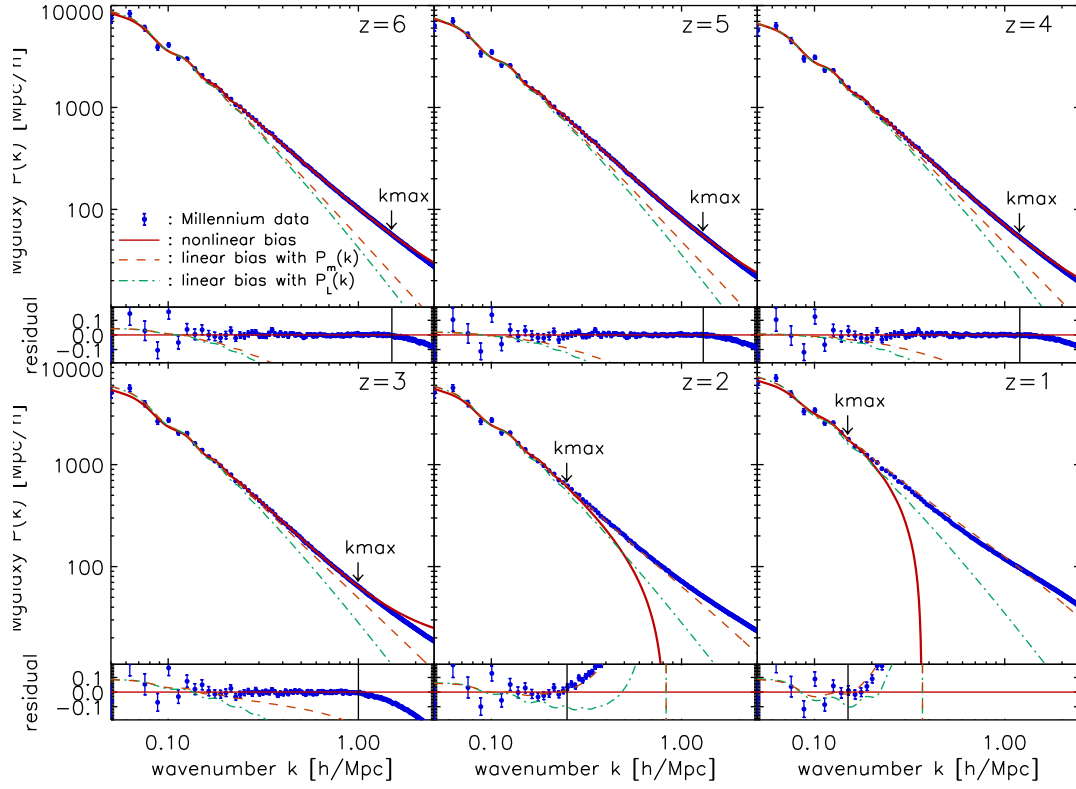


Figure 4.10: Same as Figure 4.5, but for the MPA galaxy catalogue (Mgalaxy).

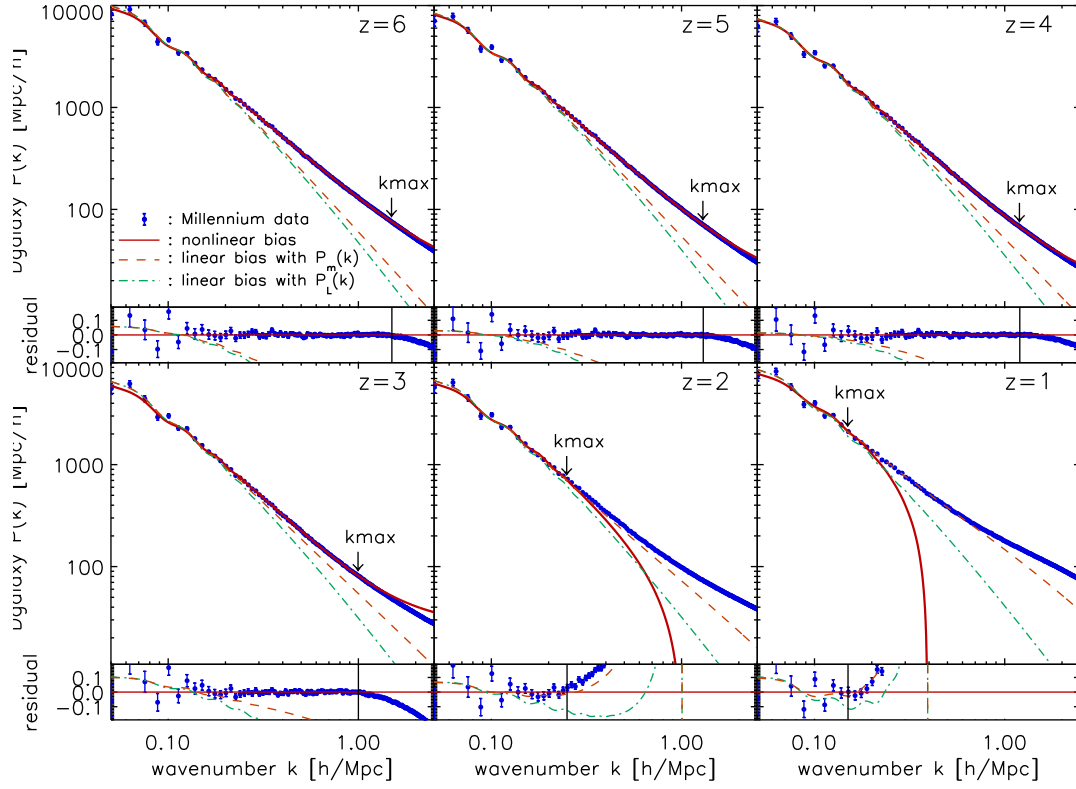


Figure 4.11: Same as Figure 4.5, but for the Durham galaxy catalogue (Dgalaxy).

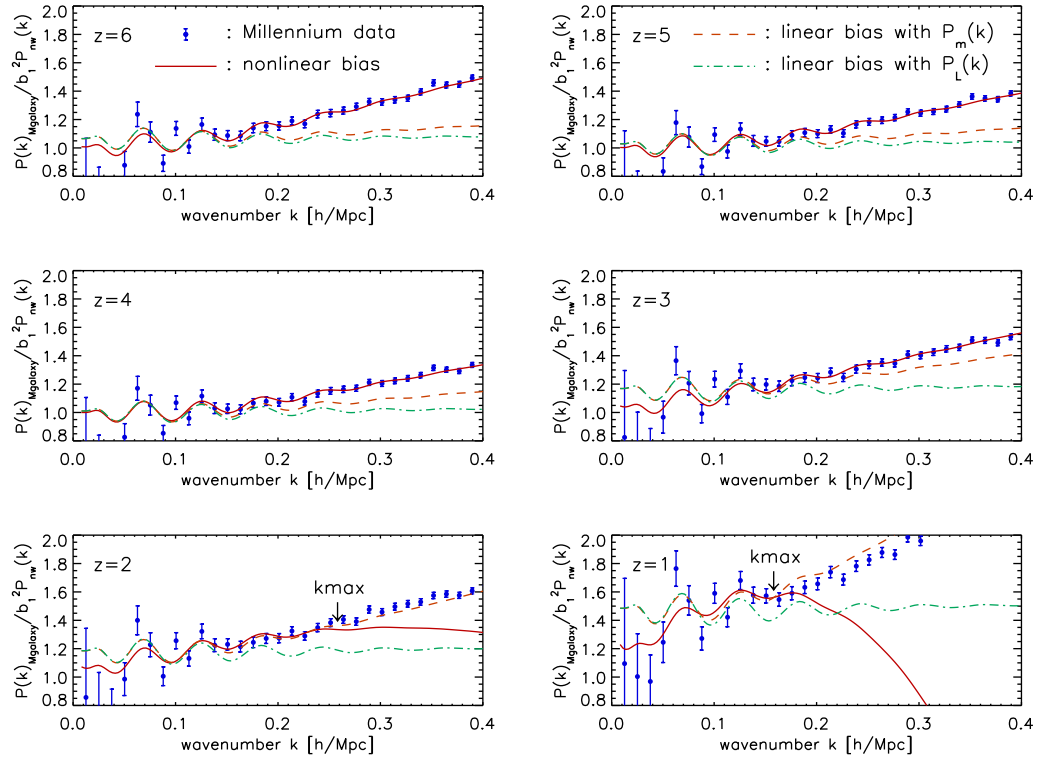


Figure 4.12: Same as Figure 4.9, but for the MPA galaxy power spectrum (M_{galaxy}).

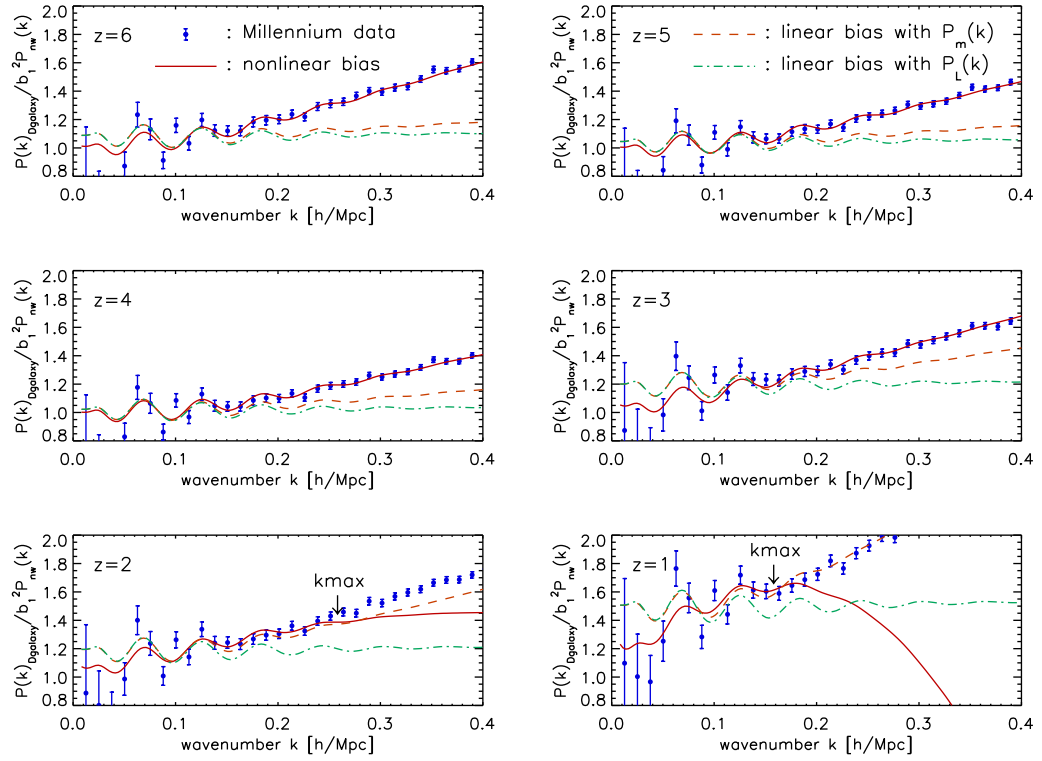


Figure 4.13: Same as Figure 4.9, but for the Durham galaxy power spectrum (D_{galaxy}).

Table 4.5: Non-linear halo bias parameters and the corresponding 68% interval estimated from the Durham galaxy power spectra

z	b_1	b_2	P_0 ($[h/\text{Mpc}]^3$)	b_1^L	b_1^{LL}	b_1^{ST}	b_2^{ST}
6	3.73 ± 0.01	1.96 ± 0.03	288.39 ± 5.82	3.90 ± 0.04	3.90 ± 0.04	3.10	0.98
5	3.07 ± 0.01	1.26 ± 0.03	143.15 ± 4.81	3.15 ± 0.03	3.15 ± 0.03	2.55	0.56
4	2.57 ± 0.01	0.83 ± 0.03	78.97 ± 3.93	2.60 ± 0.02	2.61 ± 0.02	2.13	0.26
3	1.75 ± 0.01	-2.26 ± 0.04	604.65 ± 13.8	1.92 ± 0.02	1.93 ± 0.02	1.58	-0.11
2	1.36 ± 0.08	-2.14 ± 0.65	843.49 ± 331.4	1.49 ± 0.01	1.50 ± 0.01	1.19	-0.32
1	0.96 ± 0.11	-2.94 ± 1.62	2640.20 ± 960.32	1.18 ± 0.01	1.20 ± 0.01	0.91	-0.42

z : redshift

b_1, b_2, P_0 : non-linear bias parameters

b_1^L : linear bias parameter for the linear bias model with the 3rd-order matter power spectrum

b_1^{LL} : linear bias parameter for the linear bias model with the linear power spectrum

b_1^{ST}, b_2^{ST} : non-linear bias parameters calculated from the Sheth-Tormen model, $b_2^{ST} = b_2^{ST}/b_1$

Caution: We estimate 1- σ ranges for the low redshift ($z \leq 3$) only for the peak which involves the maximum likelihood value. If two peaks in maginalized likelihood function are blended, we use only unblended side of the peak to estimate the 1- σ range.

linear bias models fit neither.

The galaxy bias parameters extracted from Mgalaxy and Dgalaxy are summarized in Table 4.4 and 4.5, respectively. While the bias parameters are different for halo, Mgalaxy and Dgalaxy, they follow the same trend: (i) b_1 becomes lower as the redshift becomes lower, and (ii) b_2 also becomes lower as the redshift becomes lower when $z > 3$, but suddenly changes to large negative values at $z \leq 3$. As we have already pointed out in § 4.3.2, this sudden peculiar change is most likely caused by the double-peak nature of the likelihood function, owing to the poor statistical power for lower k_{max} at lower z . In order to study b_2 further with better statistics, one needs a bigger simulation.

4.3.3.2 Comparison with the simplest HOD predictions

To give a rough theoretical guide for the galaxy bias parameters, we assume that each dark matter halo hosts one galaxy above a certain minimum mass. This specifies the form of the HOD completely: $\langle N|M \rangle = 1$, with the same lower mass cut-off as the minimum mass of the halo, $M_{min} = 1.72 \times 10^{10} M_\odot/h$.

This is utterly simplistic, and is probably not correct for describing Mgalaxy or Dgalaxy. Nevertheless, we give the resulting values in Table 4.4 and 4.5, which have been

computed from

$$b_i^{ST} = \frac{\int_{M_{min}}^{M_{max}} \frac{dn}{dM} b_i(M) \langle N|M \rangle dM}{\int_{M_{min}}^{M_{max}} \frac{dn}{dM} \langle N|M \rangle dM}, \quad (4.11)$$

where dn/dM is the Sheth-Tormen mass function and $b_i(M)$ is the i -th order bias parameter from Scoccimarro et al. (2001a). To compare with the non-linear bias parameters, we also calculate $b_2 = b_2^{ST}/b_1$.

While these “predictions” give values that are reasonably close to the ones obtained from the fits, they are many σ away from the best-fitting values. The freedom in the choice of the HOD may be used to make the predicted values match the best-fitting values; however, such an approach would require at least as many free parameters as the non-linear bias parameters. Also, given that the *halo* bias prediction fails to fit the halo power spectra, the HOD approach, which is still based upon knowing the halo bias, is bound to fail as well.

4.3.3.3 Effects on BAOs

In Figures 4.12 and 4.13 we show how non-linear galaxy bias distorts the structure of BAOs. Again, we find the same story as we have found for the halo bias: the galaxy bias distorts BAOs more at higher z because, for a given mass, galaxies were rarer at higher redshifts and thus more highly biased, while the quality of the fits is better at higher z because of less non-linearity in the matter clustering.

In all cases (halo, Mgalaxy and Dgalaxy) the non-linear bias model given by Eq. (4.2) provides very good fits, and describes how bias modifies BAOs.

4.3.4 Mass dependence of bias parameters and effects on BAOs

So far, we have used all the available halos and galaxies in the Millennium catalogues for computing the halo and galaxy power spectra. In this section we divide the samples into different mass bins given by $M < 5 \times 10^{10} M_\odot/h$, $5 \times 10^{10} M_\odot/h < M < 10^{11} M_\odot/h$, $10^{11} M_\odot/h < M < 5 \times 10^{11} M_\odot/h$, $5 \times 10^{11} M_\odot/h < M < 10^{12} M_\odot/h$, and study how the derived bias parameters depend on mass.

The power spectra of the selected halos and galaxies in a given mass bin are calculated and fit in the exactly same manner as before. Note that we shall use only the halo and Mgalaxy, as we expect that Dgalaxy would give similar results to Mgalaxy.

Figures 4.14 and 4.15 show the results for the halo and galaxies, respectively. To compare the power spectra of different mass bins in the same panel, and highlight the effects

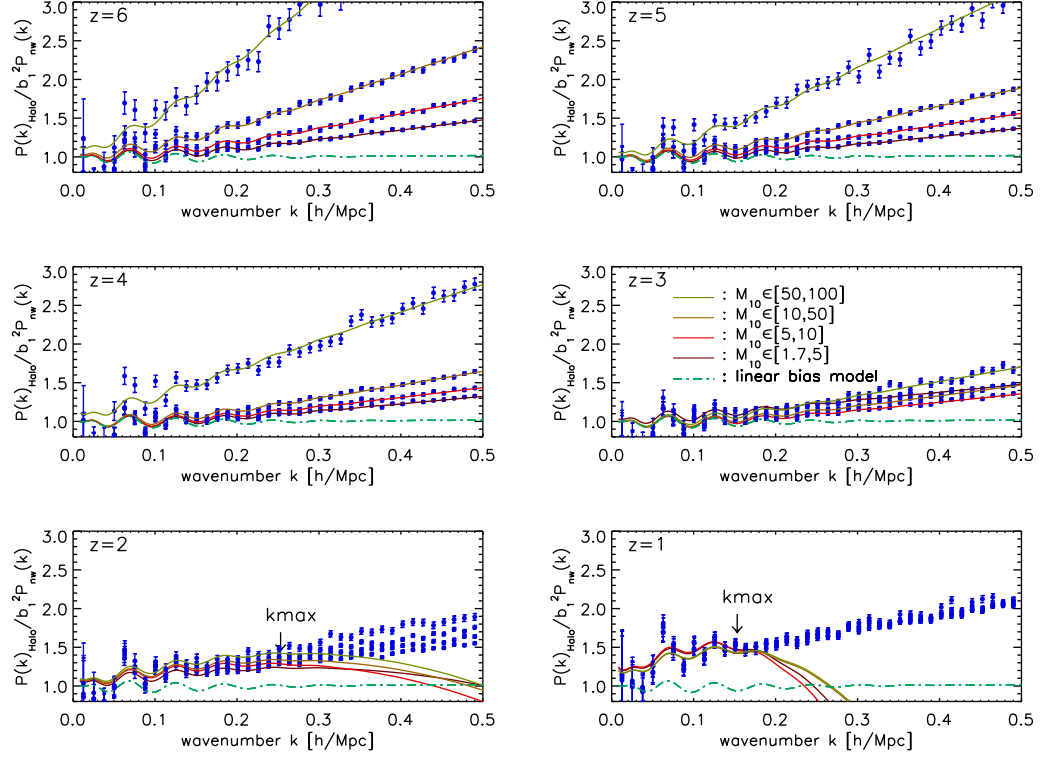


Figure 4.14: Mass dependence of distortion of BAOs due to non-linear bias. Four mass bins, $M < 5 \times 10^{10} M_{\odot}/h$, $5 \times 10^{10} M_{\odot}/h < M < 10^{11} M_{\odot}/h$, $10^{11} M_{\odot}/h < M < 5 \times 10^{11} M_{\odot}/h$, and $5 \times 10^{11} M_{\odot}/h < M < 10^{12} M_{\odot}/h$, are shown. (M_{10} stands for $M/(10^{10} M_{\odot})$.) All of the power spectra have been divided by a smooth power spectrum without baryonic oscillations from equation (29) of Eisenstein & Hu (1998). The errorbars show the Millennium Simulation data, while the solid lines show the PT calculation.

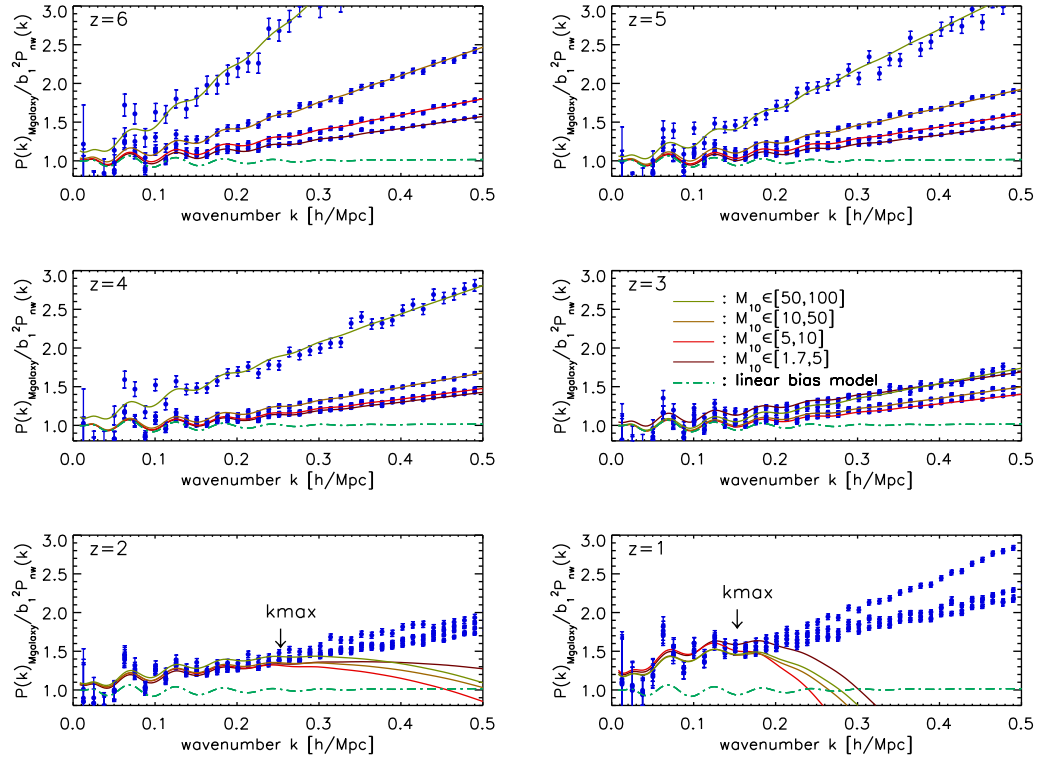


Figure 4.15: Same as Figure 4.14, but for the MPA galaxy catalogue (Mgalaxy).

on BAOs at the same time, we have divided the power spectra by a non-oscillating matter power spectrum from equation (29) of Eisenstein & Hu (1998) with the best-fitting b_1^2 from each mass bin multiplied. These figures show the expected results: the larger the mass is, the larger the non-linear bias becomes. Nevertheless, the 3rd-order PT calculation captures the dependence on mass well, and there is no evidence for failure of the PT for highly biased objects.

In Tables 4.6 and 4.7 we give values of the measured bias parameters as well as the “predicted” values. For all redshifts we see the expected trend again: the higher the mass is, the larger the effective linear bias (b_1) is. The same is true for b_2 for $z > 3$, while it is not as apparent for lower redshifts, and eventually becomes almost fuzzy for $z = 1$. Again, these are probably due to the lack of statistics due to lower values of k_{max} at lower z , and we need a bigger simulation to handle these cases with more statistics.

The high values of bias do not mean failure of PT. The PT galaxy power spectrum model fails only when $\Delta_m^2(k, z)$ exceeds ~ 0.4 (Paper I), or the locality of bias is violated. Overall, we find that the non-linear bias model given by Eq. (4.2) performs well for halos and galaxies with all mass bins, provided that we use the data only up to k_{max} determined from the matter power spectra. This implies that the locality assumption is a good approximation for $k < k_{max}$; however, is it good enough for us to extract cosmology from the observed galaxy power spectra?

4.4 Cosmological parameter estimation with the non-linear bias model

In the previous sections we have shown that the 3rd-order PT galaxy power spectrum given by Eq. (4.2) provides good fits to the galaxy power spectrum data from the Millennium Simulation.

However, we must not forget that Eq. (4.2) contains 3 *free parameters*, b_1 , b_2 , and P_0 . With 3 parameters it may seem that it should not be so difficult to fit smooth curves like those shown in, e.g., Figure 4.10.

While the quality of fits is important, it is not the end of story. We must also show that Eq. (4.2) can be used for extracting the *correct cosmological parameters* from the observed galaxy power spectra.

In this section we shall extract the distance scale from the galaxy power spectra of the Millennium Simulation, and compare them with the input values that were used to

generate the simulation. If they do not agree, Eq. (4.2) must be discarded. If they do, we should proceed to the next level by including non-linear redshift space distortion.

4.4.1 Measuring Distance Scale

4.4.1.1 Background

Dark energy influences the expansion rate of the universe as well as the growth of structure (see Copeland et al., 2006, for a recent review).

The cosmological distances, such as the luminosity distance, $D_L(z)$, and angular diameter distance, $D_A(z)$, are powerful tools for measuring the expansion rates of the universe, $H(z)$, over a wide range of redshifts. Indeed, it was $D_L(z)$ measured out to high- z ($z \leq 1.7$) Type Ia supernovae that gave rise to the first compelling evidence for the existence of dark energy (Riess et al., 1998; Perlmutter et al., 1999). The CMB power spectrum provides us with a high-precision measurement of $D_A(z_*)$ out to the photon decoupling epoch, $z_* \simeq 1090$ (see Komatsu et al., 2010, for the latest determination from the WMAP 5-year data).

The galaxy power spectrum can be used for measuring $D_A(z)$ as well as $H(z)$ over a wider range of redshifts. From galaxy surveys we find three-dimensional positions of galaxies by measuring their angular positions on the sky as well as their redshifts. We can then estimate the two-point correlation function of galaxies as a function of the angular separation, $\Delta\theta$, and the redshift separation, Δz . To convert $\Delta\theta$ and Δz into the comoving separations perpendicular to the line of sight, Δr_\perp , and those along the line of sight, Δr_\parallel , one needs to know $D_A(z)$ and $H(z)$, respectively, as

$$\Delta r_\perp = (1+z)D_A(z)\Delta\theta, \quad (4.12)$$

$$\Delta r_\parallel = \frac{c\Delta z}{H(z)}, \quad (4.13)$$

where $(1+z)$ appears because $D_A(z)$ is the proper (physical) angular diameter distance, whereas Δr_\perp is the comoving separation. Therefore, if we know Δr_\perp and Δr_\parallel *a priori*, then we may use the above equations to measure $D_A(z)$ and $H(z)$.

The galaxy power spectra contain at least three distance scales which may be used in the place of Δr_\perp and Δr_\parallel : (i) the sound horizon size at the so-called baryon drag epoch, $z_{drag} \simeq 1020$, at which baryons were released from the baryon-photon plasma, (ii) the photon horizon size at the matter-radiation equality, $z_{eq} \simeq 3200$, and (iii) the Silk damping scale (see, e.g., Eisenstein & Hu, 1998).

In Fourier space, we may write the observed power spectrum as (Seo & Eisenstein, 2003)

$$P_{obs}(k_{\parallel}, k_{\perp}, z) = \left(\frac{D_A(z)}{D_{A,true}(z)} \right)^2 \left(\frac{H_{true}(z)}{H(z)} \right) P_{true} \left(\frac{D_{A,true}(z)}{D_A(z)} k_{\perp}, \frac{H(z)}{H_{true}(z)} k_{\parallel}, z \right), \quad (4.14)$$

where k_{\perp} and k_{\parallel} are the wavenumbers perpendicular to and parallel to the line of sight, respectively, and $P_{true}(k)$, $D_{A,true}(z)$, and $H_{true}(z)$ are the true, underlying values. We then vary $D_A(z)$ and $H(z)$, trying to estimate $D_{A,true}(z)$ and $H_{true}(z)$.

There are two ways of measuring $D_A(z)$ and $H(z)$ from the galaxy power spectra:

- (1) Use BAOs. The BAOs contain the information of one of the standard rulers, the sound horizon size at z_{drag} . This method relies on measuring only the phases of BAOs, which are markedly insensitive to all the non-linear effects (clustering, bias, and redshift space distortion) (Seo & Eisenstein, 2005; Eisenstein et al., 2007; Nishimichi et al., 2007; Smith et al., 2008a; Angulo et al., 2008; Sanchez et al., 2008; Seo et al., 2008; Shoji et al., 2009), despite the fact that the amplitude is distorted by non-linearities (see Figures 4.4, 4.9, 4.12, and 4.13). Therefore, BAOs provide a robust means to measure $D_A(z)$ and $H(z)$, and they have been used for determining $D_A^2 H^{-1}$ out to $z = 0.2$ from the SDSS main galaxy sample and 2dFGRS, as well as to $z = 0.35$ from the SDSS Luminous Red Galaxy (LRG) sample (Eisenstein et al., 2005; Percival et al., 2007); however, since they use only one standard ruler, the constraints on $D_A(z)$ and $H(z)$ from the BAO-only analysis are weaker than the full analysis (Shoji et al., 2009).
- (2) Use the *entire* shape of the power spectrum. This approach gives the best determination (i.e., the smallest error) of $D_A(z)$ and $H(z)$, as it uses all the standard rulers encoded in the galaxy power spectrum; however, one must understand the distortions of the shape of the power spectrum due to non-linear effects. The question is, “is the 3rd-order (or higher) PT good enough for correcting the key non-linear effects?”

In this paper we show, for the first time, that we can extract the distance scale using the 3rd-order PT galaxy power spectrum in real space. While we have not yet included the effects of redshift space distortion, this is a significant step towards extracting $D_A(z)$ and $H(z)$ from the entire shape of the power spectrum of galaxies. We shall address the effect of non-linear redshift space distortion in the future work.

4.4.1.2 Method: Measuring “Box Size” of the Millennium Simulation

In real space simulations (as opposed to redshift space ones), there is only one distance scale in the problem: the box size of the simulation, L_{box} , which is $L_{\text{box}}^{(\text{true})} = 500 \text{ Mpc}/h$ for the Millennium simulation. Then, “estimating the distance scale from the Millennium Simulation” becomes equivalent to “estimating L_{box} from the Millennium Simulation. Eq. (4.14) now leads:

$$P_{\text{obs}}(k, L_{\text{box}}) = \left(\frac{L_{\text{box}}}{L_{\text{box}}^{(\text{true})}} \right)^3 P_{\text{true}} \left(\frac{L_{\text{box}}^{(\text{true})}}{L_{\text{box}}} k \right). \quad (4.15)$$

As we estimate the variance of power spectrum from the observed power spectrum, we need to rescale the variance when the normalization of the observed power spectrum changes :

$$\sigma_{P_i}^2(L_{\text{box}}) = \left(\frac{L_{\text{box}}}{L_{\text{box}}^{(\text{true})}} \right)^6 \sigma_{P_i}^2(L_{\text{box}}^{(\text{true})}) \quad (4.16)$$

We estimate L_{box} using the likelihood function given by

$$\mathcal{L}(b_1, b_2, P_0, L_{\text{box}}) = \prod_{k_i < k_{\text{max}}} \frac{1}{\sqrt{2\pi\sigma_{P_i}^2(L_{\text{box}})}} \exp \left[-\frac{\{P_{\text{obs}}(k_i/\alpha) - P_g(k_i/\alpha)/\alpha^3\}^2}{2\sigma_P^2(k_i/\alpha)} \right], \quad (4.17)$$

where $\alpha = L_{\text{box}}/L_{\text{box}}^{(\text{true})}$.

The likelihood function, Eq. (4.17), still depends upon the bias parameters that we wish to eliminate. Therefore we marginalize the likelihood function over all the bias parameters with flat priors.⁹ We obtain (see also Appendix G):

$$\mathcal{L}(L_{\text{box}}) = \int_0^\infty db_1^2 \int_{-\infty}^\infty db_2 \int_{-\infty}^\infty dP_0 \mathcal{L}(b_1, b_2, P_0, L_{\text{box}}). \quad (4.18)$$

Hereafter, we shall simply call L_{box} as D for ‘distance scale’. D is closely related to the angular diameter distance, $D_A(z)$, and the expansion rate, $H(z)$, in real surveys. (See, §5.1.1)

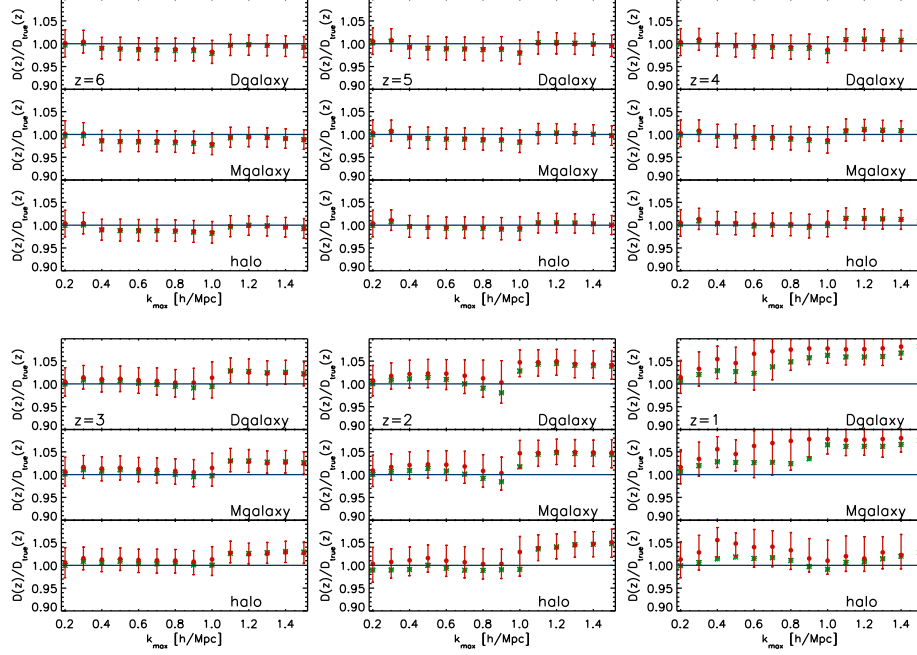


Figure 4.16: Distance scale extracted from the Millennium Simulation using the 3rd-order PT galaxy power spectrum given by Eq. (4.2), divided by the true value. The mean of the likelihood (*stars*), and the maximum likelihood values (*filled circles*) and the corresponding $1\text{-}\sigma$ intervals (*errorbars*), are shown as a function of maximum wavenumbers used in the fits, k_{max} . We find $D/D_{true} = 1$ to within the $1\text{-}\sigma$ errors from all the halo/galaxy catalogues (“halo,” “Mgalaxy,” and “Dgalaxy”) at all redshifts, provided that we use k_{max} estimated from the matter power spectra, $k_{max} = 0.15, 0.25, 1.0, 1.2, 1.3$, and 1.5 at $z = 1, 2, 3, 4, 5$, and 6 , respectively (see Table 4.1). Note that the errors on D do not decrease as k_{max} increases due to degeneracy between D and the bias parameters. See Figure (4.18) and (4.19) for further analysis.

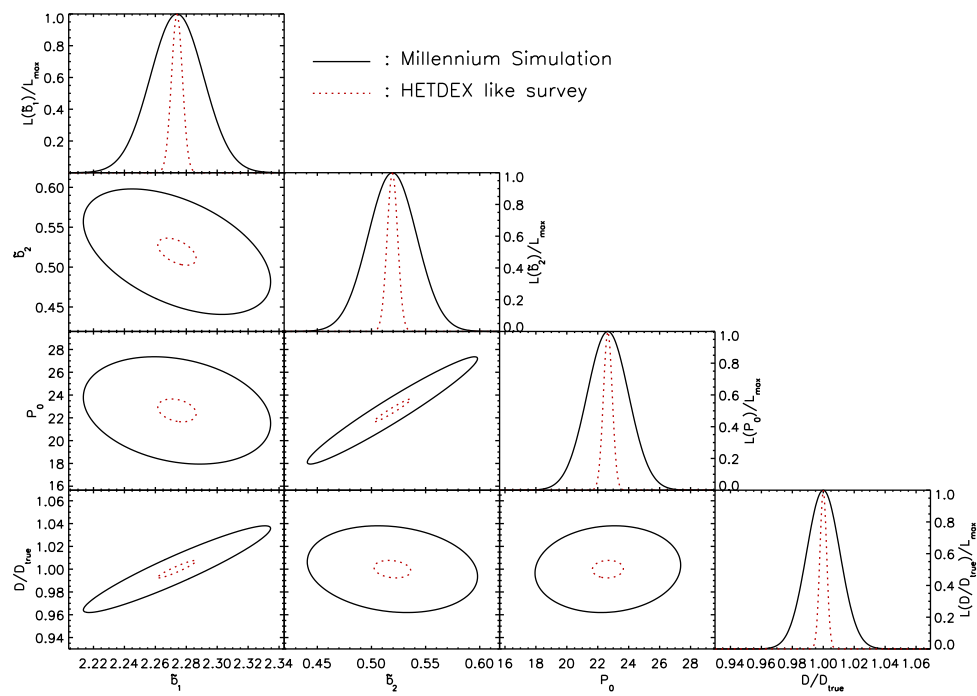


Figure 4.17: Same as Figure 4.8, but including the distance scale D/D_{true} .

4.4.1.3 Results: Unbiased Extraction of the distance scale from the Millennium Simulation

In Figure 4.16 we show $D(z)/D_{true}(z)$ estimated from the halo, Mgalaxy, and Dgalaxy catalogues at $z = 1, 2, 3, 4, 5$, and 6 . The maximum likelihood values (filled circles) and the corresponding $1-\sigma$ intervals (errorbars), as well as the mean of the likelihood (*stars*) are shown. We find $D(z)/D_{true}(z) = 1$ to within the $1-\sigma$ errors from *all of the halo/galaxy catalogues at all redshifts*, provided that we use $P_{obs}(k)$ only up to k_{max} that has been determined unambiguously from the matter power spectrum (see Table 4.1). Not only does this provide a strong support for the validity of Eq. (4.2), but also it provides a practical means for extracting D from the full shape of the observed galaxy power spectra.

Despite a small volume of the Millennium Simulation and the use of flat priors on the bias parameters upon marginalization, we could determine D to about 2.5% accuracy.

In addition, we also find that the error on D hardly decreases even though k_{max} increases. It is because of the degeneracy between D and the bias parameters. In order to see how strongly degenerate they are, we calculate correlations between pairs of parameters $(b_1, b_2, P_0, D/D_{true})$ by the Fisher information matrix from Eq. (4.9).

Figure 4.17 shows both one-dimensional marginalized constraints and two-dimensional joint marginalized constraints of $2-\sigma$ range (95.45% CL) for the bias parameters and the distance scale. This figure indicates that when we include the distance scale, the correlations between bias parameters become milder. It is mainly due to the correlation between the distance scale and b_1 making the constraint on b_1 weaker. On the other hand, the one-dimensional marginalized likelihood functions for b_2 and P_0 are hardly changed. The remaining degeneracies are those between (b_2, P_0) and $(b_1, D/D_{true})$. These degeneracies would be broken when we include the information from the bispectrum, as the bispectrum will measure b_1 and b_2 .

4.4.1.4 Optimal estimation of the distance scale

The constraint we find from the previous subsection will get better when we include the bispectrum, as the reduced bispectrum provides independent and strong constraints on b_1 and b_2 (Sefusatti et al., 2006).

⁹Note that this is the most conservative analysis one can do. In reality we can use the bispectrum for measuring b_1 and b_2 , which would give appropriate priors on them (see § 4.3.2.4). We shall report on the results from this analysis elsewhere.

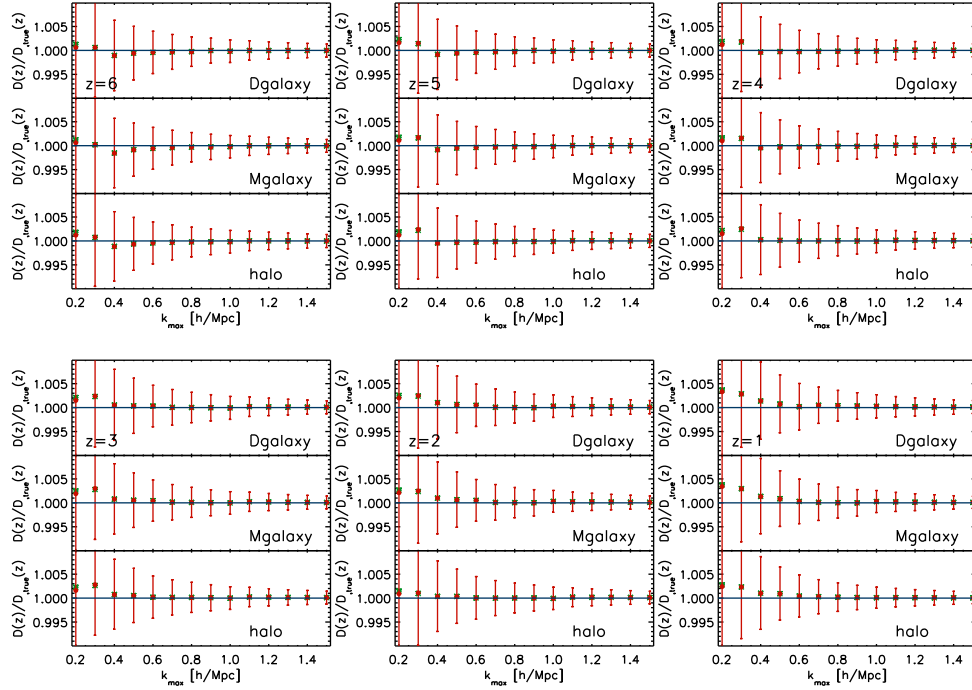


Figure 4.18: Same as Figure 4.16, but with b_1 and b_2 fixed at the best-fitting values. The $1\text{-}\sigma$ ranges for D are 1.5% and 0.15% for $k_{max} = 0.2 \text{ h/Mpc}$ and $k_{max} = 1.5 \text{ h/Mpc}$, respectively. The errors on D decrease as k_{max} increases, but the scaling is still milder than $1/\sqrt{\sum_{k < k_{max}} N_k}$.

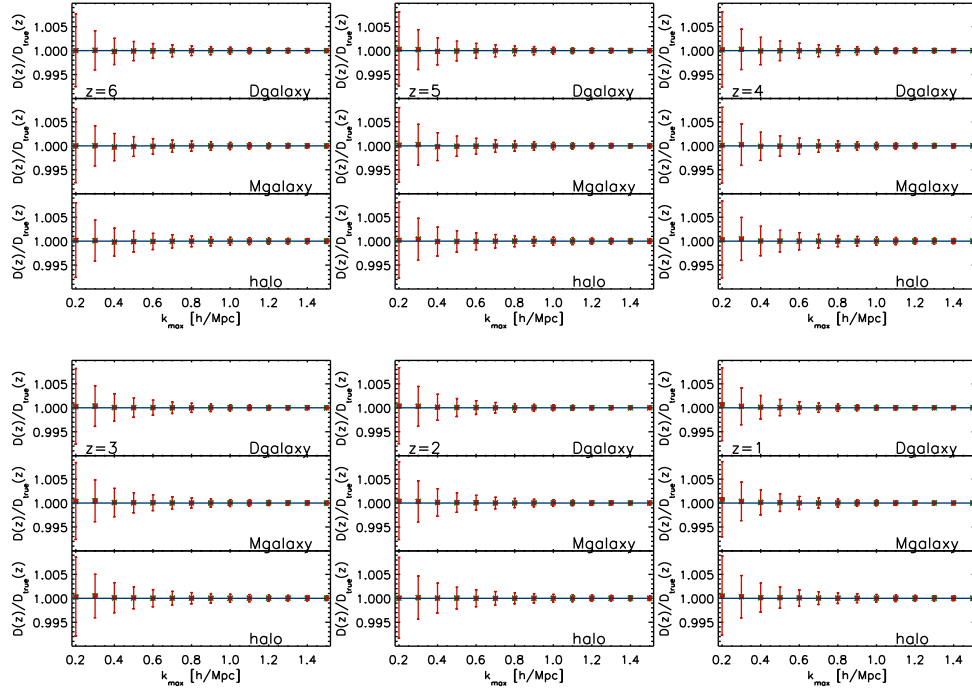


Figure 4.19: Same as Figure 4.16, but with b_1 , b_2 and P_0 fixed at the best-fitting values. The $1\text{-}\sigma$ ranges for D are 0.8% and 0.05% for $k_{\text{max}} = 0.2 h/\text{Mpc}$ and $k_{\text{max}} = 1.5 h/\text{Mpc}$, respectively. The errors on D decrease as k_{max} increases as $1/\sqrt{\sum_{k < k_{\text{max}}} N_k}$.

How much will it be better? First, let us assume that we know the exact values of b_1 and b_2 . In this case, we get the error on D by marginalizing only over P_0 while setting b_1 and b_2 to be the best-fitting values, i.e.

$$\mathcal{L}^{\text{fix } b_1 b_2}(D) = \int_{-\infty}^{\infty} dP_0 \mathcal{L}(b_1^{\text{bf}}, b_2^{\text{bf}}, P_0, D) \quad (4.19)$$

where b_1^{bf} and b_2^{bf} denote the best-fitting values of b_1 and b_2 for each k_{max} , respectively. In Figure 4.18, we show D/D_{true} estimated from Eq. (4.19). This figure shows that we can extract D to about 1.5% accuracy even for the low $k_{\text{max}} = 0.2 \text{ h/Mpc}$, and the error decreases further to 0.15% for $k_{\text{max}} = 1.5 \text{ h/Mpc}$. Note that the uncertainties on D/D_{true} decrease as k_{max} increases as expected. The reason is because fixing b_1 and b_2 breaks the degeneracy between them and the distance scale.

In reality, the bias parameters estimated from the bispectrum have finite errors, and thus the accuracy of extracting D will be somewhere in between Figure 4.16 and Figure 4.18. The result of the full analysis including both power spectrum and bispectrum of Millennium Simulation will be reported elsewhere.

In the ideal situation where we completely understand the complicated halo/galaxy formation, we may be able to calculate the three bias parameters from the first principle. This *ideal* determination of bias parameters will provide more accurate constraints on the distance scale D . In this case, we get the likelihood function by fixing all the bias parameters to their best-fitting values :

$$\mathcal{L}^{\text{fix bias}}(D) = \mathcal{L}(b_1^{\text{bf}}, b_2^{\text{bf}}, P_0^{\text{bf}}, D) \quad (4.20)$$

By knowing all the bias parameters, we can extract the distance scale D to 0.8% accuracy for $k_{\text{max}} = 0.2 \text{ h/Mpc}$. The error decreases further to 0.05% for $k_{\text{max}} = 1.5 \text{ h/Mpc}$. (See Figure (4.19))

4.4.1.5 Forecast for a HETDEX-like survey

The planned future surveys would cover a larger volume than the Millennium Simulation. Also, since the real surveys would be limited by their continuum/flux sensitivity, they would not be able to detect all galaxies that were resolved in the Millennium Simulation. In this subsection we explore how the constraints would be affected by the volume and the number of objects.

To simulate the mock data, we take a simplified approach: we take our best-fitting power spectrum at $z = 3$, i.e., Eq. (4.2) fit to the power spectrum of MPA halos in the

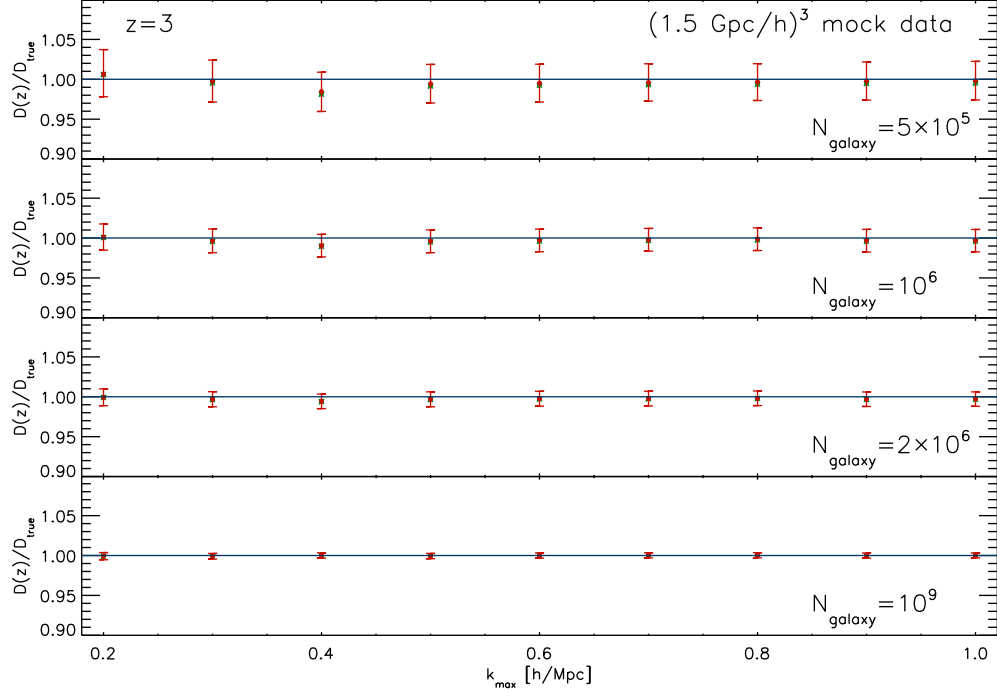


Figure 4.20: Projected constraints on D at $z = 3$ from a HETDEX-like survey with the survey volume of $(1.5 \text{ Gpc/h})^3$. We have used the best-fitting 3rd-order PT power spectrum of MPA halos in the Millennium Simulation for generating a mock simulation data. We show the results for the number of objects of $N_{\text{galaxy}} = 2 \times 10^5$, 10^6 , 2×10^6 , and 10^9 , from the top to bottom panels, respectively, for which we find the projected 1- σ errors of 2.5%, 1.5%, 1%, and 0.3%, respectively.

Millennium Simulation at $z = 3$, and add random Gaussian noise to it with the standard deviation given by Eq. (4.7). To compute the standard deviation we need to specify the survey volume, which determines the fundamental wavenumber, Δk , as $\Delta k = 2\pi/V_{survey}^{1/3}$. We use the volume that would be surveyed by the HETDEX survey (Hill et al., 2004), $V_{survey} = (1.5 \text{ Gpc}/h)^3$, which is 27 times as large as the volume of the Millennium Simulation. We then vary the number of galaxies, N_{galaxy} , which determines the shot noise as $P_{shot} = 1/n = V_{survey}/N_{galaxy}$. We have generated only one realization, and repeated the same analysis as before to extract D_A from the mock HETDEX data.

In Figure 4.20 we show D/D_{true} as a function of k_{max} and N_g . For $N_{galaxy} = 10^9$, which gives the same number density as the Millennium Simulation, the projected error on D is 0.3%, or 8 times better than the original result presented in Figure 4.16. Since the volume is 27 times bigger, the statistics alone would reduce the error by a factor of about 5.

The other factor of about 1.5 comes from the fact that the variance of the distance scale estimated from the Millennium Simulation lies on the tail of the distribution of the variance of the distance scale, (See, appendix C) while the error estimated from the HETDEX volume mock is close to the peak of PDF of the variance.

However, real surveys will not get as high the number density as the Millennium Simulation. For example, the HETDEX survey will detect about one million Ly α emitting galaxies, i.e., $N_{galaxy} = 10^6$. In Figure 4.20 we show that the errors on D increase from 0.3% for $N_{galaxy} = 10^9$ to 1%, 1.5%, and 3% for $N_{galaxy} = 2 \times 10^6$, 10^6 , and 2×10^5 , respectively.

Finally, we note that these forecasts are not yet final, as we have not included the effect of non-linear redshift space distortion. Also, eventually one needs to repeat this analysis using the “super Millennium Simulation” with a bigger volume.

4.5 Discussion and Conclusions

Two main new results that we have presented in this paper are:

- The 3rd-order PT galaxy power spectrum given by Eq. (4.2), which is based upon the assumption that the number density of galaxies at a given location is a local function of the underlying matter density at the same location (Fry & Gaztanaga, 1993) plus stochastic noise (McDonald, 2006), fits the halo as well as galaxy power spectra estimated from the Millennium Simulation at high redshifts, $1 \leq z \leq 6$, up to the maximum wavenumber, k_{max} , that has been determined from the matter power spectrum.

- When 3 galaxy bias parameters, b_1 , b_2 , and P_0 , are marginalized over, the 3rd-order PT galaxy power spectrum fit to the Millennium Simulation yields the correct (unbiased) distance scale to within the statistical error of the simulation, $\sim 3\%$.

These results suggest that the 3rd-order PT provides us with a practical means to extract the cosmological information from the observed galaxy power spectra at high redshifts, i.e., $z > 1$, accurately.

We would like to emphasize that our approach does not require simulations to calibrate the model. The 3rd-order PT is based upon the solid physical framework, and the only assumption made for the galaxy formation is that it is a local process, at least on the scales where the 3rd-order PT is valid, i.e., $k < k_{max}$. The only serious drawback so far is that the 3rd-order PT breaks down at low redshifts, and thus it cannot be applied to the current generation of survey data such as 2dFGRS and SDSS. However, the planned future high- z surveys would benefit immensely from the PT approach.

The practical application of our approach may proceed as follows:

- (1) Measure the galaxy power spectra at various redshifts. When we have N redshift bins, the number of bias parameters is $3N$, as the bias parameters evolve with z .
- (2) Calculate $k_{max}(z)$ from the condition, $\Delta_m^2(k_{max}, z) = 0.4$, where $\Delta_m^2(k, z) = k^3 P_{\delta\delta}(k, z)/(2\pi^2)$ is computed from the fiducial cosmology, e.g., the WMAP 5-year best-fitting parameters. The results should not be sensitive to the exact values of $k_{max}(z)$.
- (3) Fit Eq. (4.2) to the observed galaxy spectra up to $k_{max}(z)$ at all z simultaneously for extracting the cosmological parameters.

In addition to this, we should be able to improve upon the accuracy of parameter determinations by including the bispectrum as well, as the bispectrum basically fixes b_1 and b_2 (Sefusatti & Komatsu, 2007). Therefore, the step (3) may be replaced by

- (3') Fit Eq. (4.2) to the observed galaxy spectra up to $k_{max}(z)$, and fit the PT bispectrum to the observed galaxy bispectra up to the same $k_{max}(z)$, at all z simultaneously for extracting the cosmological parameters.

We are currently performing a joint analysis of the galaxy power spectra and bispectra on the Millennium Simulation. The results will be reported elsewhere.

There are limitations in our present study, however. First, a relatively small volume of the Millennium Simulation does not allow us to make a precision test of the 3rd-order PT. Also, this limitation does not allow us to study constraints on more than one cosmological parameter. We have picked D as the representative example because this parameter seems the most interesting one in light of the future surveys whose primary goal is to constrain the properties of dark energy. In the future we must use larger simulations to show convincingly that the bias in cosmological parameters is much lower than 1% level. Second, we have found that, due to the limited statistics of a small volume and the smaller k_{max} due to stronger non-linearities, the bias parameters are not determined very well from the galaxy power spectra alone at $z \leq 3$. This issue should disappear by including the bispectrum in the joint analysis. Last and foremost, our study has been restricted to the real space power spectra: we have not addressed the non-linearities in redshift space distortion. This is a subject of the future study.

Table 4.6: Mass dependence of non-linear halo bias parameters (MPA halos)

z	M_{\min} (M_{\odot}/h)	M_{\max} (M_{\odot}/h)	b_1	b_2	P_0 ($[h/\text{Mpc}]^3$)	b_1^{ST}	b_2^{ST}
6	1.7E+10	5.0E+10	3.19±0.01	1.28±0.03	88.76±2.97	2.96	0.93
	5.0E+10	1.0E+11	3.90±0.02	1.91±0.04	288.18±8.02	3.52	1.36
	1.0E+11	5.0E+11	4.66±0.03	3.04±0.05	1029.19±18.84	4.41	2.28
	5.0E+11	1.0E+12	6.41±0.14	5.76±0.21	6910.17±200.74	5.95	3.59
5	1.7E+10	5.0E+10	2.55±0.01	0.71±0.03	31.51±2.06	2.41	0.48
	5.0E+10	1.0E+11	3.09±0.01	1.19±0.04	120.84±5.69	2.84	0.81
	1.0E+11	5.0E+11	3.78±0.02	1.79±0.04	402.11±12.40	3.55	1.48
	5.0E+11	1.0E+12	5.14±0.07	3.55±0.11	2805.48±94.53	4.71	2.44
4	1.7E+10	5.0E+10	2.08±0.01	0.38±0.04	10.90±1.19	2.01	0.15
	5.0E+10	1.0E+11	2.51±0.01	0.66±0.04	42.34±3.52	2.33	0.40
	1.0E+11	5.0E+11	3.05±0.01	1.08±0.04	161.22±8.11	2.90	0.92
	5.0E+11	1.0E+12	3.80±0.05	-4.08±0.09	3431.19±64.81	3.79	1.77
3	1.7E+10	5.0E+10	1.39±0.01	-1.83±0.05	241.59±9.58	1.47	-0.25
	5.0E+10	1.0E+11	1.75±0.01	0.11±0.05	2.48±0.29	1.67	-0.10
	1.0E+11	5.0E+11	2.09±0.01	0.35±0.04	20.95±3.22	2.04	0.19
	5.0E+11	1.0E+12	2.78±0.02	0.82±0.06	171.31±21.03	2.57	0.60
2	1.7E+10	5.0E+10	1.01±0.05	-1.98±0.68	373.60±149.24	1.11	-0.46
	5.0E+10	1.0E+11	1.14±0.07	-2.30±0.63	627.69±204.69	1.23	-0.40
	1.0E+11	5.0E+11	1.31±0.08	-2.34±0.63	869.30±272.06	1.44	-0.28
	5.0E+11	1.0E+12	1.62±0.11	-2.53±0.69	1566.40±476.07	1.75	-0.05
1	1.7E+10	5.0E+10	0.68±0.09	-3.12±1.46	1315.40±447.14	0.86	-0.58
	5.0E+10	1.0E+11	0.75±0.10	-3.24±1.46	1699.76±571.75	0.92	-0.56
	1.0E+11	5.0E+11	0.85±0.09	-2.80±1.65	1783.07±683.57	1.02	-0.53
	5.0E+11	1.0E+12	0.99±0.11	-2.82±1.95	2443.76±972.16	1.17	-0.47

z : redshift

M_{\min} : minimum mass for a given bin

M_{\max} : maximum mass for a given bin

b_1, b_2, P_0 : non-linear bias parameters

b_1^{ST}, b_2^{ST} : bias parameters from the Sheth-Tormen model, $b_2^{ST}=b_2^{ST}/b_1$

Caution: We estimate 1- σ ranges for the low redshift ($z \leq 3$) only for the peak which involves the maximum likelihood value. If two peaks in maginalized likelihood function are blended, we use only unblended side of the peak to estimate the 1- σ range.

Table 4.7: Mass dependence of non-linear galaxy bias parameters (MPA galaxies)

z	M_{\min} (M_{\odot}/h)	M_{\max} (M_{\odot}/h)	b_1	b_2	P_0 ($[h/\text{Mpc}]^3$)	b_1^{ST}	b_2^{ST}
6	1.7E+10	5.0E+10	3.37±0.01	1.50±0.03	136.39±3.69	2.91	0.82
	5.0E+10	1.0E+11	3.96±0.02	2.00±0.04	325.38±8.44	3.49	1.31
	1.0E+11	5.0E+11	4.69±0.03	3.09±0.05	1078.72±19.25	4.23	2.01
	5.0E+11	1.0E+12	6.43±0.14	5.79±0.20	7046.28±201.94	5.89	3.49
5	1.7E+10	5.0E+10	2.77±0.01	0.93±0.03	63.17±2.99	2.38	0.40
	5.0E+10	1.0E+11	3.16±0.01	1.27±0.04	144.20±6.16	2.82	0.77
	1.0E+11	5.0E+11	3.81±0.02	1.84±0.04	432.51±12.80	3.41	1.28
	5.0E+11	1.0E+12	5.15±0.07	3.60±0.11	2897.95±95.17	4.67	2.37
4	1.7E+10	5.0E+10	2.33±0.01	0.58±0.03	32.25±2.25	1.98	0.11
	5.0E+10	1.0E+11	2.59±0.01	0.74±0.04	56.91±4.08	2.32	0.37
	1.0E+11	5.0E+11	3.09±0.02	1.13±0.04	179.81±8.52	2.79	0.77
	5.0E+11	1.0E+12	3.83±0.05	-4.09±0.09	3507.05±64.85	3.76	1.71
3	1.7E+10	5.0E+10	1.62±0.01	-2.07±0.05	431.79±12.04	1.45	-0.22
	5.0E+10	1.0E+11	1.84±0.01	0.19±0.04	7.04±1.04	1.66	-0.10
	1.0E+11	5.0E+11	2.14±0.01	0.38±0.04	27.64±3.67	1.96	0.12
	5.0E+11	1.0E+12	2.80±0.02	0.84±0.06	191.24±21.65	2.55	0.57
2	1.7E+10	5.0E+10	1.26±0.07	-2.09±0.66	683.11±240.40	1.10	-0.37
	5.0E+10	1.0E+11	1.21±0.08	-2.35±0.62	738.09±231.05	1.22	-0.38
	1.0E+11	5.0E+11	1.35±0.09	-2.32±0.63	919.65±288.79	1.40	-0.29
	5.0E+11	1.0E+12	1.65±0.11	-2.50±0.69	1602.17±480.23	1.74	-0.06
1	1.7E+10	5.0E+10	0.91±0.11	-2.96±1.59	2344.13±802.55	0.86	-0.43
	5.0E+10	1.0E+11	0.79±0.11	-3.28±1.51	1956.42±657.89	0.92	-0.53
	1.0E+11	5.0E+11	0.87±0.10	-2.88±1.63	1964.64±738.71	1.00	-0.51
	5.0E+11	1.0E+12	1.01±0.11	-2.80±2.01	2550.21±1007.23	1.16	-0.46

z : redshift

M_{\min} : minimum mass for a given bin

M_{\max} : maximum mass for a given bin

b_1, b_2, P_0 : non-linear bias parameters

b_1^{ST}, b_2^{ST} : bias parameters from the Sheth-Tormen model, $b_2^{ST}=b_2^{ST}/b_1$

Caution: We estimate 1- σ ranges for the low redshift ($z \leq 3$) only for the peak which involves the maximum likelihood value. If two peaks in maginalized likelihood function are blended, we use only unblended side of the peak to estimate the 1- σ range.

Chapter 5

Primordial non-Gaussianity, scale-dependent bias, and the bispectrum of galaxies

Why study non-Gaussianity¹? For many years it was recognized that the simple inflationary models based upon a single slowly-rolling scalar field would predict nearly Gaussian primordial fluctuations. In particular, when we parametrize the magnitude of non-Gaussianity in the primordial curvature perturbations ζ , which gives the observed temperature anisotropy in the Cosmic Microwave Background (CMB) in the Sachs–Wolfe limit as $\Delta T/T = -\zeta/5$, using the so-called non-linear parameter f_{NL} (Komatsu & Spergel, 2001) as $\zeta(\mathbf{x}) = \zeta_L(\mathbf{x}) + (3f_{\text{NL}}/5)\zeta_L^2(\mathbf{x})$, then the bispectrum of ζ is given by² $B_\zeta(k_1, k_2, k_3) = (6f_{\text{NL}}/5)[P_\zeta(k_1)P_\zeta(k_2) + (2 \text{ cyclic terms})]$, where $P_\zeta(k) \propto k^{n_s-4}$ is the power spectrum of ζ and n_s is the tilt of the power spectrum, constrained as $n_s = 0.963 \pm 0.012$ by the WMAP 7-year data (Komatsu et al., 2010). This form of the bispectrum has the maximum signal in the so-called squeezed triangle for which $k_3 \ll k_2 \approx k_1$ (Babich et al., 2004). In this limit we obtain

$$B_\zeta(k_1, k_1, k_3 \rightarrow 0) = \frac{12}{5} f_{\text{NL}} P_\zeta(k_1) P_\zeta(k_3). \quad (5.1)$$

The earlier calculations showed that f_{NL} from single-field slow-roll inflation would be of order the slow-roll parameter, $\epsilon \sim 10^{-2}$ (Salopek & Bond, 1990; Falk et al., 1993; Gangui et al., 1994). However, it is not until recent that it is finally realized that the coefficient of $P_\zeta(k_1)P_\zeta(k_3)$ from the simplest single-field slow-roll inflation with the canonical kinetic term in the squeezed limit is given precisely by Maldacena (2003); Acquaviva et al. (2003)

$$B_\zeta(k_1, k_1, k_3 \rightarrow 0) = (1 - n_s) P_\zeta(k_1) P_\zeta(k_3). \quad (5.2)$$

Comparing this result with the form predicted by the f_{NL} model, one obtains $f_{\text{NL}} = (5/12)(1 - n_s)$.

¹A significant fraction of this chapter was published in Jeong, D. & Komatsu, E. 2009, *Astrophys. J.*, 703, 1230.

²Definition of the bispectrum in terms of Fourier coefficients of ζ is $\langle \zeta_{\mathbf{k}_1} \zeta_{\mathbf{k}_2} \zeta_{\mathbf{k}_3} \rangle = (2\pi)^3 \delta(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) B_\zeta(k_1, k_2, k_3)$. Throughout this paper we shall order k_i such that $k_3 \leq k_2 \leq k_1$.

Perhaps, the most important theoretical discovery regarding primordial non-Gaussianity from inflation over the last few years is that, not only models with the canonical kinetic term, but *all* single-inflation models predict the bispectrum in the squeezed limit given by Eq. (5.2), regardless of the form of potential, kinetic term, slow-roll, or initial vacuum state (Creminelli & Zaldarriaga, 2004; Seery & Lidsey, 2005; Chen et al., 2007; Cheung et al., 2008). Therefore, the prediction from all single-field inflation models is $f_{\text{NL}} = (5/12)(1 - n_s) = 0.017$ for $n_s = 0.96$. A convincing detection of f_{NL} well above this level is a breakthrough in our understanding of the physics of very early universe (Bartolo et al., 2004; Komatsu et al., 2009). The current limit from the WMAP 7-year data is $f_{\text{NL}} = 32 \pm 21$ (68% CL) (Komatsu et al., 2010).

There are many ways of measuring f_{NL} . The most popular method has been the bispectrum of CMB (Verde et al., 2000; Wang & Kamionkowski, 2000; Komatsu & Spergel, 2001; Komatsu et al., 2002, 2003) (also see Komatsu (2001) for a pedagogical review). The other methods include the trispectrum of CMB (Okamoto & Hu, 2002; Kogo & Komatsu, 2006).

The large-scale structure of the universe can also provide alternative ways of probing primordial non-Gaussianity through abundances and clustering properties of galaxies and clusters of galaxies (Lucchin & Matarrese, 1988; Matarrese et al., 2000; Sefusatti et al., 2007; LoVerde et al., 2008). However, as the large-scale structure of the universe is more non-linear than CMB, it was generally thought that CMB would be the most promising way of constraining primordial non-Gaussianity (Verde et al., 2000).

On the other hand, Sefusatti & Komatsu (2007) have shown that observations of the large-scale structure of the universe in a *high-redshift universe*, i.e., $z > 1$, can provide competitive limits on primordial non-Gaussianity, as the other non-linear effects are weaker in a high redshift universe. Specifically, they calculate the bispectrum of the three-dimensional distribution of galaxies, $B_g(k_1, k_2, k_3)$,³ on large scales as (see also Scoccimarro

³The bispectrum, the Fourier transform of the three-point correlation function, is defined as $\langle \delta(\mathbf{k}_1)\delta(\mathbf{k}_2)\delta(\mathbf{k}_3) \rangle \equiv (2\pi)^3 B(k_1, k_2, k_3) \delta^D(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3)$.

et al. (2004))

$$\begin{aligned}
& B_g(k_1, k_2, k_3, z) \\
&= 3b_1^3 f_{\text{NL}} \Omega_m H_0^2 \left[\frac{P_m(k_1, z)}{k_1^2 T(k_1)} \frac{P_m(k_2, z)}{k_2^2 T(k_2)} \frac{k_3^2 T(k_3)}{D(z)} + (2 \text{ cyclic}) \right] \\
&+ 2b_1^3 \left[F_2^{(s)}(\mathbf{k}_1, \mathbf{k}_2) P_m(k_1, z) P_m(k_2, z) + (2 \text{ cyclic}) \right] \\
&+ b_1^2 b_2 [P_m(k_1, z) P_m(k_2, z) + (2 \text{ cyclic})], \tag{5.3}
\end{aligned}$$

where H_0 and Ω_m are the present-day value of Hubble's constant and the matter density parameter, respectively, $P_m(k, z)$ is the power spectrum of linear matter density fluctuations, $D(z)$ is the linear growth factor, $T(k)$ is the linear transfer function whose limit is $T(k) \rightarrow 1$ as $k \rightarrow 0$, and $F_2^{(s)}(\mathbf{k}_1, \mathbf{k}_2)$ is a known mathematical function given in equation (2.32)

$$F_2^{(s)}(\mathbf{k}_1, \mathbf{k}_2) = \frac{5}{7} + \frac{\mathbf{k}_1 \cdot \mathbf{k}_2}{2k_1 k_2} \left(\frac{k_1}{k_2} + \frac{k_2}{k_1} \right) + \frac{2}{7} \left(\frac{\mathbf{k}_1 \cdot \mathbf{k}_2}{k_1 k_2} \right)^2. \tag{5.4}$$

This function vanishes in the squeezed limit, $\mathbf{k}_1 = -\mathbf{k}_2$ (the triangles with the maximum angle, i.e., π , between \mathbf{k}_1 and \mathbf{k}_2 , and $|\mathbf{k}_1| = |\mathbf{k}_2|$), and takes on the maximum value, $F_2^{(s)} = (\alpha + 1)^2 / (2\alpha) \geq 2$, in the opposite limit, $\mathbf{k}_1 = \alpha \mathbf{k}_2$ where $\alpha \geq 1$ (the triangles with the vanishing angle between \mathbf{k}_1 and \mathbf{k}_2).

Here, b_1 and b_2 are the linear and non-linear galaxy bias parameters, respectively, which relate the underlying matter density contrast, δ_m , to the galaxy density contrast, δ_g , as (Fry & Gaztanaga, 1993)

$$\delta_g(\mathbf{x}) = b_1 \delta_m(\mathbf{x}) + \frac{b_2}{2} [\delta_m^2(\mathbf{x}) - \sigma^2] + \dots, \tag{5.5}$$

where $\sigma^2 \equiv \langle \delta_m^2 \rangle$, which ensures $\langle \delta_g \rangle = 0$.

The last two terms in Eq. (5.3) are the well-known results for Gaussian initial conditions (see Bernardeau et al., 2002, for a review), whereas the first term is the effect of the primordial non-Gaussianity of the “local type”.

However, Sefusatti & Komatsu's equation, Eq. (5.3), may require modifications, in light of recent analytical (Dalal et al., 2008; Matarrese & Verde, 2008; Slosar et al., 2008; Afshordi & Tolley, 2008; Taruya et al., 2008; McDonald, 2008) and numerical (Dalal et al., 2008; Desjacques et al., 2009; Pillepich et al., 2008; Grossi et al., 2009) studies of the effects of primordial non-Gaussianity on the galaxy power spectrum. These studies have discovered an unexpected signature of primordial non-Gaussianity in the form of a *scale-dependent*

galaxy bias⁴, i.e., $P_g(k, z) = b_1^2(z)P_m(k, z) \rightarrow [b_1(z) + \Delta b(k, z)]^2 P_m(k, z)$, where

$$\Delta b(k, z) = \frac{3(b_1(z) - 1)f_{\text{NL}}\Omega_m H_0^2 \delta_c}{D(z)k^2 T(k)}, \quad (5.6)$$

and $\delta_c \simeq 1.68$ is the threshold linear density contrast for a spherical collapse of an overdensity region. The k^2 factor in the denominator of $\Delta b(k)$ shows that this effect is important only on very large scales. Highly biased tracers are more sensitive to f_{NL} .

Then, several questions arise: can we still use Eq. (5.3) for the bispectrum? Should we replace b_1 by $b_1 + \Delta b(k)$? Does the first line in Eq. (5.3) somehow give the same correction as $\Delta b(k)$? How about b_2 ? We are going to address these questions in this paper.

5.1 Bispectrum of Dark Matter Halos

In this section, we derive the galaxy bispectrum for non-Gaussian initial conditions by using two different methods. In §5.1.1, we shall use the “functional integration method” for computing n -point correlation functions of *peaks* of the cosmological density fluctuations (Politzer & Wise, 1984; Grinstein & Wise, 1986). In §5.1.2, we shall present an alternative derivation of the same result by using a local bias assumption.

5.1.1 Matarrese-Lucchin-Bonometto (MLB) method

We shall use the Matarrese-Lucchin-Bonometto (MLB) formula (Matarrese et al., 1986) which allows one to calculate the n -point correlation functions of peaks for non-Gaussian initial conditions. This approach is especially well suited for our purposes, as Matarrese & Verde (2008) have applied the MLB formula to compute the scale-dependent bias of the galaxy power spectrum. We shall apply the MLB formula to compute the galaxy bispectrum for general non-Gaussian initial condition.

We study the three-point correlation function of the spatial distribution of dark matter halos. Let us consider the probability of finding three halos within three arbitrary volume elements: dV_1 , dV_2 , and dV_3 , which are at \mathbf{x}_1 , \mathbf{x}_2 , and \mathbf{x}_3 , respectively, as (Peebles, 1980)

$$\begin{aligned} P(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) &= \bar{n}^3 [1 + \xi_h(x_{12}) + \xi_h(x_{23}) + \xi_h(x_{31}) \\ &\quad + \zeta_h(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)] dV_1 dV_2 dV_3, \end{aligned} \quad (5.7)$$

⁴For the derivation of the scale dependent bias, see Appendix I.3.

where $x_{ij} \equiv |\mathbf{x}_i - \mathbf{x}_j|$, and ξ_h and ζ_h are the two- and three-point correlation functions of halos, respectively.

The next step is to relate the correlation functions of halos, ξ_h and ζ_h , to those of the underlying matter distribution function. The locations of halos coincide with those of the peaks of the matter density fluctuations; thus, one can compute ξ_h and ζ_h by computing the correlation functions of peaks above a certain threshold, above which the peaks collapse into halos.

We shall assume that halos would be formed in the region where the *smoothed* linear density contrast exceeds δ_c . For a spherical collapse in an Einstein-de Sitter universe $\delta_c = 3(12\pi)^{2/3}/20 \simeq 1.68$, and one can find other values in the ellipsoidal collapse in arbitrary cosmological models (see, e.g., Cooray & Sheth, 2002, for a review). The mass of halos is determined by the smoothing radius, R , i.e., $M = (4\pi/3)\rho_m R^3$, where ρ_m is the average mass density of the universe. The smoothed density contrast, δ_R , is related to the underlying mass fluctuations, δ_m , as $\delta_R(\mathbf{x}) = \int d^3\mathbf{x}' W_R(|\mathbf{x} - \mathbf{x}'|)\delta_m(\mathbf{x}')$, where $W_R(x)$ is a smoothing function. We shall use a top-hat filter with radius R for $W_R(x)$.

Using the MLB formula⁵, we find

$$\begin{aligned}
& 1 + \xi_h(x_{12}) + \xi_h(x_{23}) + \xi_h(x_{31}) + \zeta_h(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) \\
= & \exp \left[\frac{1}{2} \frac{\nu^2}{\sigma_R^2} \sum_{i \neq j} \xi_R^{(2)}(x_{ij}) + \sum_{n=3}^{\infty} \left\{ \sum_{m_1=0}^n \sum_{m_2=0}^{n-m_1} \frac{\nu^n \sigma_R^{-n}}{m_1! m_2! m_3!} \right. \right. \\
& \times \xi_R^{(n)} \left(\begin{array}{ccc} \mathbf{x}_1, \dots, \mathbf{x}_1, & \mathbf{x}_2, \dots, \mathbf{x}_2, & \mathbf{x}_3, \dots, \mathbf{x}_3 \\ m_1 \text{ times} & m_2 \text{ times} & m_3 \text{ times} \end{array} \right) \\
& \left. \left. - 3 \frac{\nu^n \sigma_R^{-n}}{n!} \xi_R^{(n)} \left(\begin{array}{c} \mathbf{x}, \dots, \mathbf{x} \\ n \text{ times} \end{array} \right) \right\} \right], \tag{5.8}
\end{aligned}$$

where $m_3 \equiv n - m_1 - m_2$, $\nu \equiv \delta_c/\sigma_R$, σ_R^2 is the variance of matter density fluctuations smoothed by a top-hat filter with radius R , and $\xi_R^{(n)}$ denotes the connected parts of the n -point correlation functions of the underlying matter density fields smoothed by a top-hat filter of radius R . Here, we have assumed that we are dealing with high density peaks, i.e., $\nu \gg 1$, which are equivalent to highly biased galaxies, $b_1 \gg 1$.

As $\xi_R^{(n)} \ll 1$ on the large scales that we are interested in, we expand the exponential in Eq. (5.8). We keep the terms up to the four-point function, as this term provides the

⁵MLB formula is summarized in Appendix K.

dominant contribution to the three-point function. We find

$$\begin{aligned}
\zeta_h(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) &= \frac{\nu^3}{\sigma_R^3} \xi_R^{(3)}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) \\
&+ \frac{\nu^4}{\sigma_R^4} \left[\xi_R^{(2)}(x_{12}) \xi_R^{(2)}(x_{23}) + (2 \text{ cyclic}) \right] \\
&+ \frac{\nu^4}{2\sigma_R^4} \left[\xi_R^{(4)}(\mathbf{x}_1, \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) + (2 \text{ cyclic}) \right]. \tag{5.9}
\end{aligned}$$

The bispectrum of halos in Lagrangian space, $B_h^L(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)$, is the Fourier transform of $\zeta_h(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$:

$$\begin{aligned}
&B_h^L(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) \\
&= \frac{\nu^3}{\sigma_R^3} \left[B_R(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) + \frac{\nu}{\sigma_R} \{P_R(k_1)P_R(k_2) + (2 \text{ cyclic})\} \right. \\
&\quad \left. + \frac{\nu}{2\sigma_R} \int \frac{d^3q}{(2\pi)^3} T_R(\mathbf{q}, \mathbf{k}_1 - \mathbf{q}, \mathbf{k}_2, \mathbf{k}_3) + (2 \text{ cyclic}) \right], \tag{5.10}
\end{aligned}$$

where T_R is the trispectrum, Fourier transform of $\xi_R^{(4)}$. Here, we call B_h^L the Lagrangian space bispectrum, as it relates the halo over-density to the *initial* matter overdensity with its amplitude extrapolated to the present epoch. If we assume that the halos move in the same way as matter, the observed bispectrum in Eulerian space, B_h , would be the same expression with Eq. (5.10), except for the coefficients:

$$\begin{aligned}
&B_h(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) \\
&= b_1^3 \left[B_R(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) + \frac{b_2}{b_1} \{P_R(k_1)P_R(k_2) + (2 \text{ cyclic})\} \right. \\
&\quad \left. + \frac{\delta_c}{2\sigma_R^2} \int \frac{d^3q}{(2\pi)^3} T_R(\mathbf{q}, \mathbf{k}_1 - \mathbf{q}, \mathbf{k}_2, \mathbf{k}_3) + (2 \text{ cyclic}) \right]. \tag{5.11}
\end{aligned}$$

Here, b_1 is the so-called linear Eulerian bias parameter, $b_1 = 1 + \nu/\sigma_R$, and $b_2 = (\nu/\sigma_R)^2$ is the non-linear bias parameter ⁶.

5.1.2 Alternative derivation

In this section, we present an alternative derivation of the galaxy bispectrum, Eq. (5.11). On large enough scales, we may approximate the relation between the galaxy dis-

⁶Note that these expressions, $b_1 = 1 + \nu/\sigma_R$ and $b_2 = (\nu/\sigma_R)^2$ agree with the linear halo bias parameter derived by Mo & White (1996); Mo et al. (1997), $b_1 = 1 + b_1^L$ and $b_2 = (\nu^2 - 3)/\sigma_R^2 + 8b_1^L/21$ for high density peaks, $\nu \gg 1$. Here, $b_1^L = (\nu - 1/\nu)/\sigma_R$ is the Lagrangian bias parameter.

tribution and the underlying density fluctuation as a local function. We then Taylor-expand this local function in a power series of δ_m (see Eq. (5.5)).

When computing the correlation functions of halos of a given mass M , we may smooth the matter density field with the same filter over the corresponding length scale R , $W_R(|\mathbf{x} - \mathbf{x}'|)$, which was defined in the previous section. We then Taylor-expand δ_g in a power series of the smoothed density field, $\delta_R(\mathbf{x})$, as

$$\delta_g(\mathbf{x}) = b_1 \delta_R(\mathbf{x}) + \frac{b_2}{2} [\delta_R^2(\mathbf{x}) - \sigma_R^2] + \dots \quad (5.12)$$

In Fourier space, one finds

$$\begin{aligned} \delta_g(\mathbf{k}) &= b_1 \delta_R(\mathbf{k}) \\ &+ \frac{b_2}{2} \left[\int \frac{d^3 q}{(2\pi)^3} \delta_R(\mathbf{k} - \mathbf{q}) \delta_R(\mathbf{q}) - \sigma_R^2 \delta^D(\mathbf{k}) \right] \\ &+ \dots, \end{aligned} \quad (5.13)$$

where δ^D is the Dirac delta function. We calculate the bispectrum of galaxies directly from Eq. (5.13):

$$\begin{aligned} &\langle \delta_g(\mathbf{k}_1) \delta_g(\mathbf{k}_2) \delta_g(\mathbf{k}_3) \rangle \\ &= b_1^3 \langle \delta_R(\mathbf{k}_1) \delta_R(\mathbf{k}_2) \delta_R(\mathbf{k}_3) \rangle \\ &+ \frac{b_1^2 b_2}{2} \left[\int \frac{d^3 q}{(2\pi)^3} \langle \delta_R(\mathbf{k}_1 - \mathbf{q}) \delta_R(\mathbf{q}) \delta_R(\mathbf{k}_2) \delta_R(\mathbf{k}_3) \rangle \right. \\ &\quad \left. - \sigma_R^2 \delta^D(\mathbf{k}_1) \langle \delta_R(\mathbf{k}_2) \delta_R(\mathbf{k}_3) \rangle + (2 \text{ cyclic}) \right]. \end{aligned} \quad (5.14)$$

The first term of Eq. (5.14) is the matter bispectrum,

$$\langle \delta_R(\mathbf{k}_1) \delta_R(\mathbf{k}_2) \delta_R(\mathbf{k}_3) \rangle = (2\pi)^3 B_R(k_1, k_2, k_3) \delta^D(\mathbf{k}_{123}),$$

where $\mathbf{k}_{123} \equiv \mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3$. We further calculate the ensemble average of the four-point function in the second term of Eq. (5.14). For non-Gaussian density fields, four-point function is given by a sum of connected (trispectrum) and unconnected (products of the power spectra) parts as

$$\begin{aligned} &\langle \delta_R(\mathbf{k}_1 - \mathbf{q}) \delta_R(\mathbf{q}) \delta_R(\mathbf{k}_2) \delta_R(\mathbf{k}_3) \rangle \\ &= (2\pi)^6 P_R(q) P_R(k_2) \delta^D(\mathbf{k}_1) \delta^D(\mathbf{k}_2 + \mathbf{k}_3) \\ &\quad + (2\pi)^6 P_R(k_2) P_R(k_3) \delta^D(\mathbf{k}_2 + \mathbf{q}) \delta^D(\mathbf{k}_3 + \mathbf{k}_1 - \mathbf{q}) \\ &\quad + (2\pi)^6 P_R(k_2) P_R(k_3) \delta^D(\mathbf{k}_3 + \mathbf{q}) \delta^D(\mathbf{k}_2 + \mathbf{k}_1 - \mathbf{q}) \\ &\quad + (2\pi)^3 T_R(\mathbf{k}_1 - \mathbf{q}, \mathbf{q}, \mathbf{k}_2, \mathbf{k}_3) \delta^D(\mathbf{k}_{123}), \end{aligned}$$

where T_R is the matter trispectrum. Note that the first term in the above equation cancels the last term in Eq. (5.14). Combining the above equations, Eq. (5.14) becomes

$$\begin{aligned}
& B_h(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) \\
&= b_1^3 \left[B_R(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) + \frac{b_2}{b_1} \{P_R(k_1)P_R(k_2) + (2 \text{ cyclic})\} \right. \\
&\quad \left. + \frac{1}{2} \frac{b_2}{b_1} \int \frac{d^3 q}{(2\pi)^3} T_R(\mathbf{q}, \mathbf{k}_1 - \mathbf{q}, \mathbf{k}_2, \mathbf{k}_3) + (2 \text{ cyclic}) \right]. \tag{5.15}
\end{aligned}$$

Therefore, we find that the MLB method and the locality bias assumption give formally the same results.

Although the Eq. (5.11) and Eq. (5.15) are the same, there is a subtle difference between them: the coefficient in the last term of Eq. (5.15) does not include σ_R^2 explicitly. By evaluating the last cyclic term in Eq. (5.15) for the local type of non-Gaussianity, we find that the integration of the smoothed trispectrum depends on the smoothing scale, R , up to a constant factor of $1/\sigma_R^2$ on large scales, say, $k < 0.1 \text{ h/Mpc}$. For example, the bottom right panel of Fig. 5.11 shows that $B_{f_{\text{NL}}^2}^{nG}$ and $B_{g_{\text{NL}}}^{nG}$, which are defined in Eq. (5.31) and Eq. (5.32), respectively, do not depend on the smoothing scale R , as they include $1/\sigma_R^2$ in their definitions.

Therefore, it is physically more sensible to include σ_R^2 explicitly in the equation such that the dependence on smoothing scales on large scales can be absorbed by the bias parameters. This motivates our writing the final form of the halo bispectrum, derived from the local bias assumption, as

$$\begin{aligned}
& B_h(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) \\
&= b_1^3 \left[B_R(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) + \frac{b_2}{b_1} \{P_R(k_1)P_R(k_2) + (2 \text{ cyclic})\} \right. \\
&\quad \left. + \frac{\tilde{b}_2}{b_1} \frac{1}{2\sigma_R^2} \int \frac{d^3 q}{(2\pi)^3} T_R(\mathbf{q}, \mathbf{k}_1 - \mathbf{q}, \mathbf{k}_2, \mathbf{k}_3) + (2 \text{ cyclic}) \right], \tag{5.16}
\end{aligned}$$

with three bias parameters: b_1 , b_2 and $\tilde{b}_2 \equiv b_2 \sigma_R^2$. Note that $\tilde{b}_2/b_1 \rightarrow \delta_c$ for the MLB formula. Although δ_c is known to be 1.68 for the spherically collapsed halo in the flat matter dominated universe, its precise value, in this context, needs to be tested against N-body simulations.

Eqs. (5.11) and (5.16) are the first main results of this paper, which are general and can be applied to any models of non-Gaussianities, once we know the bispectrum and

trispectrum of the underlying matter density field. Note that Eq. (5.16) was also obtained independently by Sefusatti (2009).

In principle both b_1 and b_2 are calculable from the theory of collapse of halos (see, e.g., Cooray & Sheth, 2002, for a review); however, in practice it is more convenient and safe to treat them as free parameters that one should marginalize over when extracting the cosmological information, such as f_{NL} . See Jeong & Komatsu (2009) for the same argument applied to the galaxy power spectrum.

5.2 Effects of Local-type Primordial non-Gaussianity on The Halo Bispectrum

In this section we shall evaluate Eq. (5.16) for the local-type primordial non-Gaussianity with a high-order term added:

$$\Phi(\mathbf{x}) = \phi(\mathbf{x}) + f_{\text{NL}} [\phi^2(\mathbf{x}) - \langle \phi^2 \rangle] + g_{\text{NL}} \phi^3(\mathbf{x}). \quad (5.17)$$

The cubic-order term does not generate the bispectrum of CMB anisotropy or the matter density fluctuations at the leading order; however, it does generate the trispectrum, and the CMB trispectrum has been calculated by Okamoto & Hu (2002); Kogo & Komatsu (2006). On the other hand, the bispectrum of *halos* receives the contribution from the trispectrum (see the last term in Eq. (5.16)), and thus it is necessary to include g_{NL} .

To calculate various components of the bispectra shown in Eq. (5.16), we calculate the transfer function, $T(k)$, and the power spectra with the cosmological parameters in Table 1 (“WMAP+BAO+SN”) of Komatsu et al. (2009).

As for the smoothing scale, we use $R = 1 \ h^{-1} \text{ Mpc}$. Although the smoothing scale explicitly appears in the equation, it makes negligible differences for the bispectrum on large scales, $k \ll 1/R$.

Note that we shall adopt the non-standard convention in which $\Phi(\mathbf{x})$ is Bardeen’s curvature perturbation extrapolated to the present epoch, $z = 0$, using the linear growth factor of Φ , $g(z) \equiv (1+z)D(z)$. Therefore, our f_{NL} and g_{NL} in this paper are different from those in the literature on the CMB non-Gaussianity by a factor of $g(1090)/g(0)$, i.e., $f_{\text{NL}} = [g(1090)/g(0)]f_{\text{NL}}^{\text{CMB}}$ and $g_{\text{NL}} = [g^2(1090)/g^2(0)]g_{\text{NL}}^{\text{CMB}}$ ⁷.

⁷The ratio $g(1090)/g(0)$ is 1.308 for the cosmological parameters in Table 1 (“WMAP+BAO+SN”) of Komatsu et al. (2009).

The bispectrum and trispectrum of Φ are given by

$$B_{\Phi}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = 2f_{\text{NL}} [P_{\phi}(k_1)P_{\phi}(k_2) + (2 \text{ cyclic})], \quad (5.18)$$

and

$$\begin{aligned} T_{\Phi}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) \\ = & 6g_{\text{NL}} [P_{\phi}(k_1)P_{\phi}(k_2)P_{\phi}(k_3) + (3 \text{ cyclic})] + 2f_{\text{NL}}^2 \\ & \times [P_{\phi}(k_1)P_{\phi}(k_2) \{P_{\phi}(k_{13}) + P_{\phi}(k_{14})\} + (11 \text{ cyclic})], \end{aligned} \quad (5.19)$$

with $k_{ij} = |\mathbf{k}_i + \mathbf{k}_j|$, respectively.

While Eq. (5.19) is the consequence of Eq. (5.17), general multi-field inflation models do not necessarily relate the coefficients of the trispectrum to that of the bispectrum. Therefore, one may generalize Eq. (5.19) by replacing f_{NL}^2 with a new parameter, τ_{NL} , which may or may not be related to f_{NL} :

$$\begin{aligned} T_{\Phi}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) \\ = & 6g_{\text{NL}} [P_{\phi}(k_1)P_{\phi}(k_2)P_{\phi}(k_3) + (3 \text{ cyclic})] + \frac{25}{18}\tau_{\text{NL}} \\ & \times [P_{\phi}(k_1)P_{\phi}(k_2) \{P_{\phi}(k_{13}) + P_{\phi}(k_{14})\} + (11 \text{ cyclic})]. \end{aligned} \quad (5.20)$$

Note that the coefficient of τ_{NL} reflects the definition of τ_{NL} introduced by Boubekeur & Lyth (2006). This opens up an exciting possibility that the galaxy bispectrum can test whether $\tau_{\text{NL}} = (6f_{\text{NL}}/5)^2$ or other predictions for the relation between τ_{NL} and f_{NL} are satisfied observationally.

We transform these spectra to those of the smoothed linear density contrasts, using the Poisson equation,

$$\delta_R^{(1)}(k) = \frac{2}{3} \frac{k^2 T(k)}{H_0^2 \Omega_m} \tilde{W}_R(k) \Phi(k) \equiv \mathcal{M}_R(k) \Phi(k), \quad (5.21)$$

where $\tilde{W}_R(k)$ is the Fourier transform of the top-hat filter with radius R . Note that $\tilde{W}_R(k) \rightarrow 1$ as $k \rightarrow 0$. In general $\tilde{W}_R(k) \simeq 1$ for $k \ll 1/R$. Then, we calculate the n -point function of the matter density fields from the corresponding correlator of curvature perturbations by

$$\langle \delta_R^{(1)}(\mathbf{k}_1) \cdots \delta_R^{(1)}(\mathbf{k}_n) \rangle = \prod_{i=1}^n \mathcal{M}_R(k_i) \langle \Phi(\mathbf{k}_1) \cdots \Phi(\mathbf{k}_n) \rangle.$$

5.2.1 Known Terms

5.2.1.1 Formula

The first term in Eq. (5.16) contains the bispectrum of matter density fluctuations, B_R , which consists of two pieces: (i) the non-linear evolution of gravitational clustering (B_m^G) and (ii) primordial non-Gaussianity ($B_{f_{\text{NL}}}^{nG0}$)⁸:

$$B_R(k_1, k_2, k_3) = B_m^G(k_1, k_2, k_3) + f_{\text{NL}} B_{f_{\text{NL}}}^{nG0}(k_1, k_2, k_3), \quad (5.22)$$

where

$$\begin{aligned} B_m^G(k_1, k_2, k_3) &\equiv \tilde{W}_R(k_1) \tilde{W}_R(k_2) \tilde{W}_R(k_3) 2F_2^{(s)}(\mathbf{k}_1, \mathbf{k}_2) \\ &\quad \times P_m(k_1) P_m(k_2) + (2 \text{ cyclic}), \end{aligned} \quad (5.23)$$

with $F_2^{(s)}$ given by Eq. (5.4), and

$$\begin{aligned} &B_{f_{\text{NL}}}^{nG0}(k_1, k_2, k_3) \\ &\equiv 2 \prod_{i=1}^3 \mathcal{M}_R(k_i) [P_\phi(k_1) P_\phi(k_2) + (2 \text{ cyclic})] \\ &= 2 \frac{P_R(k_1)}{\mathcal{M}_R(k_1)} \frac{P_R(k_2)}{\mathcal{M}_R(k_2)} \mathcal{M}_R(k_3) + (2 \text{ cyclic}). \end{aligned} \quad (5.24)$$

One finds that Eqs. (5.24) and (5.23) agree with the first and the second terms in Eq. (5.3) on the scales much larger than the smoothing scale, i.e., $k \ll 1/R$, for which $\tilde{W}_R \rightarrow 1$.

One might wonder if it is OK to include the bispectrum from non-linear evolution of density fluctuations in the MLB formula, as Eq. (5.11) is usually used for the Lagrangian density fluctuations, i.e., “initial” fluctuations. However, it is perfectly OK to use the evolved density fluctuations in this formula, as one can always use the evolved density fluctuations as the initial data. For example, we can think of starting our calculation at $z = 10$, and ask the MLB formula to take the initial condition at $z = 10$, including non-linear correction. Since we know how to compute the bispectrum, trispectrum, etc., of the underlying mass distribution at $z = 10$ (including non-linear effects), we can use this information in the MLB

⁸We ignore the following term in $B_R(k_1, k_2, k_3)$:

$$\prod_{i=1}^3 \tilde{W}_R(k_i) \int \frac{d^3 q}{(2\pi)^3} F_2^{(s)}(\mathbf{q}, \mathbf{k}_1 - \mathbf{q}) T(\mathbf{q}, \mathbf{k}_1 - \mathbf{q}, \mathbf{k}_2, \mathbf{k}_3) + (2 \text{ cyclic}),$$

where T is the unfiltered primordial trispectrum which contains f_{NL}^2 and g_{NL} . This term is negligibly small (Scoccimarro et al., 2004).

formula. In other words, the “initial” distribution does not need to be primordial. We can provide the evolved density field as the initial data, and compute the peak statistics. The MLB formula does not care whether the source of non-Gaussianity is truly primordial or not: the only conditions that we must respect for Eq. (5.11) to be valid are (i) high peaks ($\nu \gg 1$), and (ii) n -point correlation functions are much less than unity, $\xi_R^{(n)} \ll 1$, so that the exponential in Eq. (5.8) can be Taylor-expanded. In this case one would lose an ability to calculate the bias parameters, b_1 and b_2 , using, e.g., a halo model; however, this is not a disadvantage, as the halo model calculations of the *galaxy* bias parameter, b_1 and b_2 , are at best qualitative even for Gaussian initial conditions (see, e.g., Jeong & Komatsu, 2009). In our approach the coefficients of individual terms in Eqs. (5.11) and (5.16), including δ_c , are free parameters, and need to be determined from observations themselves.

5.2.1.2 Shape Dependence: Non-linear Gravitational Evolution and Non-linear Galaxy Bias

How about the shape dependence? First, let us review the structure of $B_m^G(k_1, k_2, k_3)$ (Eq. (5.23)), which has been studied extensively in the literature (see Bernardeau et al., 2002, for a review).

Here, let us offer a useful way of visualizing the shape dependence of the bispectrum. We can study the structure of the bispectrum by plotting the magnitude of B_m^G as a function of k_2/k_1 and k_3/k_1 for a given k_1 , with a condition that $k_1 \geq k_2 \geq k_3$ is satisfied. In order to classify various shapes of the triangles, let us use the following names: squeezed ($k_1 \simeq k_2 \gg k_3$), elongated ($k_1 = k_2 + k_3$), folded ($k_1 = 2k_2 = 2k_3$), isosceles ($k_2 = k_3$), and equilateral ($k_1 = k_2 = k_3$). See (a)–(e) of Fig. 5.1 for the visual representations of these triangles.

The top-left panel of Fig. 5.2 shows B_m^G for $k_1 = 0.01 h \text{ Mpc}^{-1}$. In this regime $P_R(k_1)$ is an increasing function of k_1 (recall that $P_R(k)$ peaks at $k \approx 0.02 h \text{ Mpc}^{-1}$). Let us then pick the first term in Eq. (5.23), $F_2^{(s)}(\mathbf{k}_1, \mathbf{k}_2)P_R(k_1)P_R(k_2)$, and ignore the cyclic terms for the moment. It follows from Eq. (5.4) and the descriptions below it that $F_2^{(s)}(\mathbf{k}_1, \mathbf{k}_2)$ vanishes in the squeezed limit ($\mathbf{k}_1 = -\mathbf{k}_2$) and reaches the maximum in the opposite limit ($\mathbf{k}_1 = \alpha\mathbf{k}_2$). Therefore, we would expect this term to give large signals in the “elongated configurations,” $k_1 = k_2 + k_3$; however, as $P_R(k)$ at $k \lesssim 0.02 h \text{ Mpc}^{-1}$ is an increasing function of k , one can also get large signals when k_1 and k_2 are equally large, $k_1 = k_2$. As we have zero signal in the squeezed limit, $k_3 = 0$, it follows that we can find a large signal in the equilateral configuration, $k_1 = k_2 = k_3$. A similar argument also

applies to the cyclic terms. As a result, for $k_1 = 0.01 \ h \text{ Mpc}^{-1}$, we find the largest signal in the equilateral configuration, and then the signal decreases as we approach the squeezed configuration, i.e., the signal decreases as we go from (e) to (a) in Fig. 5.1.

The top-right panel of Fig. 5.2 shows B_m^G for $k_1 = 0.05 \ h \text{ Mpc}^{-1}$. In this regime $P_R(k_1)$ is a *decreasing* function of k_1 , and thus the equilateral configurations are no longer as important as the folded ones. Instead we have large signals in the folded configurations as well as in the elongated configurations. Note that the exact squeezed limit is still suppressed due to the form of $F_2^{(s)}$.

In summary, B_m^G usually has the largest signal in the folded and elongated (or equilateral, depending on the wavenumber) configurations, with the squeezed configurations suppressed relative to the others. The suppression of the squeezed configurations is a generic signature of the causal mechanism such as the non-linear gravitational evolution that $F_2^{(s)}$ describes.

The bispectrum from the non-linear bias term, the second term in Eq. (5.11), has the same structure as B_m^G , but it does not contain $F_2^{(s)}$. As a result the non-linear bias term does not have as much suppression as B_m^G has in the squeezed configuration. In addition, as $F_2^{(s)}$ enhances the signal in the elongated configurations, the non-linear bias term does not have as much enhancement as B_m^G has in the elongated configurations. These properties explain the bottom panels of Fig. 5.2.

As B_m^G and the non-linear bias term contain products of $P_R(k_1)P_R(k_2)$ and the cyclic terms, it is often more convenient to deal with $Q_h(k_1, k_2, k_3)$ given by (Peebles, 1980)

$$Q_h(k_1, k_2, k_3) \equiv \frac{B_h(k_1, k_2, k_3)}{P_R(k_1)P_R(k_2) + (2 \text{ cyclic})}, \quad (5.25)$$

to reduce the dependence on the shape of the power spectrum. This combination is constant and equal to $b_1^2 b_2$ for the non-linear bias term (see the second term in Eq. (5.11)).

The left and right panels of Fig. 5.3 show $B_m^G(k_1, k_2, k_3)/[P_R(k_1)P_R(k_2) + (2 \text{ cyclic})]$ for $k_1 = 0.01 \ h \text{ Mpc}^{-1}$ and $0.05 \ h \text{ Mpc}^{-1}$, respectively. We find that Q_h better reflects the shape dependence of $F_2^{(s)}$ irrespective of k_1 : it has the largest signal in the folded and elongated configurations in both large and small scales. The squeezed configurations are still heavily suppressed relative to the others.

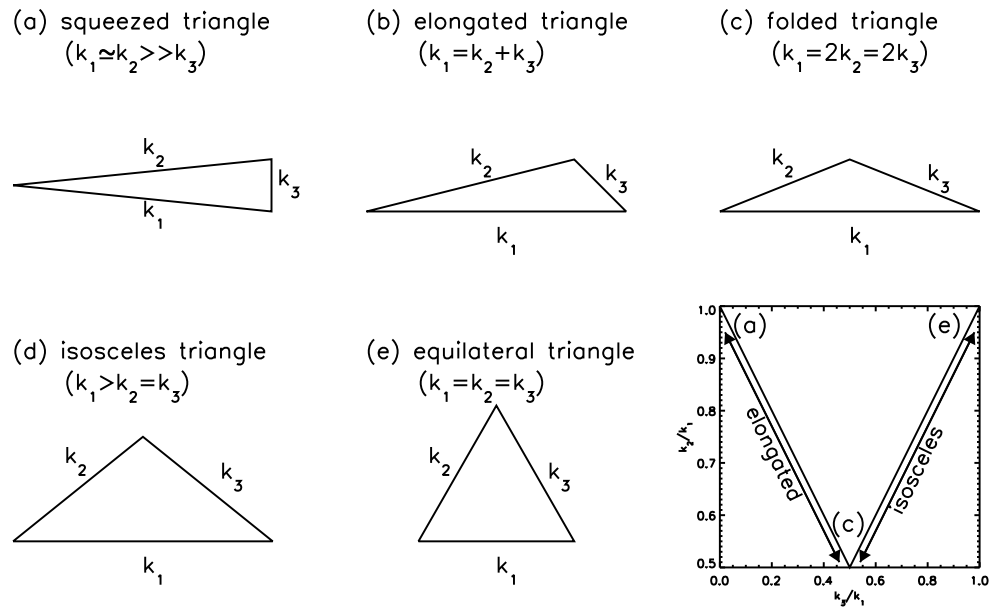


Figure 5.1: Visual representations of triangles forming the bispectrum, $B(k_1, k_2, k_3)$, with various combinations of wavenumbers satisfying $k_3 \leq k_2 \leq k_1$.

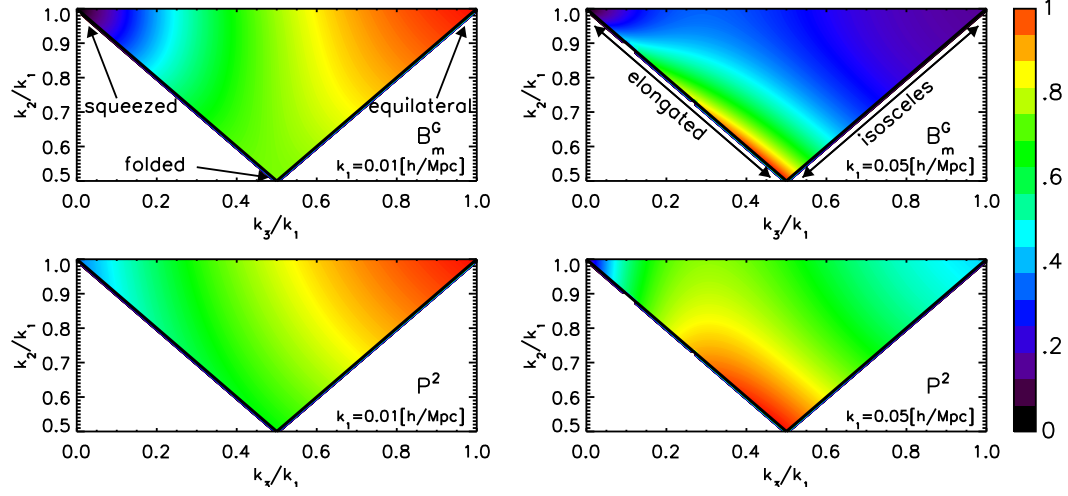


Figure 5.2: Shape of the bispectrum, $B(k_1, k_2, k_3)$. Each panel shows the amplitude of the bispectrum as a function of k_2/k_1 and k_3/k_1 for a given k_1 , with a condition that $k_3 \leq k_2 \leq k_1$ is satisfied. The amplitude is normalized such that it is unity at the point where the bispectrum takes on the maximum value. For the visual representations of the triangle names such as the squeezed, elongated, folded, isosceles, and equilateral, see Fig. 5.1. (Top Left) The bispectrum from the non-linear gravitational evolution, B_m^G (Eq. (5.23)), for $k_1 = 0.01 h \text{ Mpc}^{-1}$. (Top Right) B_m^G for $k_1 = 0.05 h \text{ Mpc}^{-1}$. (Bottom Left) The bispectrum from the non-linear galaxy biasing, $P_R(k_1)P_R(k_2) + (2 \text{ cyclic})$ (the second term in Eq. (5.16)), for $k_1 = 0.01 h \text{ Mpc}^{-1}$. (Bottom Right) $P_R(k_1)P_R(k_2) + (2 \text{ cyclic})$ for $k_1 = 0.05 h \text{ Mpc}^{-1}$.

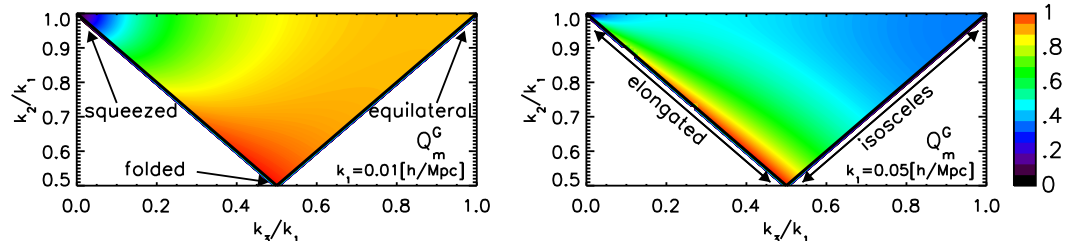


Figure 5.3: Same as the top panels of Fig. 5.2, but for $B_m^G / [P_R(k_1)P_R(k_2) + (2 \text{ cyclic})]$ (Eq. (5.25)).

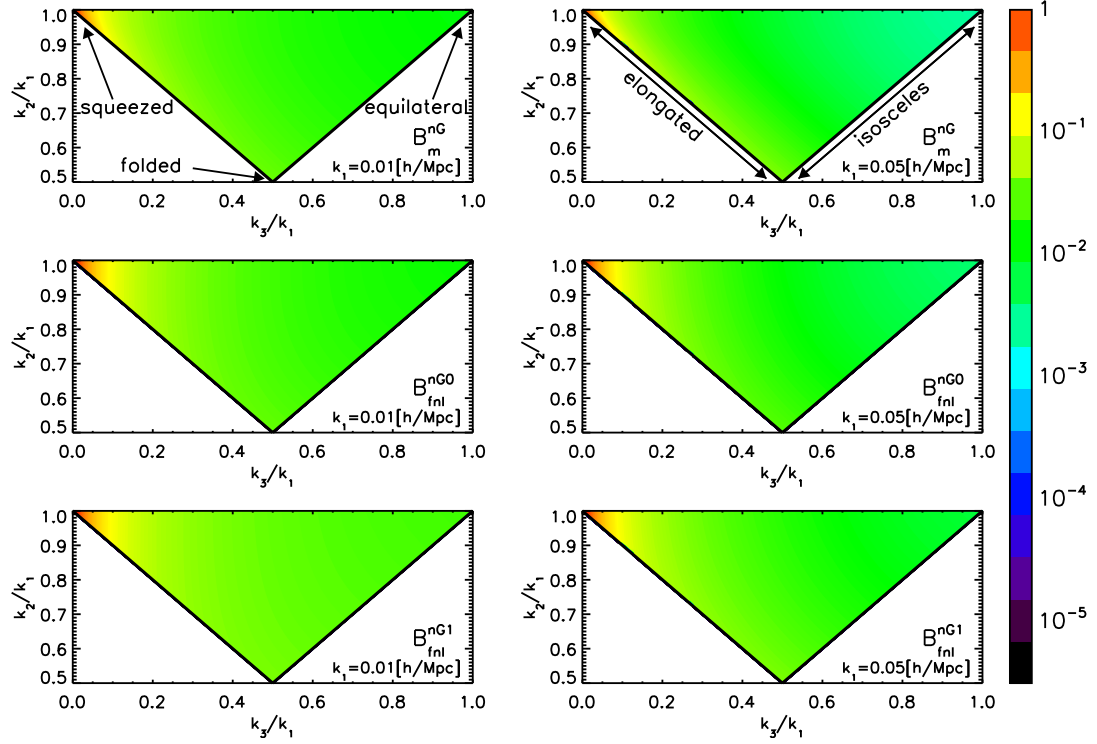


Figure 5.4: Same as Fig. 5.2, but for the terms proportional to f_{NL} . (Top) the B_m^{nG} term (Eq. (5.35)), (Middle) the $B_{f_{\text{NL}}}^{nG0}$ term (Eq. (5.24)), and (Bottom) the $B_{f_{\text{NL}}}^{nG1}$ term (Eq. (5.41)). Note that the non-Gaussian terms diverge in the exact squeezed limit, $k_3 \rightarrow 0$; thus, we show these terms normalized to be unity at $k_3/k_1 = 10^{-2}$. In order to facilitate the comparison better, we draw the dotted contour for all six panels.

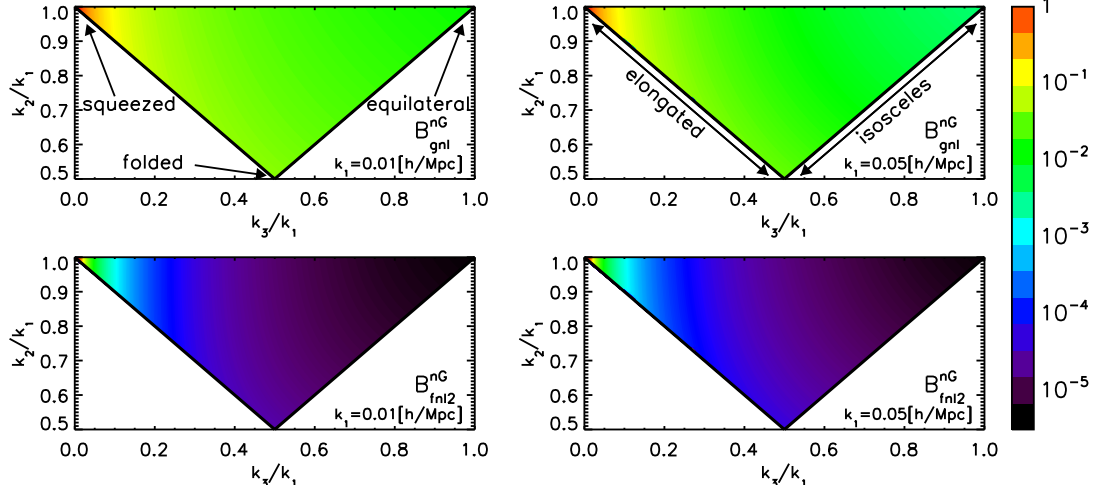


Figure 5.5: Same as Fig. 5.2, but for (Top) the g_{NL} term (Eq. (5.32)), and (Bottom) the f_{NL}^2 term (Eq. (5.33)). Note that the non-Gaussian terms diverge in the exact squeezed limit, $k_3 \rightarrow 0$; thus, we show these terms normalized to be unity at $k_3/k_1 = 10^{-2}$. In order to facilitate the comparison better, we draw the dotted contour for top panels.

5.2.1.3 Shape Dependence: f_{NL} Term

How about the f_{NL} term, $B_{f_{\text{NL}}}^{nG0}(k_1, k_2, k_3)$? This term has a completely different structure. Let us pick the first term, $\mathcal{M}_R(k_1)P_\phi(k_1)\mathcal{M}_R(k_2)P_\phi(k_2)\mathcal{M}_R(k_3)$, in Eq. (5.24). The important point is that the power spectrum of ϕ is always a decreasing function of k , i.e., $P_\phi(k) \propto 1/k^3$ for a scale-invariant spectrum. On the other hand, on large scales we have $T(k) \rightarrow 1$ and $\mathcal{M}_R(k) \propto k^2$. Therefore, collecting all the cyclic terms, we find $B_{f_{\text{NL}}}^{nG0}(k_1, k_2, k_3) \propto k_3^2/(k_1 k_2) + k_2^2/(k_1 k_3) + k_1^2/(k_2 k_3) = (k_1^3 + k_2^3 + k_3^3)/(k_1 k_2 k_3)$. In other words, it has the largest signal when one of k 's is very small, i.e., the squeezed configurations, which is opposite to the structures of B_m^G and the non-linear bias term. The middle panels of Fig. 5.4 show $B_{f_{\text{NL}}}^{nG0}$ for $k_1 = 0.01 h \text{ Mpc}^{-1}$ and $0.05 h \text{ Mpc}^{-1}$, and we find the largest signals in the squeezed configurations. We also find that Q_h from the f_{NL} term, $B_{f_{\text{NL}}}^{nG0}(k_1, k_2, k_3)/[P_R(k_1)P_R(k_2) + (2 \text{ cyclic})]$, still has the largest signal in the squeezed configurations.

These properties allow us to distinguish between the primordial non-Gaussianity and the other effects such as the non-linear gravitational evolution and non-linear bias. Sefusatti & Komatsu (2007) have studied in detail how well one can separate these effects

using Q_h .

5.2.2 New Term

5.2.2.1 Formula

Now, we shall evaluate the new term, the third term in Eq. (5.16), which was not considered in Sefusatti & Komatsu (2007). The trispectrum is generated by the primordial non-Gaussianity, as well as by the non-linear evolution of matter density fluctuations. The non-linear evolution of matter density fluctuations on large scales is given by perturbation theory (see Bernardeau et al., 2002, for a review). Let us expand the filtered non-linear matter density field in Fourier space as

$$\delta_R(\mathbf{k}) = \tilde{W}_R(k) \left[\delta^{(1)}(\mathbf{k}) + \delta^{(2)}(\mathbf{k}) + \delta^{(3)}(\mathbf{k}) + \dots \right], \quad (5.26)$$

where $\delta^{(n)}(\mathbf{k})$ is the n -th order quantity of the linear density contrast, $\delta^{(1)}(\mathbf{k})$. Then, the connected matter density trispectrum is given by

$$\begin{aligned} & T_R(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) \\ = & T_R^{1111}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) + \{T_R^{1112}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) + (3 \text{ cyclic})\} \\ & + \{T_R^{1113}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) + (3 \text{ cyclic})\} \\ & + \{T_R^{1122}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) + (5 \text{ cyclic})\} + \mathcal{O}(\phi^8), \end{aligned} \quad (5.27)$$

with T_R^{ijkl} given by

$$\begin{aligned} & (2\pi)^3 \delta^D(\mathbf{k}_{1234}) T_R^{ijkl}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) \\ \equiv & \prod_{n=1}^4 \tilde{W}_R(k_n) \langle \delta^{(i)}(\mathbf{k}_1) \delta^{(j)}(\mathbf{k}_2) \delta^{(k)}(\mathbf{k}_3) \delta^{(l)}(\mathbf{k}_4) \rangle_c. \end{aligned} \quad (5.28)$$

The leading contributions of all the terms shown in Eq. (5.27) are order of ϕ^6 .

The first term, T_R^{1111} , is the linearly evolved primordial trispectrum calculated from Eq. (5.19), and thus it contains the terms proportional to f_{NL}^2 and g_{NL} . The second term, T_R^{1112} , has a coupling between the primordial non-Gaussianity (linear in f_{NL}) and the non-linear gravitational evolution (linear in $F_2^{(s)}$). These two terms are important on large scales.

The other terms, T_R^{1113} and T_R^{1122} , do not have contributions from f_{NL} or g_{NL} at the leading-order level, but solely come from the non-linear gravitational coupling; thus,

they may be ignored on large scales we are considering in this paper. Sefusatti (2009) also derived and studied T_R^{1112} as well as T_R^{1113} and T_R^{1122} .

Therefore, we approximate the integration in the third term of Eq. (5.16) as

$$\begin{aligned} & \frac{1}{2\sigma_R^2} \int \frac{d^3q}{(2\pi)^3} [T_R(\mathbf{q}, \mathbf{k}_1 - \mathbf{q}, \mathbf{k}_2, \mathbf{k}_3) + (2 \text{ cyclic})] \\ \approx & \frac{1}{2\sigma_R^2} \left\{ \int \frac{d^3q}{(2\pi)^3} \left[T_R^{(1)}(\mathbf{q}, \mathbf{k}_1 - \mathbf{q}, \mathbf{k}_2, \mathbf{k}_3) + (2 \text{ cyclic}) \right] \right. \\ & \left. + \int \frac{d^3q}{(2\pi)^3} \left[T_R^{(2)}(\mathbf{q}, \mathbf{k}_1 - \mathbf{q}, \mathbf{k}_2, \mathbf{k}_3) + (2 \text{ cyclic}) \right] \right\}, \end{aligned} \quad (5.29)$$

where “cyclic” denotes the cyclic combinations of k_1 , k_2 , and k_3 , and $T_R^{(1)}$ and $T_R^{(2)}$ denote $T_R^{(1)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) = T_R^{1111}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4)$, and $T_R^{(2)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) = T_R^{1112}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) + (3 \text{ cyclic})$, respectively.

The first term in Eq. (5.29) is the integration of the linearly evolved primordial curvature trispectrum, which contains two pieces: one proportional to f_{NL}^2 and another to g_{NL} (see Eq. (5.19)). Therefore, we symbolically write the first line in Eq. (5.29) as

$$\begin{aligned} & \frac{1}{2\sigma_R^2} \int \frac{d^3q}{(2\pi)^3} \left[T_R^{(1)}(\mathbf{q}, \mathbf{k}_1 - \mathbf{q}, \mathbf{k}_2, \mathbf{k}_3) + (2 \text{ cyclic}) \right] \\ = & g_{\text{NL}} B_{g_{\text{NL}}}^{nG}(k_1, k_2, k_3) + f_{\text{NL}}^2 B_{f_{\text{NL}}^2}^{nG}(k_1, k_2, k_3), \end{aligned} \quad (5.30)$$

where

$$\begin{aligned}
& B_{f_{\text{NL}}^2}^{nG}(k_1, k_2, k_3) \\
& \equiv \frac{1}{2\sigma_R^2} \left[4\mathcal{M}_R(k_2)\mathcal{M}_R(k_3) \int \frac{d^3q}{(2\pi)^3} \mathcal{M}_R(q)\mathcal{M}_R(|\mathbf{k}_1 - \mathbf{q}|)P_\phi(q) \right. \\
& \quad \times [P_\phi(|\mathbf{k}_1 - \mathbf{q}|)P_\phi(|\mathbf{k}_2 + \mathbf{q}|) + P_\phi(|\mathbf{k}_1 - \mathbf{q}|)P_\phi(|\mathbf{k}_3 + \mathbf{q}|)] + (2 \text{ cyclic}) \\
& + 8\mathcal{M}_R(k_2)\mathcal{M}_R(k_3)P_\phi(k_2) \\
& \quad \times \int \frac{d^3q}{(2\pi)^3} \mathcal{M}_R(q)\mathcal{M}_R(|\mathbf{k}_1 - \mathbf{q}|)P_\phi(q)P_\phi(|\mathbf{k}_3 + \mathbf{q}|) + (2 \text{ cyclic}) \\
& + 8\mathcal{M}_R(k_2)\mathcal{M}_R(k_3)P_\phi(k_3) \\
& \quad \times \int \frac{d^3q}{(2\pi)^3} \mathcal{M}_R(q)\mathcal{M}_R(|\mathbf{k}_1 - \mathbf{q}|)P_\phi(q)P_\phi(|\mathbf{k}_2 + \mathbf{q}|) + (2 \text{ cyclic}) \\
& + 8\mathcal{M}_R(k_2)\mathcal{M}_R(k_3)P_\phi(k_1) [P_\phi(k_2) + P_\phi(k_3)] \\
& \quad \times \int \frac{d^3q}{(2\pi)^3} \mathcal{M}_R(q)\mathcal{M}_R(|\mathbf{k}_1 - \mathbf{q}|)P_\phi(q) + (2 \text{ cyclic}) \\
& + 4\mathcal{M}_R(k_2)\mathcal{M}_R(k_3)P_\phi(k_2)P_\phi(k_3) \int \frac{d^3q}{(2\pi)^3} \mathcal{M}_R(q)\mathcal{M}_R(|\mathbf{k}_1 - \mathbf{q}|) \\
& \quad \times [P_\phi(|\mathbf{k}_2 + \mathbf{q}|) + P_\phi(|\mathbf{k}_3 + \mathbf{q}|)] + (2 \text{ cyclic}) \Big], \tag{5.31}
\end{aligned}$$

$$\begin{aligned}
B_{g_{\text{NL}}}^{nG}(k_1, k_2, k_3) & \equiv \frac{1}{2\sigma_R^2} \left[6\mathcal{M}_R(k_2)\mathcal{M}_R(k_3) [P_\phi(k_2) + P_\phi(k_3)] \right. \\
& \times \int \frac{d^3q}{(2\pi)^3} \mathcal{M}_R(q)\mathcal{M}_R(|\mathbf{k}_1 - \mathbf{q}|)P_\phi(q)P_\phi(|\mathbf{k}_1 - \mathbf{q}|) + (2 \text{ cyclic}) \\
& + 12\mathcal{M}_R(k_2)\mathcal{M}_R(k_3)P_\phi(k_2)P_\phi(k_3) \\
& \quad \left. \int \frac{d^3q}{(2\pi)^3} \mathcal{M}_R(q)\mathcal{M}_R(|\mathbf{k}_1 - \mathbf{q}|)P_\phi(q) + (2 \text{ cyclic}) \right]. \tag{5.32}
\end{aligned}$$

We find that the first three cyclic terms in Eq. (5.31) are parametrically small on large scales and may be ignored for $k \lesssim 0.1 \ h \text{ Mpc}^{-1}$. Therefore, one may just calculate the last two cyclic terms:

$$\begin{aligned}
B_{f_{\text{NL}}^2}^{nG}(k_1, k_2, k_3) & \approx \frac{1}{2\sigma_R^2} \left[8\mathcal{M}_R(k_2)\mathcal{M}_R(k_3)P_\phi(k_1) [P_\phi(k_2) + P_\phi(k_3)] \right. \\
& \times \int \frac{d^3q}{(2\pi)^3} \mathcal{M}_R(q)\mathcal{M}_R(|\mathbf{k}_1 - \mathbf{q}|)P_\phi(q) + (2 \text{ cyclic}) \\
& + 4\mathcal{M}_R(k_2)\mathcal{M}_R(k_3)P_\phi(k_2)P_\phi(k_3) \int \frac{d^3q}{(2\pi)^3} \mathcal{M}_R(q)\mathcal{M}_R(|\mathbf{k}_1 - \mathbf{q}|) \\
& \quad \times [P_\phi(|\mathbf{k}_2 + \mathbf{q}|) + P_\phi(|\mathbf{k}_3 + \mathbf{q}|)] + (2 \text{ cyclic}) \Big]. \tag{5.33}
\end{aligned}$$

Next, the second term of Eq. (5.29) contains a cross-correlation between the non-linearly evolved density field ($\delta^{(2)} \sim F_2^{(s)}[\delta^{(1)}]^2$) and the primordial bispectrum, and thus it is linearly proportional to f_{NL} and $F_2^{(s)}$. We present the explicit functional form of T_R^{1112} as well as the full expression of the second term of Eq. (5.29) in Appendix J. Here, we only show the final result. We write it as

$$\begin{aligned} & \frac{1}{2\sigma_R^2} \int \frac{d^3q}{(2\pi)^3} T_R^{(2)}(\mathbf{q}, \mathbf{k}_1 - \mathbf{q}, \mathbf{k}_2, \mathbf{k}_3) + (2 \text{ cyclic}) \\ &= f_{\text{NL}} \left[B_m^G(k_1, k_2, k_3) + B_{f_{\text{NL}}}^{nG1}(k_1, k_2, k_3) \right. \\ & \quad \left. + 4B_{f_{\text{NL}}}^{nG0}(k_1, k_2, k_3) \{ \mathcal{G}_R(k_1) + \mathcal{G}_R(k_2) + \mathcal{G}_R(k_3) \} \right], \end{aligned} \quad (5.34)$$

where

$$\begin{aligned} & B_m^G(k_1, k_2, k_3) \\ & \equiv 4\mathcal{W}_R(k_1)\mathcal{W}_R(k_2)\mathcal{W}_R(k_3) \\ & \quad \times \left\{ \left[\frac{\mathcal{F}_R(k_1)}{\mathcal{M}_R(k_1)} + \frac{\mathcal{F}_R(k_2)}{\mathcal{M}_R(k_2)} \right] P_m(k_1)P_m(k_2)F_2^{(s)}(\mathbf{k}_1, \mathbf{k}_2) + (2 \text{ cyclic}) \right\}, \end{aligned} \quad (5.35)$$

and

$$\begin{aligned} & B_{f_{\text{NL}}}^{nG1}(k_1, k_2, k_3) \\ & \equiv \frac{1}{2\sigma_R^2} \left[8\mathcal{W}_R(k_2)\mathcal{W}_R(k_3)\mathcal{M}(k_3)P_m(k_2) \int \frac{d^3q}{(2\pi)^3} \mathcal{W}_R(|\mathbf{k}_1 - \mathbf{q}|)\mathcal{W}_R(q) \right. \\ & \quad \times \mathcal{M}(|\mathbf{k}_1 - \mathbf{q}|)\mathcal{M}(|\mathbf{k}_2 + \mathbf{q}|)F_2^{(s)}(-\mathbf{k}_2, \mathbf{k}_2 + \mathbf{q}) \\ & \quad \times \{ P_\phi(k_3)P_\phi(|\mathbf{k}_1 - \mathbf{q}|) + P_\phi(k_3)P_\phi(|\mathbf{k}_2 + \mathbf{q}|) + P_\phi(|\mathbf{k}_1 - \mathbf{q}|)P_\phi(|\mathbf{k}_2 + \mathbf{q}|) \} \\ & \quad + (5 \text{ permutation}) \\ & \quad + 8\mathcal{W}_R(k_2)\mathcal{W}_R(k_3)\mathcal{M}(k_3) \int \frac{d^3q}{(2\pi)^3} \mathcal{W}_R(|\mathbf{k}_1 - \mathbf{q}|)\mathcal{W}_R(q) \\ & \quad \times \mathcal{M}(|\mathbf{k}_1 - \mathbf{q}|)\mathcal{M}(|\mathbf{k}_2 + \mathbf{q}|)P_m(q)F_2^{(s)}(-\mathbf{q}, \mathbf{k}_2 + \mathbf{q}) \\ & \quad \times \{ P_\phi(|\mathbf{k}_1 - \mathbf{q}|)P_\phi(k_3) + P_\phi(|\mathbf{k}_1 - \mathbf{q}|)P_\phi(|\mathbf{k}_2 + \mathbf{q}|) + P_\phi(k_3)P_\phi(|\mathbf{k}_2 + \mathbf{q}|) \} \\ & \quad + (5 \text{ permutation}) \\ & \quad + 8(\mathcal{W}_R(k_2)\mathcal{W}_R(k_3))^2 P_m(k_3)\mathcal{M}(k_2) \int \frac{d^3p}{(2\pi)^3} \mathcal{M}(p)\mathcal{M}(|\mathbf{k}_2 - \mathbf{p}|)P_\phi(p) \\ & \quad \times \{ P_\phi(|\mathbf{k}_2 - \mathbf{p}|) + 2P_\phi(k_2) \} F_2^{(s)}(\mathbf{p}, \mathbf{k}_2 - \mathbf{p}) + (5 \text{ permutation}) \Big]. \end{aligned} \quad (5.36)$$

Here, $\mathcal{M}(k) \equiv \mathcal{M}_R(k)/\mathcal{W}_R(k)$.

In the above equations we have defined two functions, $\mathcal{F}_R(k)$ and $\mathcal{G}_R(k)$, which are given by

$$\mathcal{F}_R(k) \equiv \frac{1}{2\sigma_R^2} \int \frac{d^3q}{(2\pi)^3} P_\phi(q) \mathcal{M}_R(q) \mathcal{M}_R(|\mathbf{k} - \mathbf{q}|) \left[\frac{P_\phi(|\mathbf{k} - \mathbf{q}|)}{P_\phi(k)} + 2 \right], \quad (5.37)$$

$$\mathcal{G}_R(k) \equiv \frac{1}{2\sigma_R^2} \int \frac{d^3q}{(2\pi)^3} \frac{\mathcal{W}_R(q) \mathcal{W}_R(|\mathbf{k} - \mathbf{q}|)}{\mathcal{W}_R(k)} P_m(q) F_2^{(s)}(\mathbf{k}, -\mathbf{q}). \quad (5.38)$$

As shown in Fig. 5.6 and Fig. 5.7, both $\mathcal{F}_R(k)$ and $\mathcal{G}_R(k)$ are almost constant on large scales. If we do not have a smoothing, i.e., $R \rightarrow 0$, the large scale asymptotic value of $\mathcal{G}_R(k)$ is $17/42$. However, the presence of filter changes this asymptotic value. As $k \rightarrow 0$,

$$\mathcal{G}_R(k) \rightarrow \frac{13}{84} + \frac{1}{4\sigma_R^2} \int \frac{d^3q}{(2\pi)^3} \mathcal{W}_R(q) P_m(q) \frac{\sin(qR)}{qR}, \quad (5.39)$$

whose value depends on the smoothing scale, R , as shown in Fig. 5.8.

Let us study the structure of each term in Eq. (5.34). The first piece is B_m^{nG} . On very large scales, where $\mathcal{W}_R(k) \rightarrow 1$ and $\mathcal{F}_R(k) \rightarrow 1$, B_m^{nG} becomes a product of the usual matter bispectrum for Gaussian initial conditions, B_m^G , and the scale dependent bias shown in Eq. (5.6), as

$$\frac{2f_{\text{NL}}\mathcal{F}_R(k)}{\mathcal{M}_R(k)} = \frac{3f_{\text{NL}}H_0^2\Omega_m}{k^2T(k)} \frac{\mathcal{F}_R(k)}{\mathcal{W}_R(k)} \rightarrow \frac{3f_{\text{NL}}H_0^2\Omega_m}{k^2T(k)}, \quad (5.40)$$

as $k \rightarrow 0$. Therefore, we can interpret this term as a scale dependent bias multiplying the usual matter bispectrum for Gaussian initial conditions; however, this behavior is not generic – in fact, the other terms cannot be expressed in terms of products of the scale-dependent bias and the results in the continuous limit, Eq. (5.3).

The next piece is $B_{f_{\text{NL}}}^{nG1}(k_1, k_2, k_3)$. By numerically calculating Eq. (5.36), we find that the terms that contain $F_2^{(s)}(\mathbf{q}, \mathbf{k} - \mathbf{q})$ are parametrically small on large scales, and that the dominant contributions come from the first permutation terms. Therefore, we approximate Eq. (5.36) on large scale ($k \lesssim 0.1h\text{Mpc}^{-1}$) as

$$\begin{aligned} & B_{f_{\text{NL}}}^{nG1}(k_1, k_2, k_3) \\ & \approx \frac{1}{2\sigma_R^2} \left[8\mathcal{W}_R(k_2)\mathcal{W}_R(k_3)P_m(k_2)\mathcal{M}(k_3)P_\phi(k_3) \right. \\ & \quad \int \frac{d^3q}{(2\pi)^3} \mathcal{W}_R(|\mathbf{k}_1 - \mathbf{q}|)\mathcal{W}_R(q)\mathcal{M}(|\mathbf{k}_1 - \mathbf{q}|)\mathcal{M}(|\mathbf{k}_2 + \mathbf{q}|) \\ & \quad \times [P_\phi(|\mathbf{k}_2 + \mathbf{q}|) + P_\phi(|\mathbf{k}_1 - \mathbf{q}|)] F_2^{(s)}(-\mathbf{k}_2, \mathbf{k}_2 + \mathbf{q}) + (5 \text{ permutation}) \Big]. \quad (5.41) \end{aligned}$$

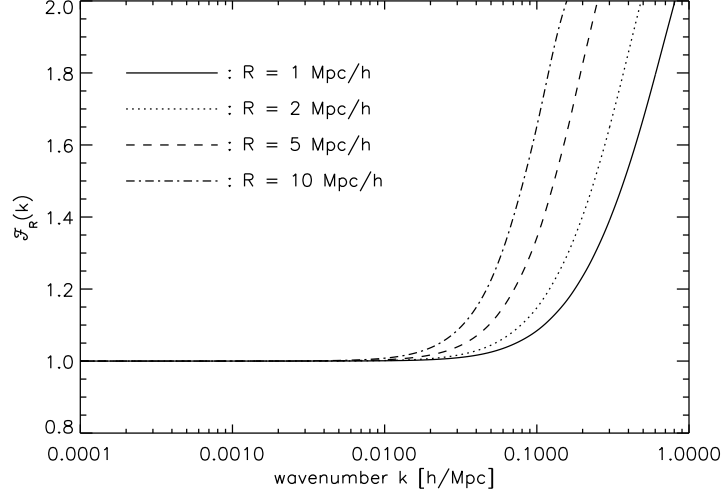


Figure 5.6: Shape of the function, $\mathcal{F}_R(k)$, defined in Eq. (5.37). We show $\mathcal{F}_R(k)$ for four different smoothing lengths: $R = 1, 2, 5, 10$ Mpc/h.

How about the last term of Eq. (5.34), $4B_{f_{\text{NL}}}^{nG0}(k_1, k_2, k_3)\{\mathcal{G}_R(k_1) + \mathcal{G}_R(k_2) + \mathcal{G}_R(k_3)\}$? As $\mathcal{G}_R(k) \rightarrow \text{constant}$ on large scales (Fig. 5.7), this piece becomes $B_{f_{\text{NL}}}^{nG0}$ multiplied by a constant factor whose exact value depends on the smoothing scale, R (Fig. 5.8).

In summary, we have derived the new terms in the galaxy bispectrum, which arise from the integration of the matter trispectrum. While we find one term, B_m^{nG} , includes the scale-dependent bias which appears on the galaxy power spectrum, we also find that there are more terms contributing to the galaxy bispectrum.

Eq. (5.30) along with Eqs. (5.31)–(5.33), and Eq. (5.34) along with Eqs. (5.35), (5.36), (5.41) are the second main results of this paper. In the next sections, we shall present the detailed assessment of the new terms we have derived in this section.

5.2.2.2 Shape Dependence

Let us consider the shape dependence. First of all, the last term of Eq. (5.34) has the same shape dependence as $B_{f_{\text{NL}}}^{nG0}$, as $\mathcal{G}_R(k)$ is almost constant on large scale. Thus, it peaks at the squeezed configurations as $B_{f_{\text{NL}}}^{nG0}$ does. How about the shape dependence of the other terms?

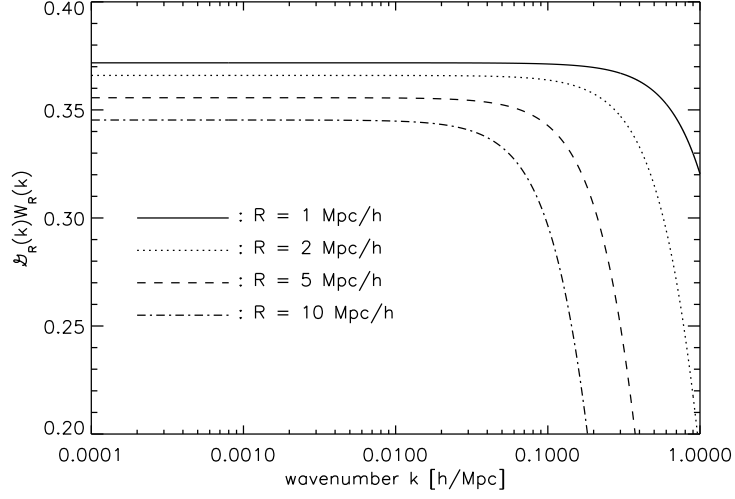


Figure 5.7: Shape of the function, $\mathcal{G}_R(k)$, defined in Eq. (5.38). We show $\mathcal{G}_R(k)$ for four different smoothing lengths: $R = 1, 2, 5, 10$ Mpc/ h .

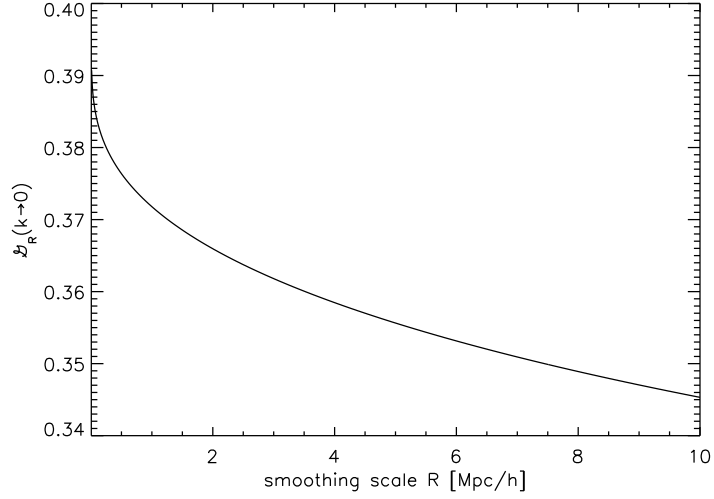


Figure 5.8: Large-scale asymptotic value of $\mathcal{G}_R(k)$ as a function of the smoothing scale R . The value for $R = 1$ [Mpc/ h], which is used for generating Figs. 5.10 to 5.14, is 0.3718.

All terms in Eqs. (5.32), (5.33), (5.41) have $P_\phi(k_i)$ outside of the integral, and Eq. (5.35) contains $1/\mathcal{M}_R(k) \propto k^{-2}$, which suggests that all of $B_{g_{\text{NL}}}^{nG}$, $B_{f_{\text{NL}}}^{nG}$, $B_{f_{\text{NL}}}^{nG1}$, and B_m^{nG} peak at the squeezed configurations. For sufficiently large scales in which $T(k) \approx 1$, and for a scale-invariant spectrum ($P_\phi(k) \propto 1/k^3$), we may write down Eqs. (5.35), (5.41), (5.33), and (5.32), as

$$B_m^{nG}(k_1, k_2, k_3) \propto \left(\frac{k_2}{k_1} + \frac{k_1}{k_2} \right) \left[\frac{5}{7} + \frac{\mathbf{k}_1 \cdot \mathbf{k}_2}{2k_1 k_2} \left(\frac{k_1}{k_2} + \frac{k_2}{k_1} \right) + \frac{2}{7} \left(\frac{\mathbf{k}_1 \cdot \mathbf{k}_2}{k_1 k_2} \right)^2 \right] + (2 \text{ cyclic}), \quad (5.42)$$

$$B_{f_{\text{NL}}}^{nG1}(k_1, k_2, k_3) \propto \frac{k_2}{k_3} \int \frac{d^3 q}{q} q^3 \left(\frac{|\mathbf{k}_1 - \mathbf{q}|^2}{|\mathbf{k}_2 + \mathbf{q}|} + \frac{|\mathbf{k}_2 + \mathbf{q}|^2}{|\mathbf{k}_1 - \mathbf{q}|} \right) \mathcal{W}_R(|\mathbf{k}_1 - \mathbf{q}|) \mathcal{W}_R(q) \times T(|\mathbf{k}_1 - \mathbf{q}|) T(|\mathbf{k}_2 + \mathbf{q}|) F_2^{(s)}(-\mathbf{k}_2, \mathbf{k}_2 + \mathbf{q}) + (5 \text{ permutation}), \quad (5.43)$$

$$B_{f_{\text{NL}}}^{nG}(k_1, k_2, k_3) \propto \frac{4}{k_2 k_3} \left[\frac{2(k_2^3 + k_3^3)}{k_1^3} \int \frac{d^3 q}{q} |\mathbf{k}_1 - \mathbf{q}|^2 T(q) \tilde{W}_R(q) \times T(|\mathbf{k}_1 - \mathbf{q}|) \tilde{W}_R(|\mathbf{k}_1 - \mathbf{q}|) + \int d^3 q q^2 |\mathbf{k}_1 - \mathbf{q}|^2 \times \left(\frac{1}{|\mathbf{k}_2 + \mathbf{q}|^3} + \frac{1}{|\mathbf{k}_3 + \mathbf{q}|^3} \right) T(q) \tilde{W}_R(q) \times T(|\mathbf{k}_1 - \mathbf{q}|) \tilde{W}_R(|\mathbf{k}_1 - \mathbf{q}|) \right] + (2 \text{ cyclic}), \quad (5.44)$$

$$B_{g_{\text{NL}}}^{nG}(k_1, k_2, k_3) \propto \frac{6}{k_2 k_3} \left[(k_2^3 + k_3^3) \int \frac{d^3 q}{q} \frac{1}{|\mathbf{k}_1 - \mathbf{q}|} T(q) \tilde{W}_R(q) \times T(|\mathbf{k}_1 - \mathbf{q}|) \tilde{W}_R(|\mathbf{k}_1 - \mathbf{q}|) + 2 \int \frac{d^3 q}{q} |\mathbf{k}_1 - \mathbf{q}|^2 T(q) \tilde{W}_R(q) \times T(|\mathbf{k}_1 - \mathbf{q}|) \tilde{W}_R(|\mathbf{k}_1 - \mathbf{q}|) \right] + (2 \text{ cyclic}), \quad (5.45)$$

respectively. For a given k_1 , all of these terms have the largest signals when k_3 is small, i.e., the squeezed configurations. Note that we do not use the exact scale-invariant spectrum for the numerical calculation, but use the WMAP 5-year best-fitting value reported in Table 1 (“WMAP+BAO+SN”) of Komatsu et al. (2009).

The top-left and bottom-left panels of Fig. 5.4 show B_m^{nG} and $B_{f_{\text{NL}}}^{nG1}$ as a function of k_2/k_1 and k_3/k_1 , respectively, for $k_1 = 0.01 h \text{ Mpc}^{-1}$. The top-right and bottom-right

panels of Fig. 5.4 show the same quantities for $k_1 = 0.05 \, h \, \text{Mpc}^{-1}$. We also show $B_{g_{\text{NL}}}^{nG}$ and $B_{f_{\text{NL}}}^{nG}$ in the top-left and bottom-left panels of Fig. 5.5 for $k_1 = 0.01 \, h \, \text{Mpc}^{-1}$, and top-right and bottom-right for $k_1 = 0.05 \, h \, \text{Mpc}^{-1}$. In all cases we find that B_m^{nG} , $B_{f_{\text{NL}}}^{nG1}$, $B_{g_{\text{NL}}}^{nG}$ and $B_{f_{\text{NL}}}^{nG}$ peak at the squeezed configurations, as expected from the above argument.

We find that the shape dependence of $B_{f_{\text{NL}}}^{nG0}$ and that of $B_{f_{\text{NL}}}^{nG1}$, $B_{g_{\text{NL}}}^{nG}$ are quite similar, whereas that of B_m^{nG} is higher toward the elongated triangles, and that of $B_{f_{\text{NL}}}^{nG}$ is more sharply peaked at the squeezed configuration.

We can understand this behavior analytically as follows. In order to simplify the analysis, we consider a scale-invariant curvature power spectrum, $P_\phi = P_{\phi 0}/k^3$, on large scales where Eqs. (5.43), (5.42), (5.44), and (5.45) are valid. On such a large scale, $\mathcal{M}_R(k)$ can be approximated as $\mathcal{M}_R(k) \simeq 2k^2/(3H_0^2\Omega_m) \equiv \mathcal{M}_0 k^2$, where $\mathcal{M}_0 \simeq 2.16 \times 10^7 (0.277/\Omega_m) [\text{Mpc}/h]^2$ is a constant. We focus on the squeezed triangle, $k_1 = k_2 = \alpha k_3 \equiv k$ ($\alpha \gg 1$), where the signals of all the primordial non-Gaussianity terms are maximized. The triangles in this configuration lie on the upper side of the triangular region of $(k_3/k_1, k_2/k_1)$ plane in Fig. 5.4 and Fig. 5.5, and the triangle approaches the exact squeezed limit as $\alpha \rightarrow \infty$. With this parametrization, we compare the dominant contributions of each of these primordial non-Gaussianity terms.

First, we shall analyze the terms proportional to f_{NL} : $B_{f_{\text{NL}}}^{nG0}$, B_m^{nG} , and $B_{f_{\text{NL}}}^{nG1}$. The largest contribution to $B_{f_{\text{NL}}}^{nG0}$ in the squeezed configurations occurs when k_3 is in the denominator:

$$\begin{aligned} B_{f_{\text{NL}}}^{nG0} &= 2\mathcal{M}_0^3 P_{\phi 0}^2 \left(\frac{k_1^2}{k_2 k_3} + \frac{k_2^2}{k_3 k_1} + \frac{k_3^2}{k_1 k_2} \right) \\ &\simeq 2\mathcal{M}_0^3 P_{\phi 0}^2 \left(\frac{k_1^2}{k_2 k_3} + \frac{k_2^2}{k_3 k_1} \right) \\ &= 4\alpha \mathcal{M}_0^3 P_{\phi 0}^2. \end{aligned} \tag{5.46}$$

To compute B_m^{nG} , which contains $F_2^{(s)}$, we note that, in the squeezed limit, the angular cosines between two wave vectors are $\mathbf{k}_1 \cdot \mathbf{k}_2/(k_1 k_2) = -1 + 1/\alpha^2 \simeq -1$ and $\mathbf{k}_2 \cdot \mathbf{k}_3/(k_2 k_3) = \mathbf{k}_1 \cdot \mathbf{k}_3/(k_1 k_3) = -1/(2\alpha)$. We thus find

$$\begin{aligned} B_m^{nG} &= 8\mathcal{M}_0^3 P_{\phi 0}^2 \left(\alpha + \frac{1}{\alpha} \right) \left[\frac{5}{7} - \frac{1}{4\alpha} \left(\alpha + \frac{1}{\alpha} \right) + \frac{1}{14\alpha^2} \right] \\ &\simeq \frac{26}{7} \alpha \mathcal{M}_0^3 P_{\phi 0}^2. \end{aligned} \tag{5.47}$$

The detailed analysis for $B_{f_{\text{NL}}}^{nG1}$ is more complicated, as Eq. (5.41) involves a non-trivial integration. We simplify the situation by only analyzing the dominant term, which

can be written as

$$\begin{aligned} B_{f_{\text{NL}}}^{nG1} &\approx 8\mathcal{M}_0^3 P_{\phi 0}^2 \left[\frac{k_2}{k_3} \mathcal{H}(\mathbf{k}_1, \mathbf{k}_2) + \frac{k_1}{k_3} \mathcal{H}(\mathbf{k}_2, \mathbf{k}_1) \right] \\ &= 8\alpha \mathcal{M}_0^3 P_{\phi 0}^2 [\mathcal{H}(\mathbf{k}_1, \mathbf{k}_2) + \mathcal{H}(\mathbf{k}_2, \mathbf{k}_1)], \end{aligned} \quad (5.48)$$

where $\mathcal{H}(\mathbf{k}_1, \mathbf{k}_2)$ is the integration that appears in Eq. (5.41) including $1/(2\sigma_R^2)$ pre-factor. Note that this integration depends only on the magnitudes of two vectors and the angle between them; thus, for the squeezed configuration we are interested in here, $\mathcal{H}(\mathbf{k}_2, \mathbf{k}_1)$ depends only weakly on α – they depend on α only through the inner product of $\mathbf{k}_1 \cdot \mathbf{k}_2 = k^2(-1 + 1/\alpha^2)$.

Second, we analyze $B_{g_{\text{NL}}}^{nG}$. We find that the first cyclic terms in Eq. (5.32) are small in the squeezed limit, and the dominant contribution to $B_{g_{\text{NL}}}^{nG}$ is given by

$$\begin{aligned} B_{g_{\text{NL}}}^{nG} &= 12\mathcal{M}_0^2 P_{\phi 0}^2 \left[\frac{\mathcal{J}(k_1)}{k_2 k_3} + \frac{\mathcal{J}(k_2)}{k_3 k_1} + \frac{\mathcal{J}(k_3)}{k_1 k_2} \right] \\ &\simeq 12\mathcal{M}_0^2 P_{\phi 0}^2 \left[\frac{\mathcal{J}(k_1)}{k_2 k_3} + \frac{\mathcal{J}(k_2)}{k_3 k_1} \right], \end{aligned} \quad (5.49)$$

where we have defined

$$\mathcal{J}(k) \equiv \frac{1}{2\sigma_R^2} \int \frac{d^3 q}{(2\pi)^3} \mathcal{M}_R(q) \mathcal{M}_R(|\mathbf{k} - \mathbf{q}|) P_{\phi}(q). \quad (5.50)$$

We find that $\mathcal{J}(k) \simeq 0.5$ and is almost independent of k on large scales (e.g., $k \lesssim 0.03 \, h \, \text{Mpc}^{-1}$ for $R = 1.0 \, \text{Mpc}/h$; see Fig. 5.9). Therefore, by writing $\mathcal{J}(k) = \mathcal{J}_0$, we obtain

$$B_{g_{\text{NL}}}^{nG} \simeq 24\alpha \mathcal{M}_0^2 P_{\phi 0}^2 \frac{\mathcal{J}_0}{k^2}. \quad (5.51)$$

These results show that all the terms we have analyzed analytically so far, $B_{f_{\text{NL}}}^{nG0}$, B_m^{nG} , $B_{f_{\text{NL}}}^{nG1}$, and $B_{g_{\text{NL}}}^{nG}$, have the same shape (i.e., α) dependence in the squeezed configurations: they both increase linearly as α increases. This explains the shape dependence computed from the full numerical calculations presented in Fig. 5.4 and the top panels of Fig. 5.5.

Finally, we analyze $B_{f_{\text{NL}}}^{nG2}$. We find that the second cyclic terms in Eq. (5.33) are

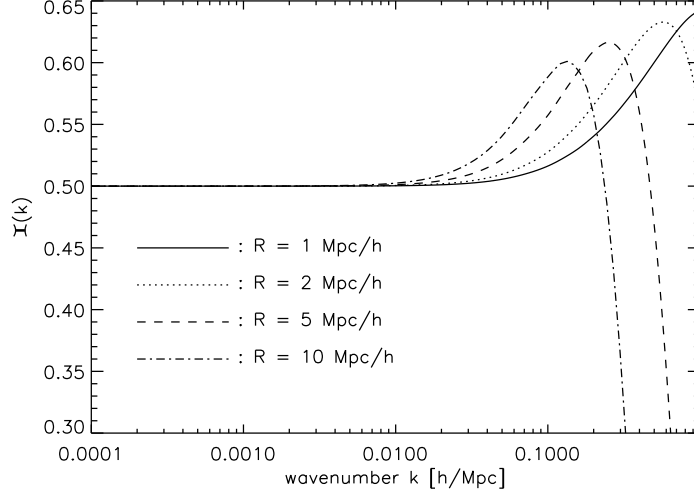


Figure 5.9: Shape of the integration that appears in the dominant term of $B_{f_{\text{NL}}^2}^{nG}$ and $B_{g_{\text{NL}}}^{nG}$, Eq. (5.50). We use four different smoothing scales: $R = 1, 2, 5$, and $10 \text{ Mpc}/h$.

small in the squeezed configurations. The dominant terms are:

$$\begin{aligned}
 B_{f_{\text{NL}}^2}^{nG} &= 8\mathcal{M}_0^2 P_{\phi 0}^2 \\
 &\times \left[\frac{k_2^3 + k_3^3}{k_2 k_3 k_1^3} \mathcal{J}(k_1) + \frac{k_3^3 + k_1^3}{k_3 k_1 k_2^3} \mathcal{J}(k_2) + \frac{k_1^3 + k_2^3}{k_1 k_2 k_3^3} \mathcal{J}(k_3) \right] \\
 &\simeq 8\mathcal{M}_0^2 P_{\phi 0}^2 \frac{k_1^3 + k_2^3}{k_1 k_2 k_3^3} \mathcal{J}(k_3) \\
 &\simeq 16\alpha^3 \mathcal{M}_0^2 P_{\phi 0}^2 \frac{\mathcal{J}_0}{k^2}.
 \end{aligned} \tag{5.52}$$

Therefore, $B_{f_{\text{NL}}^2}^{nG}$ increases more sharply as it approaches the squeezed limit, $B_{f_{\text{NL}}^2}^{nG} \propto \alpha^3$.

This sharp increase of $B_{f_{\text{NL}}^2}^{nG}$ relative to the other terms, and that there are many new terms that are of the same order of magnitude as $B_{f_{\text{NL}}}^{nG0}$, imply that the formula derived by Sefusatti & Komatsu (2007) may not be valid in the squeezed configuration, where $B_{f_{\text{NL}}^2}^{nG}$ may dominate over $B_{f_{\text{NL}}}^{nG}$. This is particularly important because it is the squeezed configuration that gives the largest signal from the primordial non-Gaussianity. We shall study this point in more detail in the next section.

A careful inspection of Eq. (5.44) shows that the second term within the square bracket diverges when $\mathbf{k}_2 + \mathbf{q} = 0$ or $\mathbf{k}_3 + \mathbf{q} = 0$. This is due to the fact that $P_\phi(k) \propto 1/k^{4-n_s}$

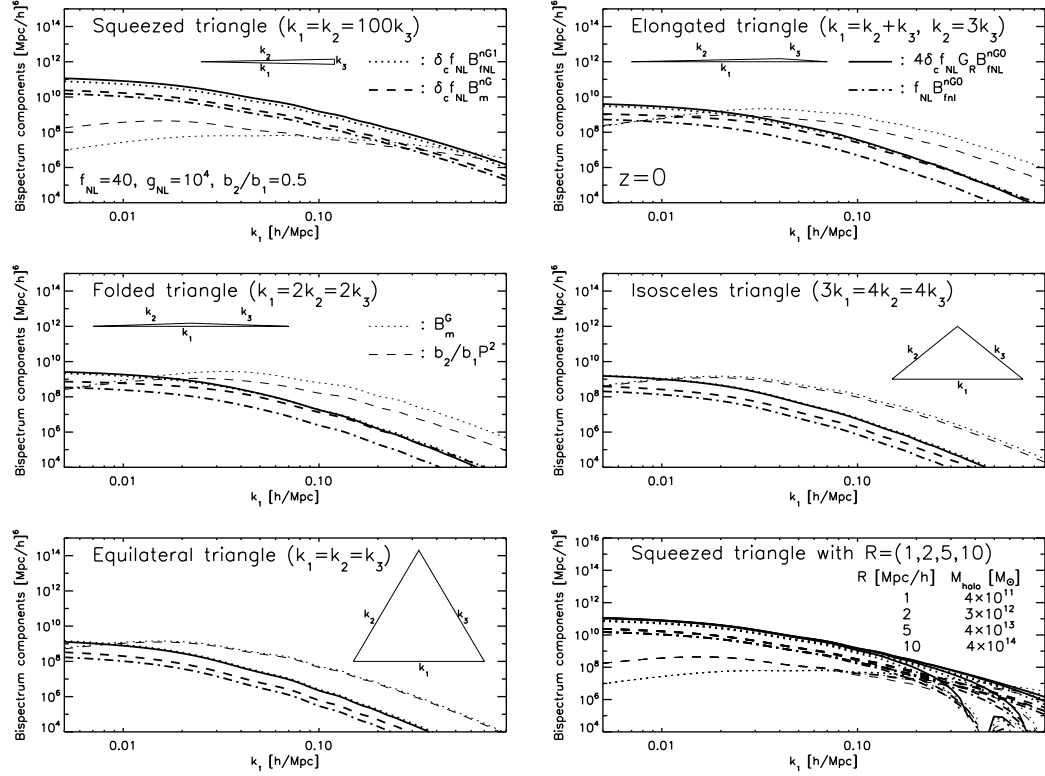


Figure 5.10: Scale and shape dependence of the galaxy bispectrum terms that are linearly proportional to f_{NL} , as a function of k_1 . Except for the bottom-right panel, we use $R = 1$ Mpc/h. (Top Left) The squeezed triangles with $k_1 = k_2 = 100k_3$, (Top Right) the elongated triangles with $k_1 = k_2 + k_3$ and $k_2 = 3k_3$, (Middle Left) the folded triangles with $k_1 = 2k_2 = 2k_3$, (Middle Right) the isosceles triangles with $3k_1 = 4k_2 = 4k_3$, and (Bottom Left) the equilateral triangles with $k_1 = k_2 = k_3$. The thick dot-dashed, dashed, solid, and dotted lines show the contributions from the primordial non-Gaussianity: the B_{fNL}^{nG0} (Eq. (5.24)), $\tilde{b}_2/b_1 B_m^{nG}$ (Eq. (5.35)), $4(\tilde{b}_2/b_1) [\mathcal{G}_R(k_1) + \mathcal{G}_R(k_2) + \mathcal{G}_R(k_3)] B_{\text{fNL}}^{nG0}$ ($\mathcal{G}_R(k)$ defined in Eq. (5.38)), and $\tilde{b}_2/b_1 B_{\text{fNL}}^{nG1}$ (Eq. (5.41)) terms, respectively. The thin dotted and dashed lines show the non-linear effects: B_m^G (Eq. (5.23)) and the non-linear bias (the second term in Eq. (5.11)), respectively. We use the standard value of $\tilde{b}_2/b_1 \equiv \delta_c \simeq 1.686$ from spherical collapse model. (Bottom Right) Dependence of the squeezed bispectrum on the smoothing scale, R , showing that the dependence is negligible for $k_1 \ll 1/R$.

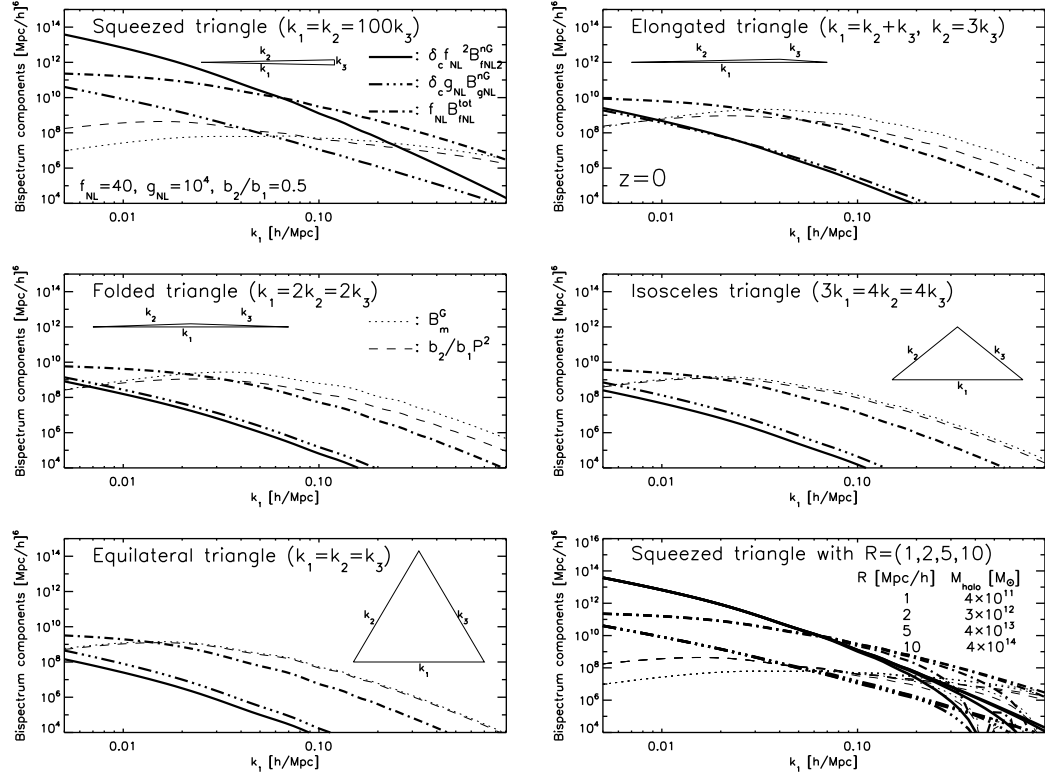


Figure 5.11: Scale and shape dependence of various bispectrum terms, $B(k_1, k_2, k_3)$, as a function of k_1 . For the figure except for the bottom right, we use $R = 1 \text{ Mpc}/h$. (Top Left) The squeezed triangles with $k_1 = k_2 = 100k_3$, (Top Right) the elongated triangles with $k_1 = k_2 + k_3$ and $k_2 = 3k_3$, (Middle Left) the folded triangles with $k_1 = 2k_2 = 2k_3$, (Middle Right) the isosceles triangles with $3k_1 = 4k_2 = 4k_3$, and (Bottom Left) the equilateral triangles with $k_1 = k_2 = k_3$. The thick dot-dashed, triple-dot-dashed, and solid lines show the contributions from the primordial non-Gaussianity: the $f_{\text{NL}} B_{f_{\text{NL}}}^{\text{tot}}$ (Eq. (5.53)), $\tilde{b}_2/b_1 g_{\text{NL}} B_{g_{\text{NL}}}^{\text{nG}}$ (Eq. (5.32)), and $\tilde{b}_2/b_1 f_{\text{NL}}^2 B_{f_{\text{NL}}^2}^{\text{nG}}$ (Eq. (5.33)) terms, respectively. The thin dotted and dashed lines show the non-linear effects: B_m^G (Eq. (5.23)) and the non-linear bias (the second term in Eq. (5.11)), respectively. We use the standard value of $\tilde{b}_2/b_1 \equiv \delta_c \simeq 1.686$ from spherical collapse model. (Bottom Right) Dependence of the squeezed bispectrum on the smoothing scale, R , showing that the dependence is negligible for $k_1 \ll 1/R$.

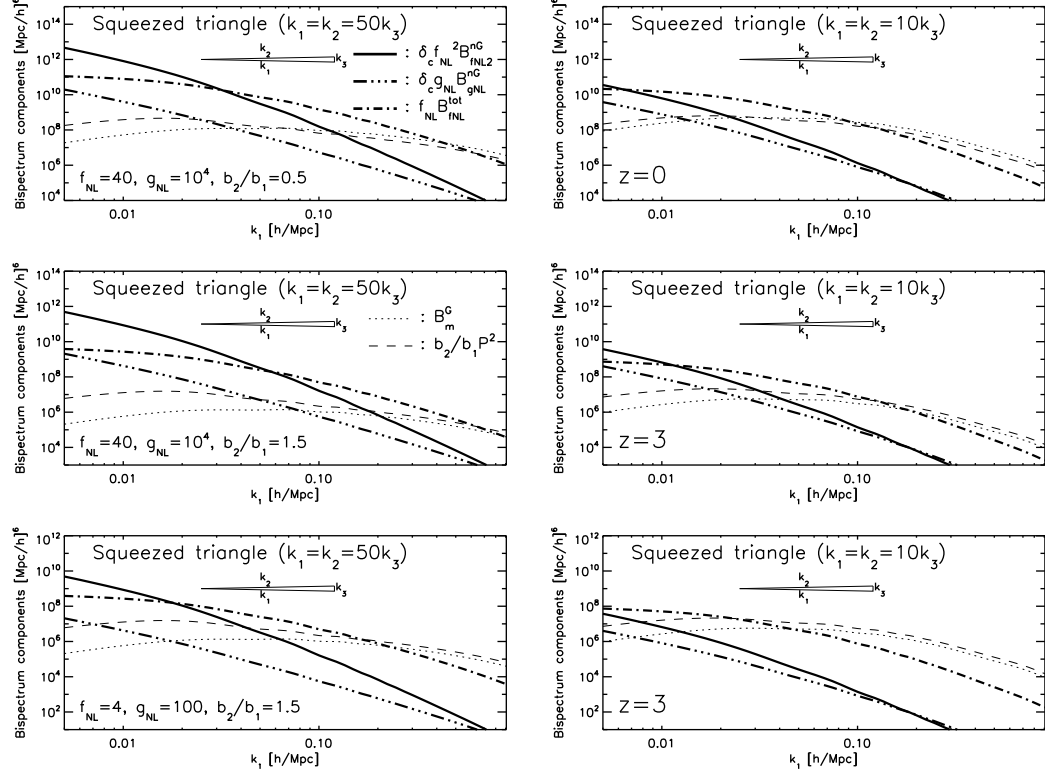


Figure 5.12: Same as Fig. 5.11, but for squeezed triangles with different ratios: $\alpha = 50$ and $\alpha = 10$. (Top) All the parameters are the same as in Fig. 5.11. (Middle) $z = 3$ and $b_2/b_1 = 1.5$. The non-Gaussianity parameters, $f_{\text{NL}} = 40$ and $g_{\text{NL}} = 10^4$, are the same as in Fig. 5.11. (Bottom) $z = 3$ and $b_2/b_1 = 1.5$. The non-Gaussianity parameters, $f_{\text{NL}} = 4$ and $g_{\text{NL}} = 100$.

and thus $P_\phi(k)$ diverges as $k \rightarrow 0$ for $n_s < 4$. To avoid the divergence we set $P_\phi(k) = 0$ at $k \leq k_{min}$, and use $k_{min} = 10^{-6} h \text{ Mpc}^{-1}$. Fortunately the divergence is mild and the results on the squeezed configurations, for which $B_{f_{NL}^2}^{nG}$ gives the most important contribution, are insensitive to k_{min} : changing $k_{min} = 10^{-6} h \text{ Mpc}^{-1}$ to $k_{min} = 10^{-9} h \text{ Mpc}^{-1}$ results in negligible changes in the squeezed configurations.

On the other hand, the folded and equilateral configurations are more sensitive to k_{min} , and we find that the difference between $k_{min} = 10^{-6} h \text{ Mpc}^{-1}$ and $k_{min} = 10^{-9} h \text{ Mpc}^{-1}$ is scale-dependent: at $k_1 = 0.01 h \text{ Mpc}^{-1}$ the differences are negligible for all shapes, whereas the differences reach $\sim 40\%$ at $k \sim 1 h \text{ Mpc}^{-1}$. (Note that the difference in the squeezed configuration reaches 1% at $k \sim 1 h \text{ Mpc}^{-1}$, being totally negligible on larger scales.) While this divergence does not have much observational consequences (because the signals in the other configurations at $k \gtrsim 0.01 h \text{ Mpc}^{-1}$ would be dominated by the other non-linear effects: B_m^G , non-linear bias and terms proportional to f_{NL} , as we show in the next section), there may be a better treatment of the divergence than setting $P_\phi(k) = 0$ at $k \leq k_{min}$.

5.2.3 Scale Dependence

How important are the primordial non-Gaussianity terms, $B_{f_{NL}}^{nG0}$, $B_{f_{NL}}^{nG1}$, B_m^{nG} , $B_{f_{NL}^2}^{nG}$, and $B_{g_{NL}}^{nG}$, relative to B_m^G and the non-linear bias term? Which one is the most dominant of the primordial terms, terms proportional to f_{NL} , $B_{f_{NL}^2}^{nG}$, or $B_{g_{NL}}^{nG}$? How about the scale-dependence? How about the shape dependence?

We collect all the terms proportional to f_{NL} , and call it $B_{f_{NL}}^{tot}$:

$$B_{f_{NL}}^{tot} \equiv B_{f_{NL}}^{nG0} + \frac{\tilde{b}_2}{b_1} [B_m^{nG} + B_{f_{NL}}^{nG1} + 4(\mathcal{G}_R(k_1) + \mathcal{G}_R(k_2) + \mathcal{G}_R(k_3)) B_{f_{NL}}^{nG0}]. \quad (5.53)$$

Throughout this section, we use the standard value of $\tilde{b}_2/b_1 = 3(12\pi)^{2/3}/20 \simeq 1.68$ from a spherical collapse model.

Figure 5.10 shows the scale and shape dependence of each term in Eq. (5.53) evaluated at $z = 0$. For all configurations shown in this figure, the primordial non-Gaussian term calculated in Sefusatti & Komatsu (2007) is the smallest among four f_{NL} terms, which means that the non-Gaussian signal on large scales is much bigger than recognized before.

For the squeezed triangle, all of the terms in Eq. (5.53) depend on k_1 in a similar

way. We find their ratios by comparing Eqs. (5.46), (5.47), and (5.48):

$$B_{f_{\text{NL}}}^{nG0} : B_m^{nG} : B_{f_{\text{NL}}}^{nG1} \simeq 1 : \frac{26}{28} : 2.96. \quad (5.54)$$

Note that we have used the numerical value of $\mathcal{H}(\mathbf{k}_1, \mathbf{k}_2) \simeq 0.741$ for $\alpha = 100$, and this value slightly increases when α decreases⁹. Therefore, for the squeezed triangle, we find a simple and illuminating result:

$$B_{f_{\text{NL}}}^{\text{tot}}(k_1, k_2, k_3) \simeq 15 B_{f_{\text{NL}}}^{nG0}(k_1, k_2, k_3). \quad (5.55)$$

This is an important result, showing that the statistical error on f_{NL} from the galaxy bispectrum will be smaller by at least a factor of 15, compared to what was obtained in Sefusatti & Komatsu (2007). Note that this result is valid only for the high-density peak limit, i.e., $\nu \gg 1$, which implies $\tilde{b}_2/b_1 = 1.68$. For lower density peaks we would find a smaller factor.

Figure 5.11 and the top panels of Figure 5.12 show various bispectrum terms in various triangle configurations (see Fig. 5.1 for the visual representations of the triangles), evaluated at $z = 0$. As an example we use the following bias and non-Gaussianity parameters: $b_2/b_1 = 0.5$, $f_{\text{NL}} = 40$, and $g_{\text{NL}} = 10^4$. The value of the linear bias, b_1 , is irrelevant here as it does not change the relative importance of terms in Eq. (5.16), and thus we show the bispectrum terms divided by b_1^3 .

The message is quite simple: it is the squeezed configuration that provides the best window into the primordial non-Gaussianity. The other non-linear effects become more and more dominant as we move from the squeezed to the equilateral, i.e., (a) to (e) in Fig. 5.1. Even with this generous amount of non-Gaussian signals, $f_{\text{NL}} = 40$ and $g_{\text{NL}} = 10^4$, only f_{NL} term can be visible in the isosceles and equilateral configurations on large scales.

For the the non-squeezed configurations, the f_{NL}^2 and g_{NL} terms with the above chosen parameters are comparable and the f_{NL} term is order of magnitude greater than the f_{NL}^2 and g_{NL} terms; however, the f_{NL}^2 term is the most dominant of all on large scales in the squeezed configuration ($\alpha > 10$).

We can understand these results analytically by comparing Eqs. (5.55), (5.46), (5.51), and (5.52). For the squeezed triangles with $k_1 = k_2 = \alpha k_3$ ($\alpha \gg 1$) and a scale-

⁹On large scales, $k < 0.01 \text{ h/Mpc}$, the numerical ratio $B_{f_{\text{NL}}}^{nG1}/B_{f_{\text{NL}}}^{nG0}$ is constant, and is equal to 3.15, 3.06, 3.00, and 2.98 for $\alpha = 10, 20, 50$, and 100, respectively.

invariant spectrum, $P_\phi \propto k^{-3}$, we find

$$\begin{aligned} \frac{f_{\text{NL}} B_{f_{\text{NL}}}^{\text{tot}}}{f_{\text{NL}}^2 B_{f_{\text{NL}}}^{nG}} &\simeq \frac{15}{f_{\text{NL}} \alpha^2} \frac{\mathcal{M}_0 k^2}{4\mathcal{J}_0(\tilde{b}_2/b_1)} \\ &\simeq 0.0240 \left(\frac{100}{\alpha}\right)^2 \frac{40}{f_{\text{NL}}} \left(\frac{k}{0.01 \text{ h Mpc}^{-1}}\right)^2, \end{aligned} \quad (5.56)$$

$$\begin{aligned} \frac{g_{\text{NL}} B_{g_{\text{NL}}}^{nG}}{f_{\text{NL}}^2 B_{f_{\text{NL}}}^{nG}} &\simeq \frac{3}{2\alpha^2} \frac{g_{\text{NL}}}{f_{\text{NL}}^2} \\ &\simeq 0.000938 \left(\frac{100}{\alpha}\right)^2 \left(\frac{40}{f_{\text{NL}}}\right)^2 \frac{g_{\text{NL}}}{10^4}, \end{aligned} \quad (5.57)$$

$$\begin{aligned} \frac{f_{\text{NL}} B_{f_{\text{NL}}}^{\text{tot}}}{g_{\text{NL}} B_{g_{\text{NL}}}^{nG}} &\simeq 15 \frac{f_{\text{NL}}}{g_{\text{NL}}} \frac{\mathcal{M}_0 k^2}{6\mathcal{J}_0} \\ &\simeq 25.6 \frac{f_{\text{NL}}}{40} \frac{10^4}{g_{\text{NL}}} \left(\frac{k}{0.01 \text{ h Mpc}^{-1}}\right)^2. \end{aligned} \quad (5.58)$$

These estimates confirm that $B_{f_{\text{NL}}^2}$ dominates over $B_{f_{\text{NL}}}$ and $B_{g_{\text{NL}}}$ in the squeezed configurations on large scales, $k \lesssim 0.05 \text{ h Mpc}^{-1}$ for $\alpha = 100$, and $k \lesssim 0.03 \text{ h Mpc}^{-1}$ for $\alpha = 50$. For $\alpha = 10$, f_{NL}^2 term dominates only on the extremely large scales: $k \lesssim 0.006 \text{ h Mpc}^{-1}$.

Note that for a given configuration (for a given α), $B_{f_{\text{NL}}}^{nG}/B_{f_{\text{NL}}}^{nG} \propto k^2$ and $B_{f_{\text{NL}}}^{nG}/B_{g_{\text{NL}}}^{nG} \propto k^2$ while $B_{g_{\text{NL}}}^{nG}/B_{f_{\text{NL}}}^{nG}$ is independent of k , which is consistent with what we show in Fig. 5.11 on $k \lesssim 0.1 \text{ h Mpc}^{-1}$.

In summary, the most unexpected and important results of our study are as follows.

- The terms that are linearly proportional to f_{NL} , derived in Sefusatti & Komatsu (2007), receive additional contributions, and are enhanced by a factor of ~ 15 for the squeezed triangles (see Eq. (5.55)).
- The f_{NL}^2 (or τ_{NL}) term actually dominates over the f_{NL} term by a large factor for the squeezed triangles (see the top-left panel of Fig. 5.11).

This suggests that the galaxy bispectrum is more sensitive to f_{NL} than previously recognized by Sefusatti & Komatsu (2007), greatly enhancing our ability to detect the primordial non-Gaussianity of local type. On very large scales, $k_1 \ll 0.01 \text{ h Mpc}^{-1}$, even the g_{NL} term (with $g_{\text{NL}} = 10^4$) dominates over the f_{NL} term, giving us a hope that perhaps we can obtain a meaningful limit on this term using the galaxy bispectrum.

5.2.4 Redshift Dependence

The quantities we have calculated so far are evaluated at the present epoch, $z = 0$. At higher redshift, each quantity needs to be scaled with some powers of the linear growth factor $D(z)$, which is normalized to 1 at the present epoch. We find $P_R \propto D^2(z)$, $B_m^G \propto D^4(z)$, $B_{f_{\text{NL}}}^{nG0} \propto D^3(z)$, $B_m^{nG} \propto D^3(z)$, $B_{f_{\text{NL}}}^{nG1} \propto D^3(z)$, $B_{f_{\text{NL}}}^{nG} \propto D^2(z)$, and $B_{g_{\text{NL}}}^{nG} \propto D^2(z)$. Therefore, the final result for the halo bispectrum from the local type non-Gaussianity is

$$\begin{aligned}
& B_g(k_1, k_2, k_3, z) \\
= & b_1^3(z) D^4(z) \left[B_m^G(k_1, k_2, k_3) + \frac{b_2(z)}{b_1(z)} \{P_R(k_1)P_R(k_2) + (2 \text{ cyclic})\} \right. \\
& + \frac{f_{\text{NL}}}{D(z)} B_{f_{\text{NL}}}^{nG0}(k_1, k_2, k_3) + \frac{\tilde{b}_2(z)}{b_1(z)} \left\{ \frac{f_{\text{NL}}}{D(z)} \left(B_m^{nG}(k_1, k_2, k_3) \right. \right. \\
& \left. \left. + 4(\mathcal{G}_R(k_1) + \mathcal{G}_R(k_2) + \mathcal{G}_R(k_3)) B_{f_{\text{NL}}}^{nG0}(k_1, k_2, k_3) + B_{f_{\text{NL}}}^{nG1}(k_1, k_2, k_3) \right) \right. \\
& \left. \left. + \frac{f_{\text{NL}}^2}{D^2(z)} B_{f_{\text{NL}}}^{nG}(k_1, k_2, k_3) + \frac{g_{\text{NL}}}{D^2(z)} B_{g_{\text{NL}}}^{nG}(k_1, k_2, k_3) \right\} \right], \tag{5.59}
\end{aligned}$$

where B_m^G , P_R , $B_{f_{\text{NL}}}^{nG0}$, B_m^{nG} , $B_{f_{\text{NL}}}^{nG1}$, $B_{f_{\text{NL}}}^{nG}$, and $B_{g_{\text{NL}}}^{nG}$ are evaluated at $z = 0$.

From equation (5.59) it is clear that the contributions from non-Gaussian initial conditions become more and more important as we go to higher redshifts. The new terms that we have derived in this paper, the $B_{f_{\text{NL}}}^{nG}$ and $B_{g_{\text{NL}}}^{nG}$ terms, are even more important than the term derived by Sefusatti & Komatsu (2007), $B_{f_{\text{NL}}}^{nG0}$. This property makes high-redshift galaxy surveys particularly a powerful probe of primordial non-Gaussianity.

Fig. 5.13 and the middle panel of Fig. 5.12 show the bispectrum terms at $z = 3$. Note that we use a larger value for the non-linear bias, $b_2/b_1 = 1.5$, in accordance with a halo model (Cooray & Sheth, 2002). At this redshift, with $f_{\text{NL}} = 40$ and $g_{\text{NL}} = 10^4$, the g_{NL} and f_{NL}^2 terms dominate over the non-linear effects also in the elongated, folded and isosceles configurations at $k \lesssim 0.01 \ h \text{ Mpc}^{-1}$, as well as in the squeezed ones. The f_{NL} terms dominate over the non-linear effects on even smaller scales, and the importance of the f_{NL}^2 and g_{NL} terms relative to the f_{NL} term is greater, as expected from their dependence on $D(z)$.

5.3 Discussion and Conclusions

Let us come back to this question, “can we still use Sefusatti & Komatsu’s equation, Eq. (5.3), with b_1 replaced by the scale-dependent bias, Eq. (5.6)?” The answer is clearly

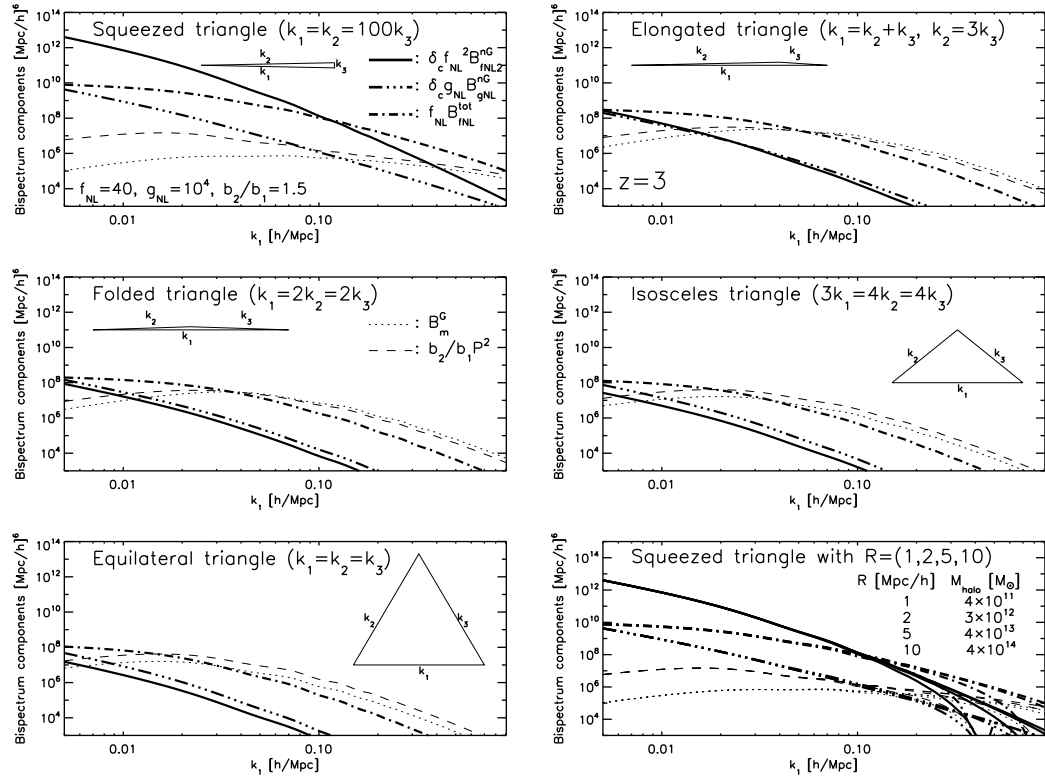


Figure 5.13: Same as Fig. 5.11, but for $z = 3$ and $b_2/b_1 = 1.5$. The non-Gaussianity parameters, $f_{\text{NL}} = 40$ and $g_{\text{NL}} = 10^4$, are the same as in Fig. 5.11.

no: the primordial non-Gaussianity gives the largest signal in the squeezed limit, whereas the non-linear gravitational evolution and non-linear bias give the minimal signals in the same limit. This means that these effects are physically totally distinct, and thus a mere scale-dependent rescaling of one effect does not give another. Therefore, replacing b_1 in Eq. (5.3) with the scale-dependent bias in Eq. (5.6) results in an incorrect prediction. For example, even though we have a term similar to that of the scale-dependent bias, B_m^{nG} , in our final expression of the galaxy bispectrum for the local-type primordial non-Gaussianity, there are many more terms that do not look like the scale-dependent bias that appears in the galaxy power spectrum. Furthermore, B_m^{nG} is by no means the most dominant term.

In this paper, we have derived a general expression for the bispectrum of density peaks in the presence of primordial non-Gaussianity (Eq. (5.11)), using the MLB formula as well as using the local bias ansatz. This result is general as long as we consider the bispectrum of high density peaks, i.e., $\nu = \delta_c/\sigma_R \gg 1$, which is equivalent to highly biased galaxy populations, $b_1 \gg 1$, on large scales in which the n -point correlation functions are much smaller than unity. (This condition was necessary for us to Taylor expand the exponential in Eq. (5.8).)

We have applied our formula to the local form of primordial non-Gaussianity in Bardeen’s curvature perturbations, $\Phi = \phi + f_{\text{NL}}\phi^2 + g_{\text{NL}}\phi^3$, and found new terms that are proportional to f_{NL} , f_{NL}^2 and g_{NL} , which were absent in the formula derived by Sefusatti & Komatsu (2007). We have examined the shape and scale dependence of these new terms as well as those of the known terms, and found that the primordial non-Gaussianity contributions yield the largest signals in the squeezed triangle configurations, where the non-linear gravitational evolution and non-linear bias yield the minimal signals. This is a good news: this property enables us to distinguish the primordial and non-primordial effects easily.

The effects of primordial non-Gaussianity on the galaxy bispectrum are more important in a high redshift universe, and thus high-redshift galaxy surveys are particularly a potent probe of the physics of inflation via measurements of primordial non-Gaussianity.

The most significant conclusion of this paper is that, in the squeezed configurations, the f_{NL}^2 term actually dominates over the f_{NL} term by a large factor, and, on large scales, newly derived f_{NL} term dominates over the non linear terms for all configurations. Because of this, the galaxy bispectrum should be more sensitive to f_{NL} than previously recognized: in the high density peak limit, we have found a factor of ~ 15 enhancement for the f_{NL} term studied in Sefusatti & Komatsu (2007). In addition it is also sensitive to a new term, g_{NL} . Figure 5.14 and the bottom panel of Fig. 5.12 shows the bispectrum at $z = 3$ with much

reduced primordial non-Gaussianity parameters, $f_{\text{NL}} = 4$ and $g_{\text{NL}} = 100$. In the squeezed configurations ($\alpha = 100$), the f_{NL}^2 term is still well above the usual terms from Gaussian initial conditions at $k \lesssim 0.1 h \text{ Mpc}^{-1}$, the f_{NL} term is above at $k \lesssim 0.4 \text{ Mpc}^{-1}$, and the g_{NL} term is above at $k \lesssim 0.01 h \text{ Mpc}^{-1}$. Even with the milder squeezed limit for $\alpha = 10$, the f_{NL} term still is above the Gaussian term at $k \lesssim 0.02 \text{ Mpc}^{-1}$.

The fact that the f_{NL}^2 term dominates in the squeezed limit is particularly interesting, as it provides us with the unique window into the physics of inflation in the following way. Recently, a number of groups (e.g., Boubekur & Lyth, 2006; Huang & Shiu, 2006; Byrnes et al., 2006; Buchbinder et al., 2008) have shown that the primordial trispectrum can in general be written as

$$\begin{aligned} T_{\Phi}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) \\ = 6g_{\text{NL}} [P_{\phi}(k_1)P_{\phi}(k_2)P_{\phi}(k_3) + (3 \text{ cyclic})] + \frac{25}{18}\tau_{\text{NL}} \\ \times [P_{\phi}(k_1)P_{\phi}(k_2) \{P_{\phi}(k_{13}) + P_{\phi}(k_{14})\} + (11 \text{ cyclic})], \end{aligned} \quad (5.60)$$

instead of Eq. (5.19). Different models of the early universe predict different relations between τ_{NL} and f_{NL} . Therefore, separately detecting the τ_{NL} (i.e., f_{NL}^2) and f_{NL} terms can be a powerful tool for constraining the model of the early universe.

How well can one constrain these parameters with the current or planned future high-redshift galaxy surveys? As we based our analysis in this paper on the assumption of high density peaks, i.e., $\nu \gg 1$, the relative importance of the new terms depends on how high the peaks (in which the observed galaxies reside) are. For example, for SDSS-LRG sample ($z = 0.315$, $n_g = 1.36 \times 10^{-4} [h/\text{Mpc}]^3$) where $\delta_c/\sigma_R \simeq 1.57$, the halo model (See Sefusatti & Komatsu, 2007, for a detailed method) gives $\tilde{b}_2/b_1 \simeq 0.5$, which is about a third of the value from the high peak limit that we have used in this paper. Therefore, we expect that the non-Gaussian signal for the SDSS-LRG bispectrum is smaller by the same factor. Detailed analysis will be presented in a forthcoming paper.

5.4 Summary of Equations

As various terms contributing to the galaxy bispectrum are scattered over various places in the paper, we collect them together in this Appendix. For galaxies of size R (or mass $M = (4\pi/3)R^3\bar{\rho}_m$, where $\bar{\rho}_m$ is the cosmic mean matter density), the galaxy

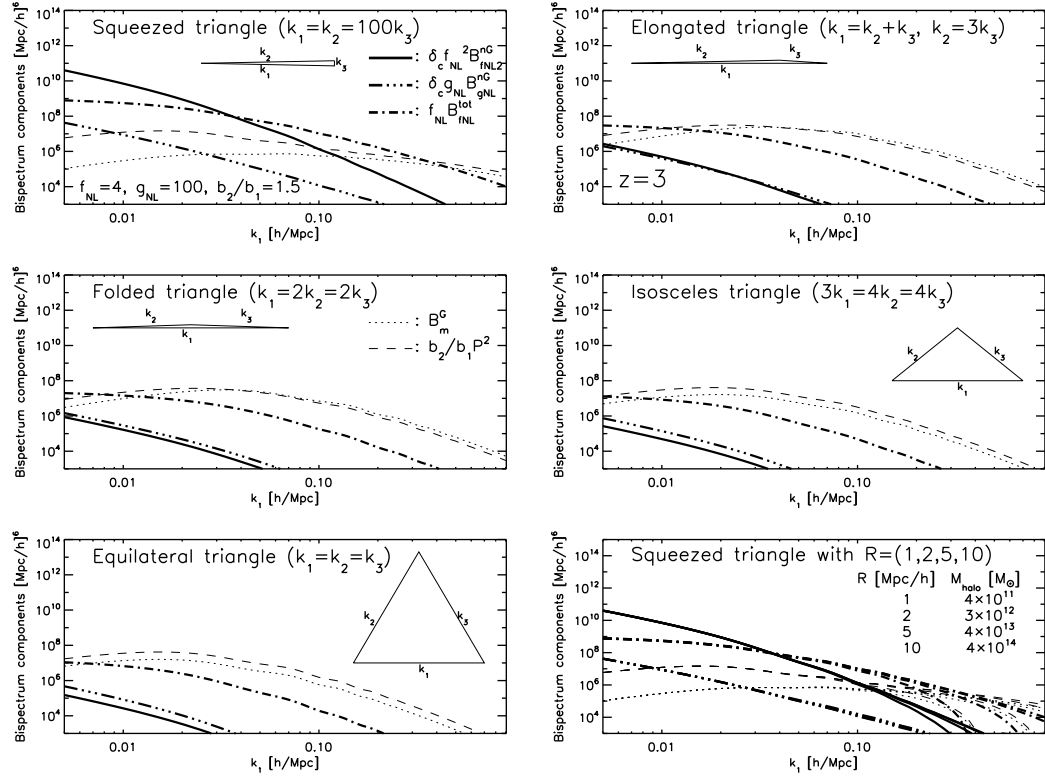


Figure 5.14: Same as Fig. 5.13, but for smaller non-Gaussianity parameters, $f_{\text{NL}} = 4$ and $g_{\text{NL}} = 100$.

bispectrum at redshift z is given by

$$\begin{aligned}
& B_g(k_1, k_2, k_3, z) \\
= & b_1^3(z) D^4(z) \left[B_m^G(k_1, k_2, k_3) + \frac{b_2(z)}{b_1(z)} \{P_R(k_1)P_R(k_2) + (2 \text{ cyclic})\} \right. \\
& + \frac{f_{\text{NL}}}{D(z)} B_{f_{\text{NL}}}^{nG0}(k_1, k_2, k_3) + \frac{\tilde{b}_2(z)}{b_1(z)} \left\{ \frac{f_{\text{NL}}}{D(z)} \left(B_m^{nG}(k_1, k_2, k_3) \right. \right. \\
& + 4(\mathcal{G}_R(k_1) + \mathcal{G}_R(k_2) + \mathcal{G}_R(k_3)) B_{f_{\text{NL}}}^{nG0}(k_1, k_2, k_3) + B_{f_{\text{NL}}}^{nG1}(k_1, k_2, k_3) \Big) \\
& \left. \left. + \frac{f_{\text{NL}}^2}{D^2(z)} B_{f_{\text{NL}}}^{nG}(k_1, k_2, k_3) + \frac{g_{\text{NL}}}{D^2(z)} B_{g_{\text{NL}}}^{nG}(k_1, k_2, k_3) \right\} \right], \tag{5.61}
\end{aligned}$$

where $b_1(z)$ and $b_2(z)$ are the linear and non-linear bias parameters, respectively. As we mentioned in §2, $\tilde{b}_2(z)/b_1(z)$ would be equal to δ_c within the context of the MLB formalism, but the precise value has to be measured from N-body simulations.

Note that the redshift evolution of each term is explicitly given by the powers of the linear growth factor $D(z)$, and various contributions, B_m^G , P_R , B_m^{nG} , $B_{f_{\text{NL}}}^{nG0}$, $B_{f_{\text{NL}}}^{nG1}$, $B_{f_{\text{NL}}}^{nG}$, and $B_{g_{\text{NL}}}^{nG}$, are evaluated at $z = 0$ with

$$\begin{aligned}
B_m^G(k_1, k_2, k_3) &= 2F_2^{(s)}(\mathbf{k}_1, \mathbf{k}_2) \mathcal{W}_R(k_1) \mathcal{W}_R(k_2) \mathcal{W}_R(k_3) P_m(k_1) P_m(k_2) \\
&+ (2 \text{ cyclic}) \tag{5.62}
\end{aligned}$$

$$B_{f_{\text{NL}}}^{nG0}(k_1, k_2, k_3) = 2 \frac{P_R(k_1)}{\mathcal{M}_R(k_1)} \frac{P_R(k_2)}{\mathcal{M}_R(k_2)} \mathcal{M}_R(k_3) + (2 \text{ cyclic}) \tag{5.63}$$

$$\begin{aligned}
B_m^{nG}(k_1, k_2, k_3) &= 4 \mathcal{W}_R(k_1) \mathcal{W}_R(k_2) \mathcal{W}_R(k_3) \left[\frac{\mathcal{F}_R(k_1)}{\mathcal{M}_R(k_1)} + \frac{\mathcal{F}_R(k_2)}{\mathcal{M}_R(k_2)} \right] \\
&\times P_m(k_1) P_m(k_2) F_2^{(s)}(\mathbf{k}_1, \mathbf{k}_2) + (2 \text{ cyclic}) \tag{5.64}
\end{aligned}$$

$$\begin{aligned}
B_{f_{\text{NL}}}^{nG1}(k_1, k_2, k_3) &\approx \frac{1}{2\sigma_R^2} \left[8\mathcal{W}_R(k_2)\mathcal{W}_R(k_3)P_m(k_2)\mathcal{M}(k_3)P_\phi(k_3) \right. \\
&\times \int \frac{d^3q}{(2\pi)^3} \mathcal{W}_R(|\mathbf{k}_1 - \mathbf{q}|)\mathcal{W}_R(q)\mathcal{M}(|\mathbf{k}_1 - \mathbf{q}|)\mathcal{M}(|\mathbf{k}_2 + \mathbf{q}|) \\
&\times [P_\phi(|\mathbf{k}_2 + \mathbf{q}|) + P_\phi(|\mathbf{k}_1 - \mathbf{q}|)] F_2^{(s)}(-\mathbf{k}_2, \mathbf{k}_2 + \mathbf{q}) \\
&\left. + (5 \text{ permutation}) \right] \quad (5.65)
\end{aligned}$$

$$\begin{aligned}
B_{f_{\text{NL}}}^{nG}(k_1, k_2, k_3) &\approx \frac{1}{2\sigma_R^2} \left[8\mathcal{M}_R(k_2)\mathcal{M}_R(k_3)P_\phi(k_1) [P_\phi(k_2) + P_\phi(k_3)] \int \frac{d^3q}{(2\pi)^3} \right. \\
&\times \mathcal{M}_R(q)\mathcal{M}_R(|\mathbf{k}_1 - \mathbf{q}|)P_\phi(q) + (2 \text{ cyclic}) \\
&+ 4\mathcal{M}_R(k_2)\mathcal{M}_R(k_3)P_\phi(k_2)P_\phi(k_3) \int \frac{d^3q}{(2\pi)^3} \mathcal{M}_R(|\mathbf{k}_1 - \mathbf{q}|) \\
&\times \mathcal{M}_R(q) [P_\phi(|\mathbf{k}_2 + \mathbf{q}|) + P_\phi(|\mathbf{k}_3 + \mathbf{q}|)] + (2 \text{ cyclic}) \left. \right] \quad (5.66)
\end{aligned}$$

$$\begin{aligned}
B_{g_{\text{NL}}}^{nG}(k_1, k_2, k_3) &= \frac{1}{2\sigma_R^2} \left[6\mathcal{M}_R(k_2)\mathcal{M}_R(k_3) [P_\phi(k_2) + P_\phi(k_3)] \int \frac{d^3q}{(2\pi)^3} \right. \\
&\times \mathcal{M}_R(q)\mathcal{M}_R(|\mathbf{k}_1 - \mathbf{q}|)P_\phi(q)P_\phi(|\mathbf{k}_1 - \mathbf{q}|) + (2 \text{ cyclic}) \\
&+ 12\mathcal{M}_R(k_2)\mathcal{M}_R(k_3)P_\phi(k_2)P_\phi(k_3) \int \frac{d^3q}{(2\pi)^3} \mathcal{M}_R(q) \\
&\times \mathcal{M}_R(|\mathbf{k}_1 - \mathbf{q}|)P_\phi(q) + (2 \text{ cyclic}) \left. \right]. \quad (5.67)
\end{aligned}$$

Note that we show only dominant terms for $B_{f_{\text{NL}}}^{nG1}$ and $B_{f_{\text{NL}}}^{nG}$ on large scales. One can find the exact definitions in Eq. (5.31) and Eq. (5.36). Finally, $\mathcal{F}_R(k)$ and $\mathcal{G}_R(k)$ are defined as follows.

$$\mathcal{F}_R(k) \equiv \frac{1}{2\sigma_R^2} \int \frac{d^3q}{(2\pi)^3} P_\phi(q)\mathcal{M}_R(q)\mathcal{M}_R(|\mathbf{k} - \mathbf{q}|) \left[\frac{P_\phi(|\mathbf{k} - \mathbf{q}|)}{P_\phi(k)} + 2 \right] \quad (5.68)$$

$$\mathcal{G}_R(k) \equiv \frac{1}{2\sigma_R^2} \int \frac{d^3q}{(2\pi)^3} \frac{\mathcal{W}_R(q)\mathcal{W}_R(|\mathbf{k} - \mathbf{q}|)}{\mathcal{W}_R(k)} P_m(q)F_2^{(s)}(\mathbf{k}, -\mathbf{q}) \quad (5.69)$$

Chapter 6

Galaxy-CMB and galaxy-galaxy lensing on large scales: sensitivity to primordial non-Gaussianity

We can also probe the scale dependent bias generated from Primordial non-Gaussianity by the weak gravitational lensing (see Bartelmann & Schneider (2001) for a review)¹. The scale-dependent bias² was theoretically discovered when the authors of Dalal et al. (2008) studied the form of the *cross-correlation* power spectrum between the dark matter halos and the underlying matter density fluctuations, $P_{hm}(k) = [b_1 + \Delta b(k)]P_m(k)$, where $P_m(k)$ is the power spectrum of matter density fluctuations, and

$$\Delta b(k) = \frac{3(b_1 - 1)f_{\text{NL}}\Omega_m H_0^2 \delta_c}{D(z)k^2 T(k)}. \quad (6.1)$$

Here, $D(z)$ and $T(k)$ are the growth rate and the transfer function for linear matter density fluctuations, respectively, and $\delta_c = 1.68$ is the threshold linear density contrast for a spherical collapse of an overdensity region. The k^2 factor in the denominator of $\Delta b(k)$ shows that this effect is important only on very large scales. Highly biased tracers are more sensitive to f_{NL} .

We can observe $P_{hm}(k)$ by cross-correlating the locations of galaxies or clusters of galaxies with the matter density fluctuations traced by the weak gravitational lensing. In this chapter, we study the modification of two-point cross-correlation statistics of weak lensing - galaxy-galaxy lensing and galaxy-Cosmic Microwave Background (CMB) cross-correlation - due to f_{NL} ³.

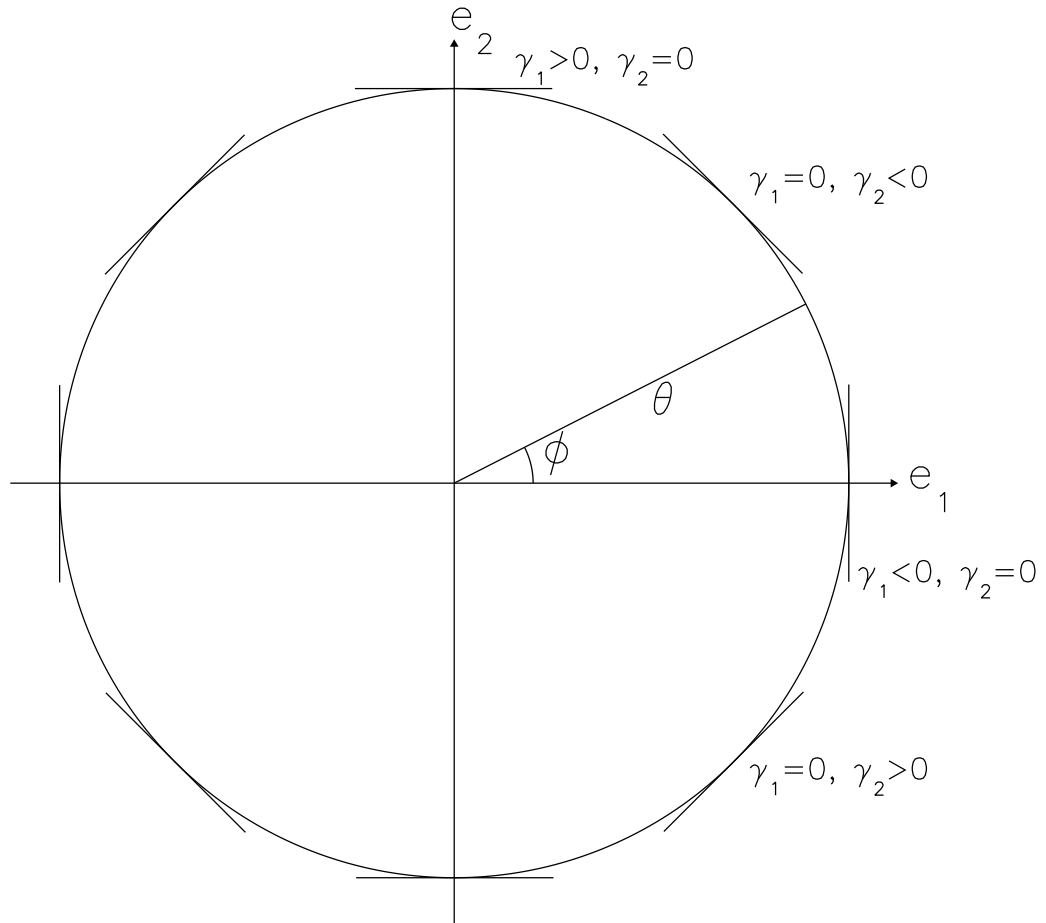


Figure 6.1: Coordinate system and γ_1 and γ_2 . The shear along e_1 has $\gamma_1 > 0$ and $\gamma_2 = 0$, whereas the shear along e_2 has $\gamma_1 < 0$ and $\gamma_2 = 0$. The shear along $e_1 + e_2$ has $\gamma_1 = 0$ and $\gamma_2 > 0$, whereas The shear along $e_1 - e_2$ has $\gamma_1 = 0$ and $\gamma_2 < 0$.

6.1 Halo-mass correlation from galaxy-galaxy lensing

6.1.1 Formula

One efficient way of measuring $P_{hm}(k)$ is to use the so-called galaxy-galaxy lensing technique Tyson et al. (1984); Brainerd et al. (1996); McKay et al. (2001); Guzik & Seljak (2002); Sheldon et al. (2004); Mandelbaum et al. (2006b,a): choose one lens galaxy at a redshift z_L , and measure the mean of *tangential* shears in images of lensed (source or background) galaxies around the chosen central lensing galaxy as a function of radii from that central galaxy. Finally, average those mean tangential shears over all lensing galaxies at the same redshift, z_L .

We begin with the definition of the tangential shear, γ_t , on the flat sky⁴

$$\gamma_t(\boldsymbol{\theta}) = -\gamma_1(\boldsymbol{\theta}) \cos(2\phi) - \gamma_2(\boldsymbol{\theta}) \sin(2\phi), \quad (6.2)$$

where $\boldsymbol{\theta} = (\theta \cos \phi, \theta \sin \phi)$, and γ_1 and γ_2 are components of the shear field.⁵ The coordinate system and the meaning of γ_1 and γ_2 are explained in Fig. 6.1. For purely tangential shears shown in Fig. 6.1, γ_t is always positive. This property allows us to average γ_t over the ring around the origin to estimate the mean tangential shear, $\bar{\gamma}_t$:

$$\bar{\gamma}_t(\theta) \equiv \int_0^{2\pi} \frac{d\phi}{2\pi} \gamma_t(\theta, \phi). \quad (6.3)$$

On the flat sky, γ_1 and γ_2 are related to the projected mass density fluctuation in Fourier space, $\kappa(\mathbf{l})$, as

$$\gamma_1(\boldsymbol{\theta}) = \int \frac{d^2\mathbf{l}}{(2\pi)^2} \kappa(\mathbf{l}) \cos(2\varphi) e^{i\mathbf{l} \cdot \boldsymbol{\theta}}, \quad (6.4)$$

$$\gamma_2(\boldsymbol{\theta}) = \int \frac{d^2\mathbf{l}}{(2\pi)^2} \kappa(\mathbf{l}) \sin(2\varphi) e^{i\mathbf{l} \cdot \boldsymbol{\theta}}, \quad (6.5)$$

¹Previous version of this chapter was published in Jeong, D., Komatsu, E. & Jain, B. 2009, Phys. Rev. D, 80, 123527.

²For the derivation of the scale dependent bias, see Appendix I.3.

³A significant fraction of this chapter has been published in Jeong, D., Komatsu, E., & Jain B. 2009 Phys. Rev. D, 80, 123527.

⁴For an all-sky analysis, this relation needs to be replaced with the exact relation using the spin-2 harmonics Stebbins (1996).

⁵As the shear has two independent components, we are ignoring another linear combination of γ_1 and γ_2 by only focusing on the tangential shear. In particular, on large scales there is information in the other component of the shear, and thus the full analysis including both shear components (not just tangential one) yields a modest (smaller than a factor of $\sqrt{2}$) improvement in the signal-to-noise ratio. Moreover, using magnification (in addition to shears), which is proportional to the convergence field κ , can also yield a modest improvement.

where φ is the angle between \mathbf{l} and \mathbf{e}_1 , i.e., $\mathbf{l} = (l \cos \varphi, l \sin \varphi)$. Using Eqs. (6.4) and (6.5) in Eq. (6.2), we write the tangential shear in terms of $\kappa(\mathbf{l})$ as

$$\gamma_t(\boldsymbol{\theta}) = - \int \frac{d^2 \mathbf{l}}{(2\pi)^2} \kappa(\mathbf{l}) \cos[2(\phi - \varphi)] e^{il\theta \cos(\phi - \varphi)}. \quad (6.6)$$

The mean tangential shear (Eq. (6.3)) is then given by

$$\begin{aligned} \bar{\gamma}_t(\theta) &= - \int \frac{d^2 \mathbf{l}}{(2\pi)^2} \kappa(\mathbf{l}) \int_0^{2\pi} \frac{d\phi}{2\pi} \cos[2(\phi - \varphi)] e^{il\theta \cos(\phi - \varphi)} \\ &= \int \frac{d^2 \mathbf{l}}{(2\pi)^2} \kappa(\mathbf{l}) J_2(l\theta). \end{aligned} \quad (6.7)$$

Here, we have used the identity

$$J_m(x) = \int_{\alpha}^{2\pi+\alpha} \frac{d\psi}{2\pi} e^{i(m\psi - x \sin \psi)}, \quad (6.8)$$

with $m = 2$, $\psi = \phi - \varphi - \pi/2$, $\alpha = \varphi + \pi/2$, and $\int_0^{2\pi} d\psi \sin(2\psi) e^{ix \cos \psi} = 0$.

The ensemble average of the mean tangential shear vanishes, i.e., $\langle \bar{\gamma}_t \rangle = 0$, as $\langle \kappa \rangle = 0$. This simply means that the average of the mean tangential shears, measured with respect to random points on the sky, vanishes. We obtain non-zero values when we average the mean tangential shears measured with respect to the locations of halos (galaxies or clusters of galaxies). This quantity, called the galaxy-galaxy lensing or cluster-galaxy lensing, can be used to measure the halo-mass cross correlation.

While clusters of galaxies may be identified directly with dark matter halos of a given mass, how are galaxies related to halos? Some galaxies (“field galaxies”) may also be identified directly with dark matter halos; however, galaxies residing within groups or clusters of galaxies should be identified with subhalos moving in a bigger dark matter halo. For such subhalos our argument given below may not be immediately used. However, it is observationally feasible to identify the central galaxies in groups or clusters of galaxies and measure the mean tangential shear around them. A number of studies of Luminous Red Galaxies (LRGs) extracted from the Sloan Digital Sky Survey (SDSS) have shown that these are typical central galaxies in galaxy groups Mandelbaum et al. (2006a); Sheldon et al. (2007); Johnston et al. (2007). Scalings such as the mass-luminosity scaling imply that LRGs provide a useful proxy for the halos within which they reside. We will assume in this study that such tracers will enable the halo-shear cross-correlation to be measured. There are some caveats such as bimodal mass distributions in galaxy groups Johnston et al.

(2007) and the extrapolation to higher redshift, but we will leave a detailed exploration to real galaxy tracers for later work.

The ensemble average of the mean tangential shears relative to the locations of halos at a given redshift z_L , denoted as $\langle \bar{\gamma}_t^h \rangle(\theta, z_L)$, is related to the angular cross-correlation power spectrum of halos and κ , $C_l^{h\kappa}$, as Hu & Jain (2004)

$$\langle \bar{\gamma}_t^h \rangle(\theta, z_L) = \int \frac{l dl}{2\pi} C_l^{h\kappa}(z_L) J_2(l\theta). \quad (6.9)$$

We give the derivation of this result in Appendix L.1.

With the lens redshift z_L known (from spectroscopic observations), we can calculate the comoving radius, R , corresponding to the angular separation on the sky, θ , as $R = \theta d_A(0; z_L)$ where $d_A(0; z_L)$ is the *comoving* angular diameter distance from $z = 0$ to $z = z_L$. Using Limber's approximation (Limber, 1954; Kaiser, 1992) on the flat sky relating $C_l^{h\kappa}$ to $P_{hm}(k)$,⁶ we can write Eq. (6.9) as Hu & Jain (2004)

$$\langle \bar{\gamma}_t^h \rangle(R, z_L) = \frac{\rho_0}{\Sigma_c(z_L)} \int \frac{k dk}{2\pi} P_{hm}(k, z_L) J_2(kR). \quad (6.10)$$

Here, ρ_0 is the mean comoving mass density of the universe, and $\Sigma_c(z_L)$ is the so-called critical surface density:

$$\Sigma_c^{-1}(z_L) = \frac{4\pi G}{c^2} (1 + z_L) d_A(0; z_L) \int_{z_L}^{\infty} dz_S p(z_S) \frac{d_A(z_L; z_S)}{d_A(0; z_S)}, \quad (6.11)$$

where $p(z_S)$ is the redshift distribution of sources normalized to unity, $\int dz p(z) = 1$, and $d_A(0; z)$ and $d_A(z; z_S)$ are the comoving angular diameter distances out to z and between z and z_S , respectively. The numerical value of $4\pi G/c^2$ is 6.01×10^{-19} Mpc/M $_{\odot}$, and $4\pi G\rho_0/c^2$ is $1.67 \times 10^{-7}(\Omega_m h^2)$ Mpc $^{-2}$.

Eq. (6.10) is often written as

$$\langle \bar{\gamma}_t^h \rangle(R, z_L) = \frac{\Delta \Sigma(R, z_L)}{\Sigma_c(z_L)}. \quad (6.12)$$

To simplify the analysis, let us define the “effective source redshift” of a given survey from the following equation:

$$\frac{d_A(z_L; z_{S,\text{eff}})}{d_A(0; z_{S,\text{eff}})} \equiv \int_{z_L}^{\infty} dz_S p(z_S) \frac{d_A(z_L; z_S)}{d_A(0; z_S)}. \quad (6.13)$$

⁶As we are dealing with correlations on very large angular scales, one may worry about the validity of Limber's approximation. In Appendix M we give a detailed study of the validity and limitation of Limber's approximation for the galaxy-galaxy lensing.

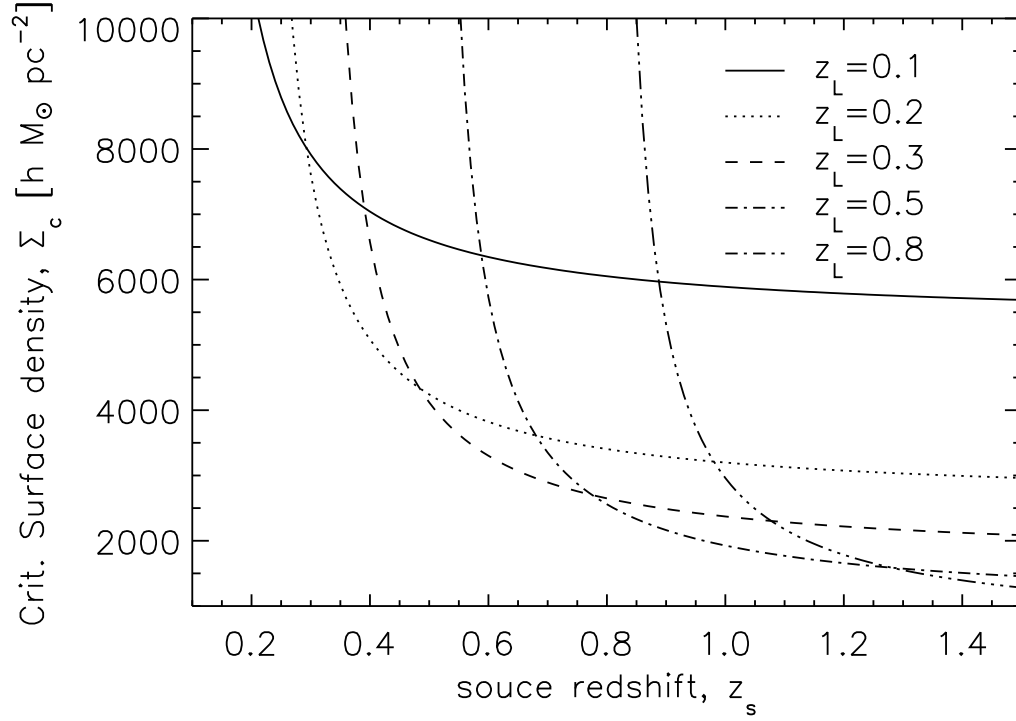


Figure 6.2: Critical surface density, $\Sigma_c(z_L; z_S)$, as a function of the source redshift, z_S , for various lens redshifts that roughly correspond to the Two Degree Field Galaxy Redshift Survey (2dFGRS; $z_L = 0.1$, solid), the main sample of the Sloan Digital Sky Survey (SDSS; $z_L = 0.2$, dotted), the Luminous Red Galaxies (LRGs) of SDSS ($z_L = 0.3$, dashed), and the Large Synoptic Survey Telescope (LSST; $z_L = 0.5$ and 0.8 , dot-dashed and triple-dot-dashed, respectively).

Henceforth we shall use z_S to denote $z_{S,\text{eff}}$, and write

$$\Sigma_c^{-1}(z_L; z_S) = \frac{4\pi G}{c^2} (1 + z_L) d_A(0; z_L) \frac{d_A(z_L; z_S)}{d_A(0; z_S)}. \quad (6.14)$$

Fig. 6.2 shows Σ_c for $z_L = 0.1$ (2dFGRS, Two Degree Field Galaxy Redshift Survey), 0.2 (SDSS main), 0.3 (SDSS LRG), and 0.5 and 0.8 (both LSST, Large Synoptic Survey Telescope). The smaller Σ_c is, the larger the observed mean tangential shear is.

6.1.2 Results

We can now calculate the observable, $\Delta\Sigma(R, z_L)$, for various values of f_{NL} . We use

$$\begin{aligned} & \Delta\Sigma(R, z_L) \\ &= \rho_0 b_1 \int \frac{k dk}{2\pi} P_m(k, z_L) J_2(kR) \\ & \quad + \rho_0 \int \frac{k dk}{2\pi} \Delta b(k, z_L) P_m(k, z_L) J_2(kR), \end{aligned} \quad (6.15)$$

where the scale-dependent bias, $\Delta b(k, z)$, is given by Eq. (6.1). As we are interested in large scales, i.e., $R > 10 h^{-1}$ Mpc, we shall use the linear matter spectrum for $P_m(k)$.

Fig. 6.3 shows, for the Gaussian initial condition ($f_{\text{NL}} = 0$), $\Delta\Sigma(R, z_L)$ from $R = 50$ to $200 h^{-1}$ Mpc. We have chosen the bias parameters and lens redshifts to represent the existing data sets as well as the future ones: $b_1 = 2$ at $z_L = 0.3$ (similar to the observed values from SDSS LRGs (Tegmark et al., 2006), top-left), $b_1 = 2$ at $z_L = 0.5$ (higher- z LRGs (Schlegel et al., 2009), top-right), $b_1 = 2$ at $z_L = 0.8$ (galaxies that can be observed by LSST, (Zhan, 2006), bottom-left), and $b_1 = 5$ at $z_L = 0.8$ (clusters of galaxies that can be observed by LSST, bottom-right). While LSST is an imaging survey, we assume that we can obtain spectroscopic redshifts of some ($\sim 10^6$) lens galaxies by follow-up observations. It is also straightforward to extend our analysis to lenses selected by photometric redshifts.

At $R \sim 110 h^{-1}$ Mpc we see a clear “shoulder” due to the baryonic feature in the linear matter power spectrum (often called Baryon Acoustic Oscillations; BAO). The sound horizon at the drag epoch (which is more relevant to the matter power spectrum than the photon decoupling epoch for the CMB power spectrum) calculated from the cosmological model that we use, the “WMAP+BAO+SN ML” parameters in Table 1 of Komatsu et al. (2009), is $106.9 h^{-1}$ Mpc, as shown as the vertical line in this figure. The magnitude of $\Delta\Sigma$ on this scale is $\sim 0.1 h M_\odot \text{ pc}^{-2}$. Assuming a range of Σ_c from future surveys, $\Sigma_c \sim 1000 - 4000 h M_\odot \text{ pc}^{-2}$ (see Fig. 6.2), this value corresponds to the mean tangential shear of order 2.5×10^{-5} to 10^{-4} . Is this observable?

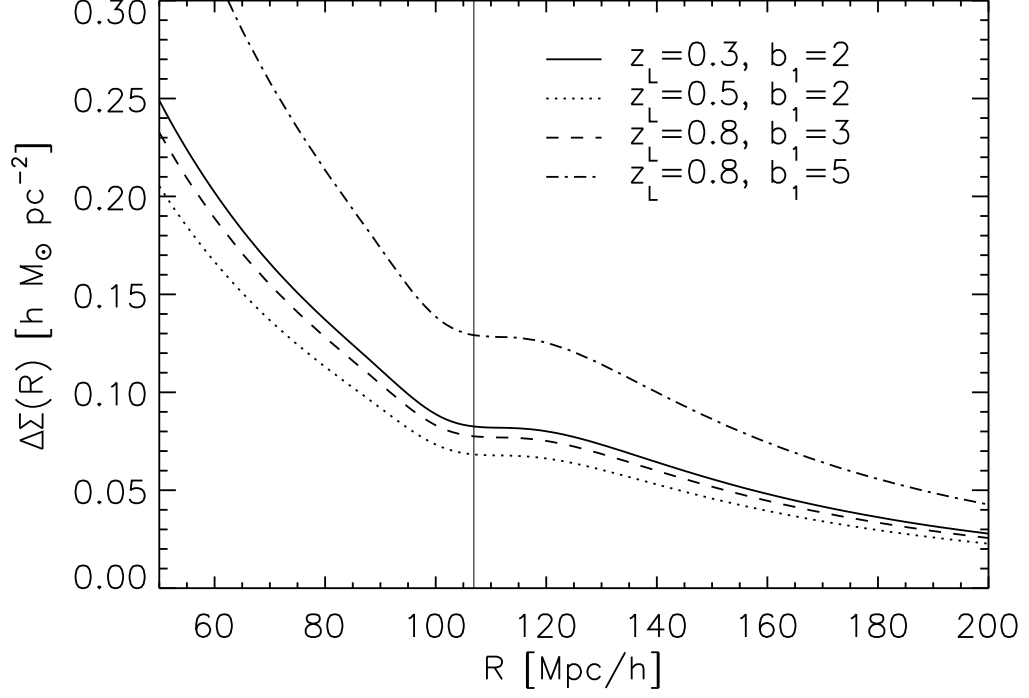


Figure 6.3: The baryonic feature in the matter power spectrum, as seen in the galaxy-galaxy lensing, $\Delta\Sigma(R)$, for several populations of lens galaxies with $b_1 = 2$ at $z_L = 0.3$ (similar to SDSS LRGs, solid), $b_1 = 2$ at $z_L = 0.5$ (higher- z LRGs, dotted), $b_1 = 2$ at $z_L = 0.8$ (galaxies that can be observed by LSST, dashed), and $b_1 = 5$ at $z_L = 0.8$ (clusters of galaxies that can be observed by LSST, dot-dashed). The vertical line shows the location of the baryonic feature, $R_{\text{BAO}} = 106.9 h^{-1} \text{ Mpc}$, calculated from the “WMAP+BAO+SN ML” parameters in Table 1 of Komatsu et al. (2009). Note that we have used the linear matter power spectrum and the Gaussian initial condition ($f_{\text{NL}} = 0$) for this calculation.

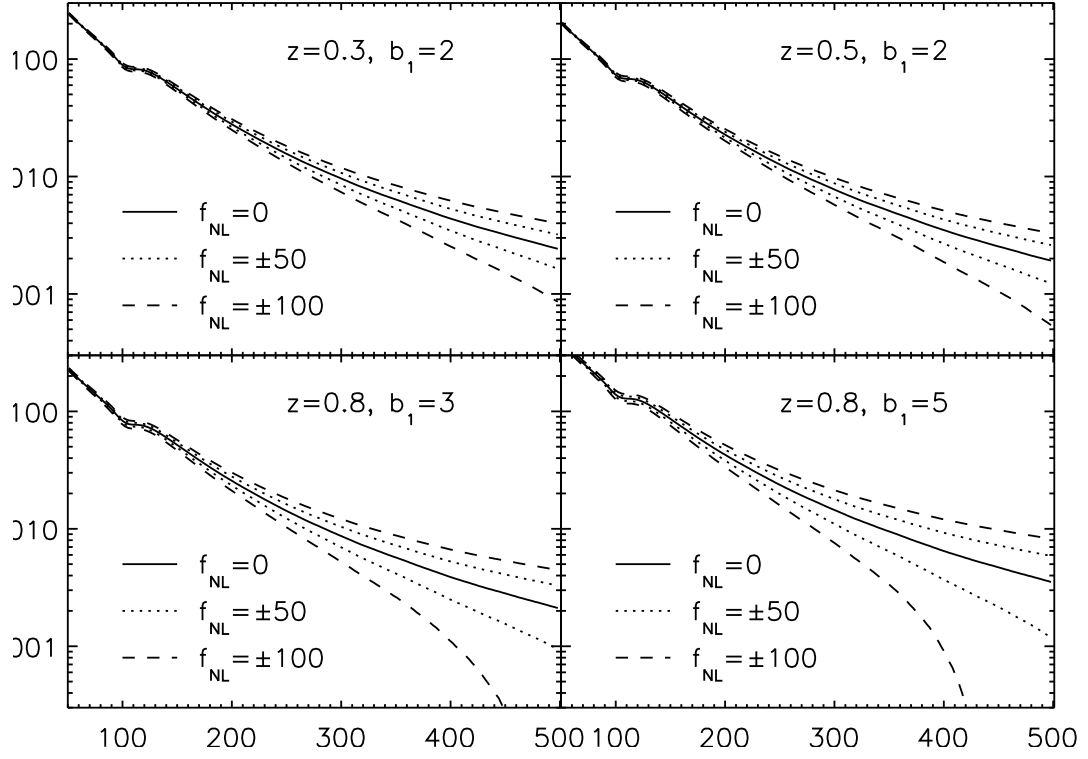


Figure 6.4: Imprints of the local-type primordial non-Gaussianity in the galaxy-galaxy lensing, $\Delta\Sigma(R)$, for the same populations of lens galaxies as in Fig. 6.3. The solid, dashed, and dotted lines show $f_{\text{NL}} = 0, \pm 50$, and ± 100 , respectively.

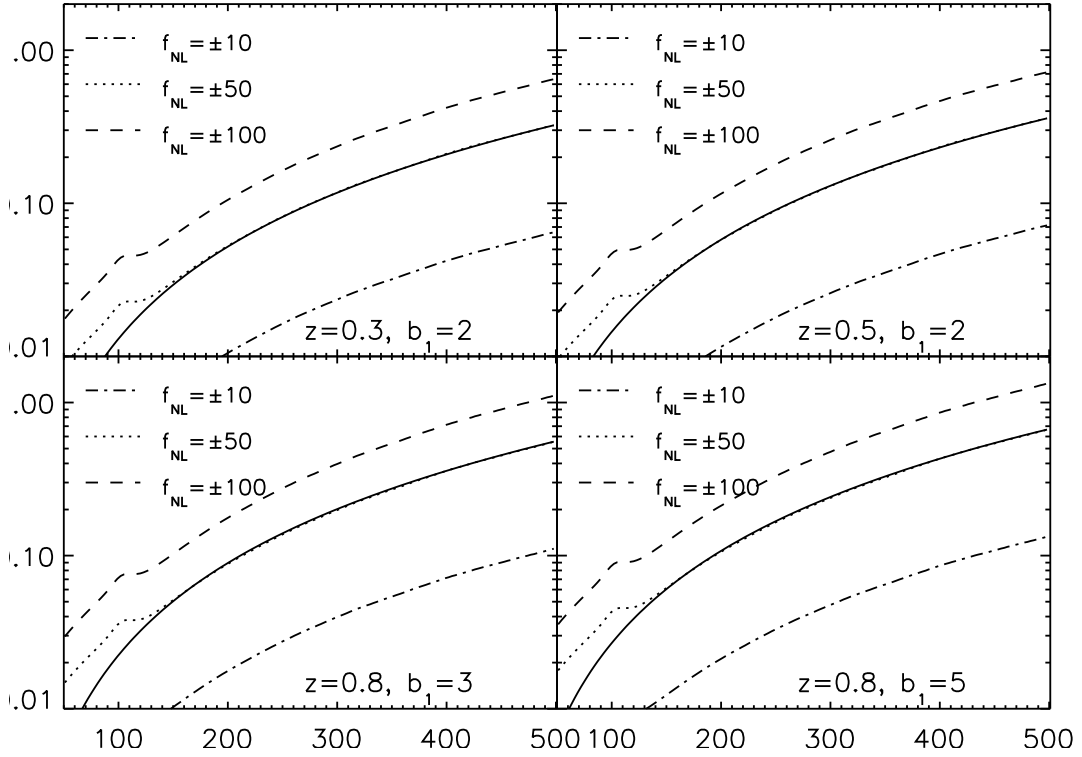


Figure 6.5: Fractional differences between $\Delta\Sigma(R)$ from non-Gaussian initial conditions and the Gaussian initial condition, $|\Delta\Sigma(R; f_{\text{NL}})/\Delta\Sigma(R; f_{\text{NL}}=0) - 1|$, calculated from the curves shown in Fig. 6.4. The dot-dashed, dashed, and dotted lines show $f_{\text{NL}} = \pm 10$, ± 50 , and ± 100 , respectively, while the thin solid line shows $\propto R^2$ with an arbitrary normalization.

For comparison, Sheldon et al. (2007) measured $\Delta\Sigma(R) \sim 0.5 h M_\odot \text{pc}^{-2}$ at $R \sim 30 h^{-1} \text{Mpc}$ from clusters of galaxies in the SDSS main sample. The mean lens redshift for these data is $z_L \sim 0.2$, which would give $\Sigma_c \sim 5000 h M_\odot \text{pc}^{-2}$ (see Fig. 6.2 for $z_L = 0.2$ and $z_S \sim 0.4$); thus, the magnitude of the mean tangential shear that they were able to measure is of order 10^{-4} , which is only ~ 1 to 4 times larger than the magnitude of the signal expected from the BAO. Therefore, detecting the BAO signature in $\Delta\Sigma(R)$ should be quite feasible with the future observations. We shall give a more quantitative discussion on the detectability of BAO from the galaxy-galaxy lensing effect in Sec 6.1.4.

How about f_{NL} ? As expected, the effect of f_{NL} is enhanced on very large scales, i.e., hundreds of Mpc (see Fig. 6.4). For $f_{\text{NL}} = \pm 50$, $\Delta\Sigma(R)$ is modified by 10–20% at $R \sim 300 h^{-1} \text{Mpc}$ (depending on b_1 and z_L ; see Fig. 6.5). The modification grows rapidly toward larger scales, in proportion to R^2 . On such a large scale ($R \sim 300 h^{-1} \text{Mpc}$), the galaxy-galaxy lensing signal is on the order of $\Delta\Sigma \sim 0.01 h M_\odot \text{pc}^{-2}$, and thus we need to measure the mean tangential shear down to the level of $\bar{\gamma}_t^h \sim 2.5 \times 10^{-6}$ to 10^{-5} , i.e., 10–40 times smaller than the level of sensitivity achieved by the current observations. Can we observe such a small shear?

6.1.3 Covariance matrix of the mean tangential shear

In order to study the feasibility of measuring the tangential shear of order 10^{-6} , we compute the covariance matrix of the mean tangential shears averaged over N_L lens galaxies. As derived in Appendix L.2, the covariance matrix of the mean tangential shear is

$$\begin{aligned} & \langle \bar{\gamma}_t^h(\theta) \bar{\gamma}_t^h(\theta') \rangle - \langle \bar{\gamma}_t^h(\theta) \rangle \langle \bar{\gamma}_t^h(\theta') \rangle \\ &= \frac{1}{4\pi f_{\text{sky}}} \int \frac{ldl}{2\pi} J_2(l\theta) J_2(l\theta') \\ & \quad \times \left[(C_l^{h\kappa})^2 + \left(C_l^h + \frac{1}{n_L} \right) \left(C_l^\kappa + \frac{\sigma_\gamma^2}{n_S} \right) \right]. \end{aligned} \quad (6.16)$$

This expression includes the cosmic variance, the shot noise of lens halos, as well as the shape noise σ_γ . As far as we know this formula has not been derived before. Note that we have assumed a single source and lens redshift. For multiple source and lens redshifts, the covariance matrix needs to be suitably generalized.

Here, C_l^h and C_l^κ are the angular power spectra of the lens halos (galaxies or cluster of galaxies) and κ , respectively, and n_L and n_S are the number densities of the lens halos and the lensed (source) galaxies, respectively. These angular power spectra, $C_l^{h\kappa}$, C_l^h , C_l^κ , will be related to the corresponding three-dimensional power spectrum, $P(k)$, in Sec 6.2.3.

In the limit that the cosmic variance is unimportant, we recover the usual expression used in the literature:

$$\langle \bar{\gamma}_t^h(\theta) \bar{\gamma}_t^h(\theta') \rangle - \langle \bar{\gamma}_t^h(\theta) \rangle \langle \bar{\gamma}_t^h(\theta') \rangle = \frac{\sigma_\gamma^2}{N_L} \frac{\delta_D(\theta - \theta')}{2\pi\theta n_S}, \quad (6.17)$$

where $N_L = 4\pi f_{\text{sky}} n_L$ is the total number of lens halos available in the data. In this limit the errors in different radial bins are uncorrelated, and they are simply given by the shape noise, σ_γ , reduced by the square-root of the number of source galaxies available within each radial bin and the total number of lens halos that we can use for averaging the mean tangential shear. In particular, at each bin with a width $\Delta\theta$, we find the variance of

$$\text{Var}[\bar{\gamma}_t^h(\theta)] = \frac{\sigma_\gamma^2}{2\pi\theta(\Delta\theta)n_S N_L}, \quad (6.18)$$

in the absence of the cosmic variance.

When would the cosmic variance become important? There is the maximum surface number density of sources, $n_{S,\text{max}} = \sigma_\gamma^2/C_l^\kappa$, above which the shape noise becomes irrelevant. This gives the maximum number of sources within a given radial bin of a width $\Delta\theta$ ($\ll \theta$) above which the shape noise becomes irrelevant:

$$N_{S,\text{max}} = 2\pi\theta(\Delta\theta)n_{S,\text{max}} = (l\theta)^2 \left(\frac{\Delta\theta}{\theta} \right) \frac{\sigma_\gamma^2}{l^2 C_l^\kappa / (2\pi)}. \quad (6.19)$$

For $l\theta = \pi$ (the usual relation between l and θ) and $\sigma_\gamma \simeq 0.3$ (realistic shape noise), we find

$$N_{S,\text{max}} \simeq \left(\frac{\Delta\theta}{\theta} \right) \frac{1}{l^2 C_l^\kappa / (2\pi)}. \quad (6.20)$$

At $l \sim 100$, $l^2 C_l^\kappa / (2\pi) \sim 10^{-5}$ Hu & Jain (2004); thus, we do not gain sensitivity any further by having more than, say, 10^4 galaxies (for $\Delta\theta/\theta = 0.1$) within a single radial bin.

Alternatively, one can define the minimum multipole, l_{min} , below which the cosmic variance term dominates:

$$l_{\text{min}} = \sqrt{\frac{2\pi n_S l^2 C_l^\kappa}{\sigma_\gamma^2} \frac{1}{2\pi}}. \quad (6.21)$$

For LSST, we expect to have the surface density of sources on the order of $n_S = 30 \text{ arcmin}^{-2} = 3.5 \times 10^8 \text{ sr}^{-1}$. For $\sigma_\gamma = 0.3$, we find $l_{\text{min}}(\text{LSST}) \sim 1.6 \times 10^5 \sqrt{l^2 C_l^\kappa / (2\pi)}$. At $l \lesssim 10^3$, $l^2 C_l^\kappa / (2\pi) \lesssim 10^{-4}$ Hu & Jain (2004); thus, at $l \lesssim 10^3$ the cosmic variance term dominates.

In the limit that the covariance matrix is dominated by the cosmic variance terms, we have

$$\begin{aligned} & \langle \bar{\gamma}_t^h(\theta) \bar{\gamma}_t^h(\theta') \rangle - \langle \bar{\gamma}_t^h(\theta) \rangle \langle \bar{\gamma}_t^h(\theta') \rangle \\ &= \frac{1}{4\pi f_{\text{sky}}} \int \frac{ldl}{2\pi} J_2(l\theta) J_2(l\theta') C_l^h C_l^\kappa (1 + r_l^2), \end{aligned}$$

where $r_l \equiv C_l^{h\kappa} / \sqrt{C_l^\kappa C_l^h}$ is the cross-correlation coefficient. The variance at a given radial bin is

$$\begin{aligned} & \text{Var}[\bar{\gamma}_t^h(\theta)] \\ &= \frac{1}{4\pi f_{\text{sky}}} \int \frac{ldl}{2\pi} [J_2(l\theta)]^2 C_l^h C_l^\kappa (1 + r_l^2). \end{aligned} \quad (6.22)$$

6.1.4 Detectability of the mean tangential shear

In this section, we shall calculate the expected uncertainties in radially binned measurements of the mean tangential shear.

The mean tangential shear averaged within the i -th bin, $\langle \hat{\gamma}_t^h \rangle(\theta_i)$, i.e., the mean tangential shear averaged within an annulus between $\theta_{i,\min}$ and $\theta_{i,\max}$, is given by

$$\begin{aligned} \langle \hat{\gamma}_t^h \rangle(\theta_i) &= \frac{2\pi}{A(\theta_i)} \int_{\theta_{i,\min}}^{\theta_{i,\max}} \theta d\theta \langle \bar{\gamma}_t^h \rangle(\theta) \\ &\equiv \int \frac{ldl}{2\pi} C_l^{h\kappa} \hat{J}_2(l\theta_i), \end{aligned} \quad (6.23)$$

where $A(\theta_i) = \pi(\theta_{i,\max}^2 - \theta_{i,\min}^2)$ is the area of the annulus, and

$$\hat{J}_2(l\theta_i) = \frac{2\pi}{A(\theta_i)} \int_{\theta_{i,\min}}^{\theta_{i,\max}} \theta d\theta J_2(l\theta), \quad (6.24)$$

is the Bessel function averaged within a bin.

Similarly, the covariance matrix of the binned mean tangential shears is given by

$$\begin{aligned} C_{ij} &\equiv \langle \hat{\gamma}_t^h(\theta_i) \hat{\gamma}_t^h(\theta_j) \rangle - \langle \hat{\gamma}_t^h(\theta_i) \rangle \langle \hat{\gamma}_t^h(\theta_j) \rangle \\ &= \frac{1}{4\pi f_{\text{sky}}} \int \frac{ldl}{2\pi} \hat{J}_2(l\theta_i) \hat{J}_2(l\theta_j) \\ &\quad \times \left[(C_l^{h\kappa})^2 + \left(C_l^h + \frac{1}{n_L} \right) \left(C_l^\kappa + \frac{\sigma_\gamma^2}{n_S} \right) \right]. \end{aligned} \quad (6.25)$$

This matrix contains the full information regarding the statistical errors of the binned measurements of the mean tangential shear, which includes the cosmic variance errors due to the cosmic shear (C_l^κ), clustering of lens galaxies (C_l^h) and their correlations ($C_l^{h\kappa}$), the finite number density of lenses, and the noise in intrinsic shapes of source galaxies.

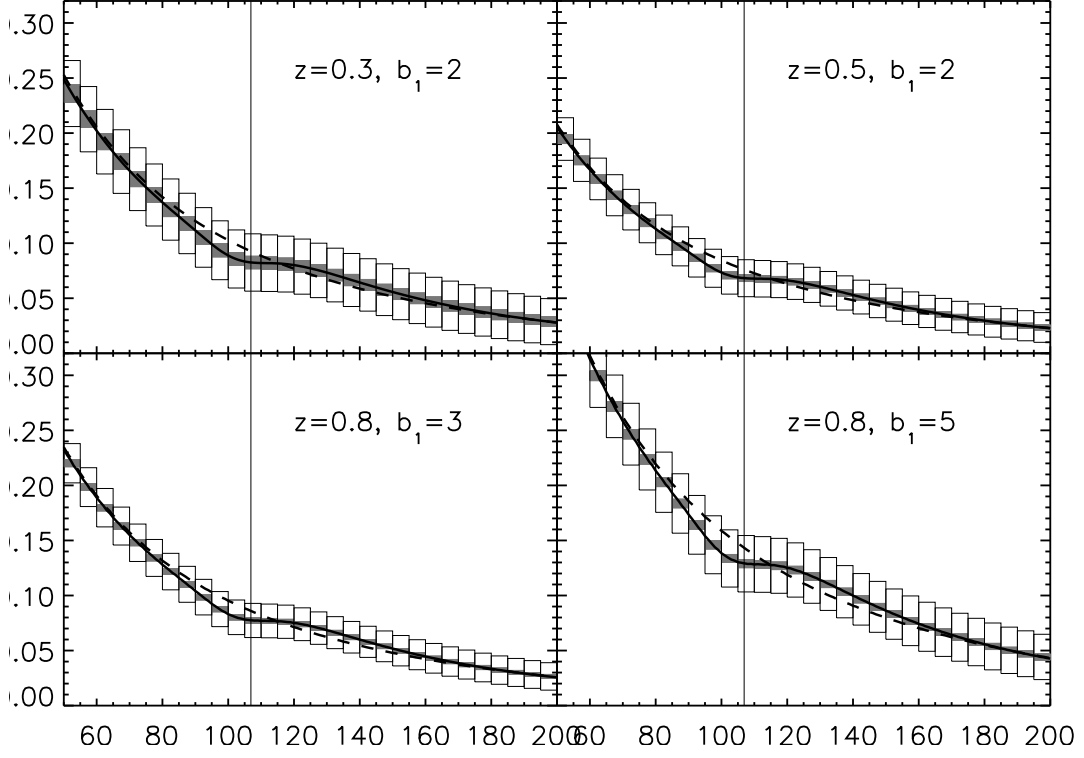


Figure 6.6: Same as Fig. 6.3, but with the expected $1\text{-}\sigma$ uncertainties for full-sky lens surveys and a single lens redshift. Adjacent bins are highly correlated, with the correlation coefficients shown in Fig. 6.7. The open (filled) boxes show the binned uncertainties with (without) the cosmic variance term due to the cosmic shear field included. See Eq. (6.26) and (6.27) for the formulae giving open and filled boxes, respectively. We use the radial bin of size $\Delta R = 5 h^{-1}$ Mpc. For comparison, we also show $\Delta\Sigma(R)$ computed from the smooth power spectrum without the baryonic feature (Eisenstein & Hu, 1998) (dashed lines). Note that the uncertainties are calculated for a single lens redshift slice, and thus they will go down as we add more lens redshift slices.

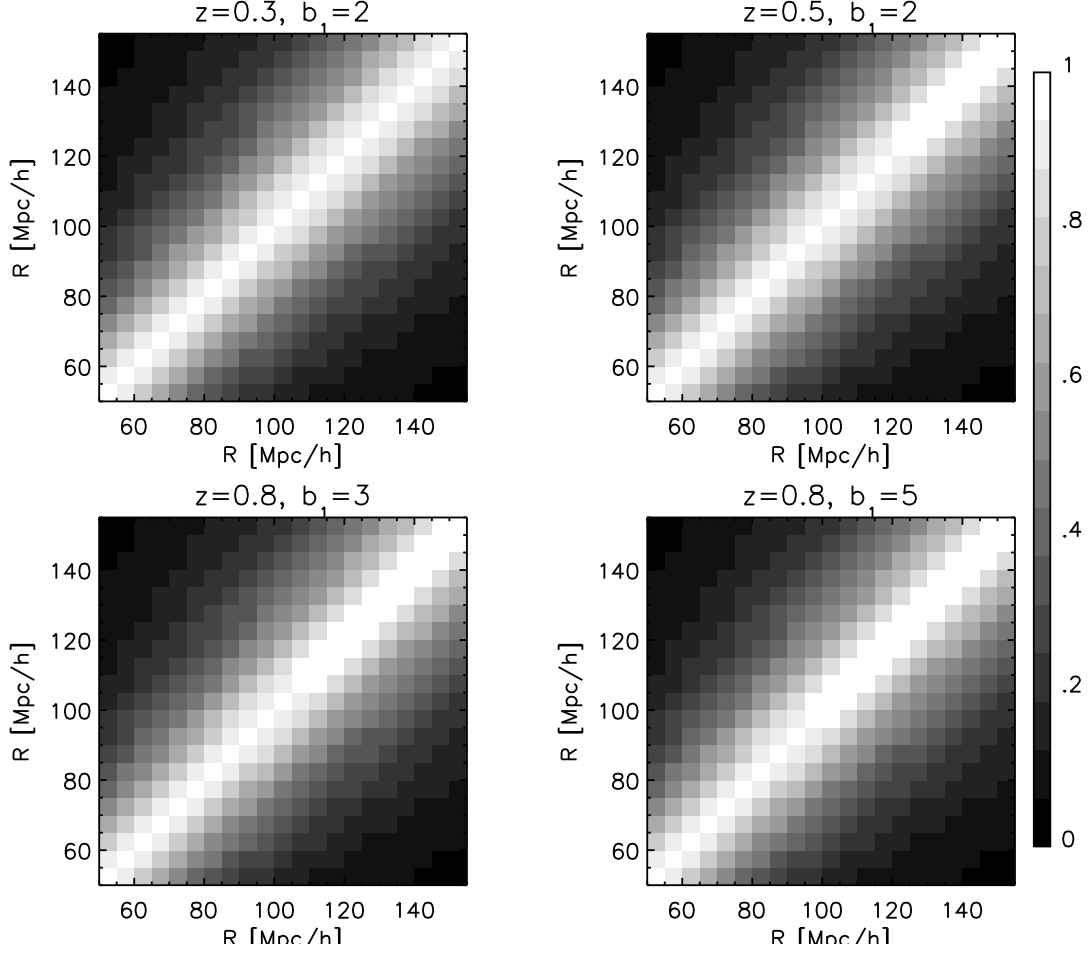


Figure 6.7: The cross-correlation-coefficient matrix, $r_{ij} \equiv C_{ij} / \sqrt{C_{ii}C_{jj}}$, where C_{ij} is the covariance matrix given in Eq. (6.25), for a radial bin of $\Delta R = 5 h^{-1}$ Mpc. We show r_{ij} for the same populations of lens galaxies as shown in Fig. 6.3 and 6.6. We use the same number of source galaxies and the same shape noise as in Fig. 6.6. The neighboring bins are highly correlated for $\Delta R < 10 h^{-1}$ Mpc.

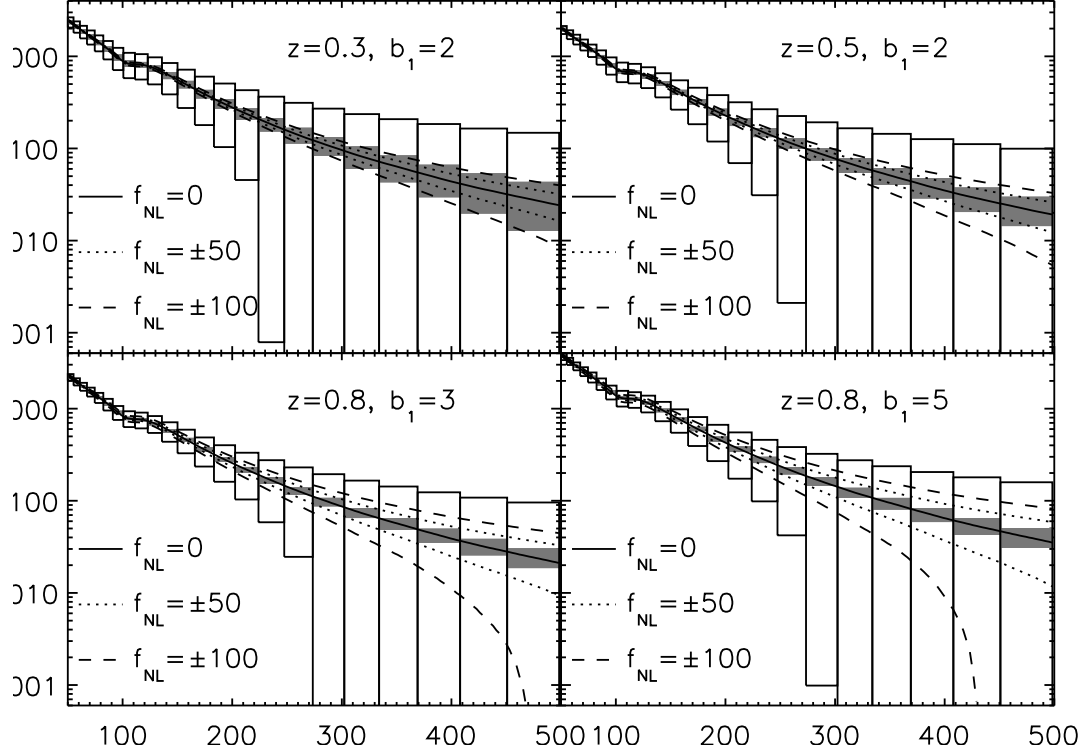


Figure 6.8: Same as Fig. 6.4, but with the expected $1\text{-}\sigma$ uncertainties for full-sky lens surveys and a single lens redshift. Adjacent bins are highly correlated. The open (filled) boxes show the binned uncertainties with (without) the cosmic variance term due to the cosmic shear field included. See Eq. (6.26) and (6.27) for the formulae giving open and filled boxes, respectively. We use logarithmic bins with $\Delta R = R/10$. Note that the uncertainties are calculated for a single lens redshift slice, and thus they will go down as we add more lens redshift slices.

The variance at a given radial bin is

$$\begin{aligned} \text{Var}[\hat{\gamma}_t^h(\theta_i)] &= \frac{1}{4\pi f_{\text{sky}}} \int \frac{ldl}{2\pi} [\hat{J}_2(l\theta_i)]^2 \\ &\times \left[(C_l^{h\kappa})^2 + \left(C_l^h + \frac{1}{n_L} \right) \left(C_l^\kappa + \frac{\sigma_\gamma^2}{n_S} \right) \right]. \end{aligned} \quad (6.26)$$

In the analysis of the galaxy-galaxy lensing effects in the literature, the cosmic variance due to cosmic shear is usually ignored:

$$\begin{aligned} \text{Var}[\hat{\gamma}_t^h(\theta_i)] \Big|_{\kappa=0} &= \frac{1}{4\pi f_{\text{sky}}} \int \frac{ldl}{2\pi} [\hat{J}_2(l\theta_i)]^2 \\ &\times \left[\left(C_l^h + \frac{1}{n_L} \right) \frac{\sigma_\gamma^2}{n_S} \right]. \end{aligned} \quad (6.27)$$

This is probably a reasonable approximation for the current measurements at $R \lesssim 30 h^{-1}$ Mpc; however, on larger scales which will be probed by the next-generation lens surveys, the cosmic variance due to cosmic shear must be included, as we show in Fig. 6.6.

For estimating the expected uncertainties, we assume a million lens galaxies with very narrow (delta-function like) redshift distribution centered at z_L ($N_L = 10^6$) over the full sky, $f_{\text{sky}} = 1$. We also assume $\sigma_\gamma = 0.3$, and $n_S = 3.5 \times 10^8 \text{ sr}^{-1}$. As the covariance matrix is dominated by the cosmic variance terms, the size of open boxes is insensitive to the exact values of N_L , σ_γ , or n_S . (See Sec. 6.2.3.) First, we calculate the binned uncertainties in the region close to the baryonic feature, $R \sim 110 h^{-1}$ Mpc. In Fig. 6.6, the open boxes show the full uncertainties including the cosmic variance due to cosmic shear (Eq. (6.26)), while the filled boxes show the uncertainties without the cosmic shear term (Eq. (6.27)). The latter is clearly negligible compared to the former on large scales, $R \gtrsim 50 h^{-1}$ Mpc.

Can we distinguish $\Delta\Sigma(R)$ with and without the baryonic feature? Without baryons, we do not see any features in $\Delta\Sigma(R)$; see dashed lines in Fig. 6.6 which are calculated from the smooth linear power spectrum without the baryonic feature (Eisenstein & Hu, 1998). To see if we can detect this feature in $\Delta\Sigma(R)$, we estimate the χ^2 difference between $\Delta\Sigma(R)$ with and without the baryonic feature:

$$\Delta\chi^2 \equiv \sum_{i,j} (\Delta\Sigma_i - \Delta\Sigma_{i,\text{nw}}) C_{ij}^{-1} (\Delta\Sigma_j - \Delta\Sigma_{j,\text{nw}}),$$

where $\Delta\Sigma_i$ is the mean tangential shear of i -th bin, $\Delta\Sigma_{\text{nw}}$ is $\Delta\Sigma$ without the baryonic feature, and C_{ij}^{-1} is the inverse of the binned covariance matrix (Eq. 6.25). Using only a single lens redshift slice, we find $\Delta\chi^2 = 0.85$ ($z_L = 0.3$, $b = 2$), 1.07 ($z_L = 0.5$, $b = 2$),

1.32 ($z_L = 0.8$, $b = 3$), and 1.34 ($z_L = 0.8$, $b = 5$). For example, if we add up all these measurements at different slices ($z_L = 0.3, 0.5$ and 0.8), significance of detection of the baryonic feature is $\Delta\chi^2 = 3.2$, i.e., 93% C.L. As we expect to have many more lens redshift slices from the future lens surveys, detection and measurement of the baryonic feature in $\Delta\Sigma$ are quite feasible. For multiple lens slices the gain in the signal-to-noise ratio will be approximately $\sqrt{N_{\text{lens}}}$; thus, for 10 lens slices the errors would be a factor of 3 smaller. At best we can expect ~ 25 slices, which gives a factor of 5 reduction in errors.

What about f_{NL} ? We show the expected $1\text{-}\sigma$ uncertainties for the mean tangential shears, $\Delta\Sigma(R)$, on larger scales in Fig. 6.8. For this figure we use logarithmic bins with the radial size of $\Delta R/R = 0.1$. We find that $\Delta\Sigma(R)$ on $R \simeq 250 h^{-1}$ Mpc is detectable, even from a single lens redshift slice. This is remarkable; however, the predicted uncertainties are too large for us to distinguish between $f_{\text{NL}} = 0$ and $f_{\text{NL}} = 100$ using a single lens redshift slice. In order to obtain a tight limit on f_{NL} , we would need to include many lens redshift slices.

Note that the uncertainty at a given R is larger for a smaller lens redshift. This is because a given R corresponds to a larger angular size for a lower lens redshift, making the cosmic variance contribution greater.

6.2 Harmonic Space Approach

6.2.1 Formula

The mean tangential shear, $\langle\bar{\gamma}_t^h\rangle$ or $\Delta\Sigma$, is currently widely used for measuring the halo-shear cross correlation, as this method is easy to implement and is less sensitive to systematic errors.

In this section, we shall study the effects of f_{NL} on the equivalent quantity in harmonic space: the halo-convergence cross power spectrum, $C_l^{h\kappa}$. The mean tangential shear is related to $C_l^{h\kappa}$ by the 2-dimensional Fourier integral given in Eq. (6.9).

The convergence field, $\kappa(\mathbf{n})$, is the matter density fluctuations projected on the sky:

$$\kappa(\mathbf{n}) = \int_0^\infty dz W_\kappa(z) \delta_m[d_A(0; z)\mathbf{n}, z], \quad (6.28)$$

where $\delta_m(\mathbf{r}, z) \equiv \rho_m(\mathbf{r}, z)/\bar{\rho}_m(z) - 1$, and $W_\kappa(z)$ is a lens kernel which describes the efficiency of lensing for a given redshift distribution of sources, $p(z_S)$:

$$W_\kappa(z) = \frac{\rho_0}{\Sigma_c(z; z_S)H(z)}, \quad (6.29)$$

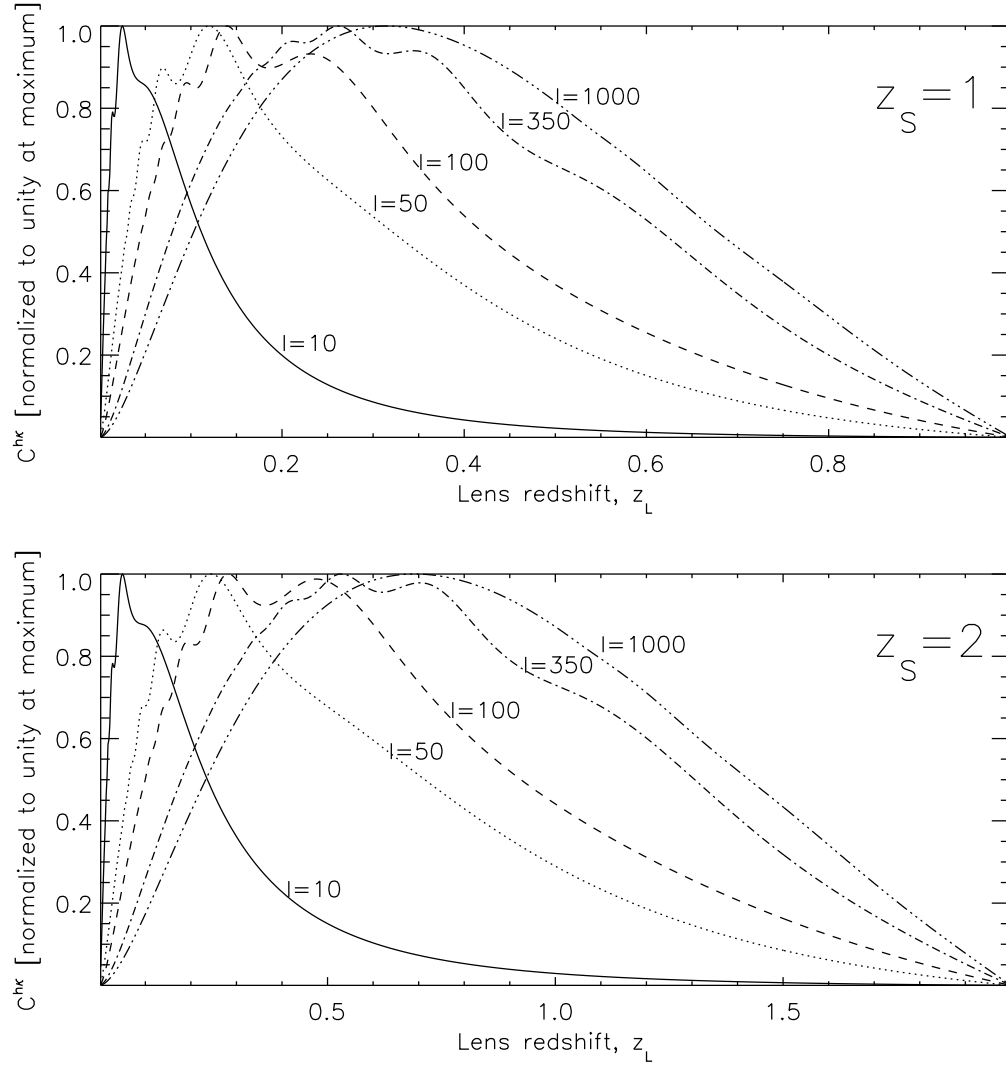


Figure 6.9: Angular power spectrum of the galaxy-convergence cross correlation, $C_l^{h\kappa}$, at various multipoles as a function of the lens redshift, z_L , for two effective source redshifts, $z_s = 1$ (top) and 2 (bottom). We have divided $C_l^{h\kappa}$ by its maximum value. The solid, dotted, dashed, dot-dashed, and triple-dot-dashed lines show $l = 10, 50, 100, 350$, and 1000, respectively.

where the critical density, Σ_c , is defined in Eq. (6.14).

Again using Limber's approximation (whose validity and limitation are studied in Appendix M), we find the relation between the angular cross-correlation power spectrum of the convergence field and the halo density at a given lens redshift z_L , $C_l^{h\kappa}(z_L)$, and the halo-mass cross-correlation power spectrum at the same redshift, $P_{hm}(k, z_L)$, as

$$\begin{aligned} C_l^{h\kappa}(z_L) &= \frac{\rho_0}{\Sigma_c(z_L; z_S) d_A^2(0; z_L)} P_{hm} \left[k = \frac{l + 1/2}{d_A(0; z_L)}, z_L \right] \\ &= \frac{4\pi G \rho_0}{c^2} (1 + z_L) \frac{d_A(z_L; z_S)}{d_A(0; z_L) d_A(0; z_S)} \\ &\quad \times P_{hm} \left[k = \frac{l + 1/2}{d_A(0; z_L)}, z_L \right]. \end{aligned} \quad (6.30)$$

Fig. 6.9 shows $C_l^{h\kappa}(z_L)$ for the Gaussian density field as a function of lens redshifts, z_L . The convergence fields at low (high) multipoles are better correlated with low- z (high- z) galaxies. This is due to the shape of the matter power spectrum: on very large scales (i.e., low l), the matter power spectrum is given by the initial power spectrum, $P_{hm}(k) \propto k$, and thus we get $1/d_A(0; z_L)$ from $P_{hm}[k = l/d_A(0; z_L)]$. This gives a larger weight to low- z galaxies. On smaller scales where $P_{hm}(k) \propto k^{n_{\text{eff}}}$ with $n_{\text{eff}} \simeq -3$, we get positive powers of $d_A(0; z_L)$ from $P_{hm}[k = l/d_A(0; z_L)]$, which gives a larger weight to high- z galaxies.

6.2.2 Result

We can now calculate $C_l^{h\kappa}$ for various values of f_{NL} . We use

$$\begin{aligned} C_l^{h\kappa}(z_L) &= \frac{4\pi G \rho_0}{c^2} (1 + z_L) \frac{d_A(z_L; z_S)}{d_A(0; z_L) d_A(0; z_S)} \\ &\quad \times \left[b_1(z_L) + \Delta b \left(k = \frac{l + 1/2}{d_A(0; z_L)}, z_L \right) \right] \\ &\quad \times P_m \left[k = \frac{l + 1/2}{d_A(0; z_L)}, z_L \right], \end{aligned} \quad (6.31)$$

where the scale-dependent bias, $\Delta b(k, z)$, is given by Eq. (6.1).

Figure 6.10 shows $C_l^{h\kappa}(z_L)$ for $f_{\text{NL}} = \pm 50$ and ± 100 for populations of galaxies that we have considered in the previous sections. For each lens redshift, we calculate the “effective” source redshift by requiring that the angular diameter distance to the source redshift is twice as large as that to the lens redshift, i.e., $d_A(0; z_S) = 2d_A(0; z_L)$. With this requirement, the source redshifts are $z_s = 0.65, 1.19$, and 2.25 for $z_L = 0.3, 0.5$, and 0.8 , respectively.

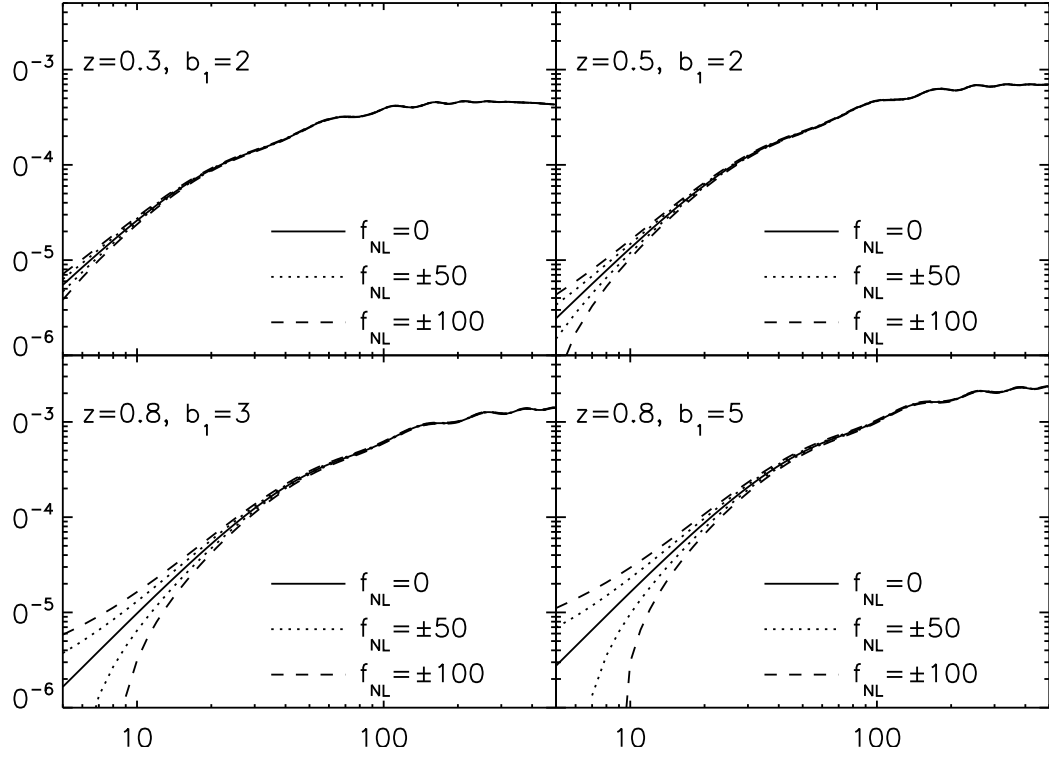


Figure 6.10: Imprints of the local-type primordial non-Gaussianity in the galaxy-convergence cross power spectrum, $l(l+1)C_l^{h\kappa}/(2\pi)$, for the same populations of lens galaxies as in Fig. 6.3. The solid, dashed, and dotted lines show $f_{NL} = 0, \pm 50$, and ± 100 , respectively.

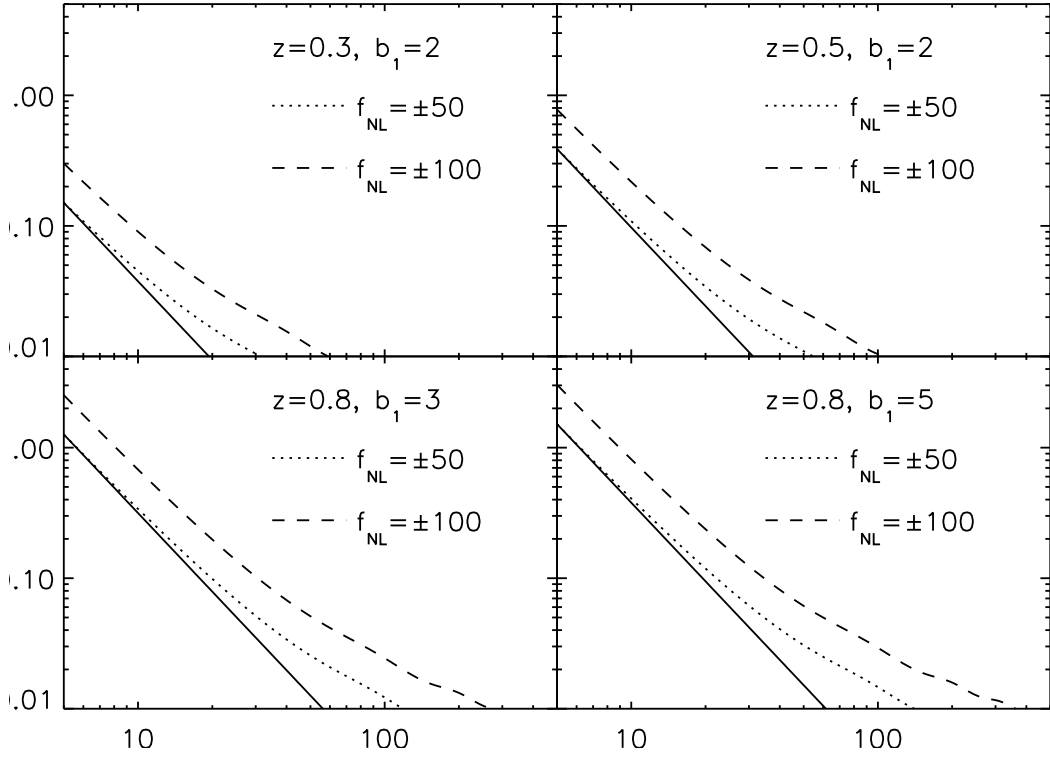


Figure 6.11: Fractional differences between C_l^{hk} from non-Gaussian initial conditions and the Gaussian initial condition, calculated from the curves shown in Fig. 6.10. These differences are equal to $|\Delta b(l = k/d_A, z_L)|/b_1(z_L)$. The dashed and dotted lines show $f_{NL} = \pm 50$ and ± 100 , respectively, while the thin solid lines show l^{-2} with an arbitrary normalization.

Figure 6.11 shows the fractional differences between non-Gaussian predictions and the Gaussian prediction ($f_{\text{NL}} = 0$), which are simply equal to $\Delta b(k, z_L)/b_1(z_L)$ where $k = l/d_A(0; z_L)$. As expected from the form of the scale-dependent bias, the difference grows toward small multipoles as roughly $1/l^2$. While lower redshift populations do not show more than 10% difference at $l \geq 10$ for $f_{\text{NL}} = \pm 50$, a higher- z population of lens galaxies or clusters of galaxies at $z_L = 0.8$ show the differences at the level of $\sim 10\%$ at $l \sim 20$ and $\sim 30\%$ at $l \sim 10$. Are these effects detectable?

6.2.3 Covariance matrix of the galaxy-convergence cross power spectrum

The covariance matrix of the galaxy-convergence cross-correlation power spectrum is given by

$$\begin{aligned} & \langle C_l^{h\kappa} C_{l'}^{h\kappa} \rangle - \langle C_l^{h\kappa} \rangle \langle C_{l'}^{h\kappa} \rangle \\ &= \frac{\delta_{ll'}}{(2l+1)f_{\text{sky}}} \left[(C_l^{h\kappa})^2 + \left(C_l^h + \frac{1}{n_L} \right) \left(C_l^\kappa + \frac{\sigma_\gamma^2}{n_S} \right) \right], \end{aligned} \quad (6.32)$$

where $\delta_{ll'}$ is Kronecker's delta symbol showing that the angular power spectra at different multipoles are uncorrelated. Again, C_l^h and C_l^κ are the angular power spectra of the lens halos (galaxies or cluster of galaxies) and κ , respectively, and n_L and n_S are the number densities of the lens halos and the lensed (source) galaxies, respectively.

We calculate C_l^κ by using Limber's approximation as

$$C_l^\kappa = \int_0^{z_S} dz \frac{\rho_0^2}{\Sigma_c^2(z; z_S)} \frac{P_m \left[k = \frac{l+1/2}{d_A(0; z)}; z \right]}{H(z) d_A^2(0; z)}. \quad (6.33)$$

However, we cannot use Limber's approximation for C_l^h unless one considers lens redshift slices that are broad. As we are assuming a thin lens redshift slice throughout this paper, we must not use Limber's approximation, but evaluate the exact integral relation:

$$C_l^h = \frac{2}{\pi} \int dk k^2 P_g(k, z_L) j_l^2[k d_A(z_L)], \quad (6.34)$$

where j_l is the spherical Bessel function, and $P_g(k, z)$ is the linear galaxy power spectrum: $P_g(k) = b_1^2 P_m(k)$.

Fig. 6.12 shows the galaxy-galaxy, galaxy-convergence, and convergence-convergence angular power spectra for Gaussian ($f_{\text{NL}} = 0$) initial conditions. We also show the shot noise of the galaxy angular power spectrum, $1/n_L$, and the shape noise of the convergence power spectrum, σ_γ^2/n_S , with the following representative values: $N_L = 4\pi n_L = 10^6$,

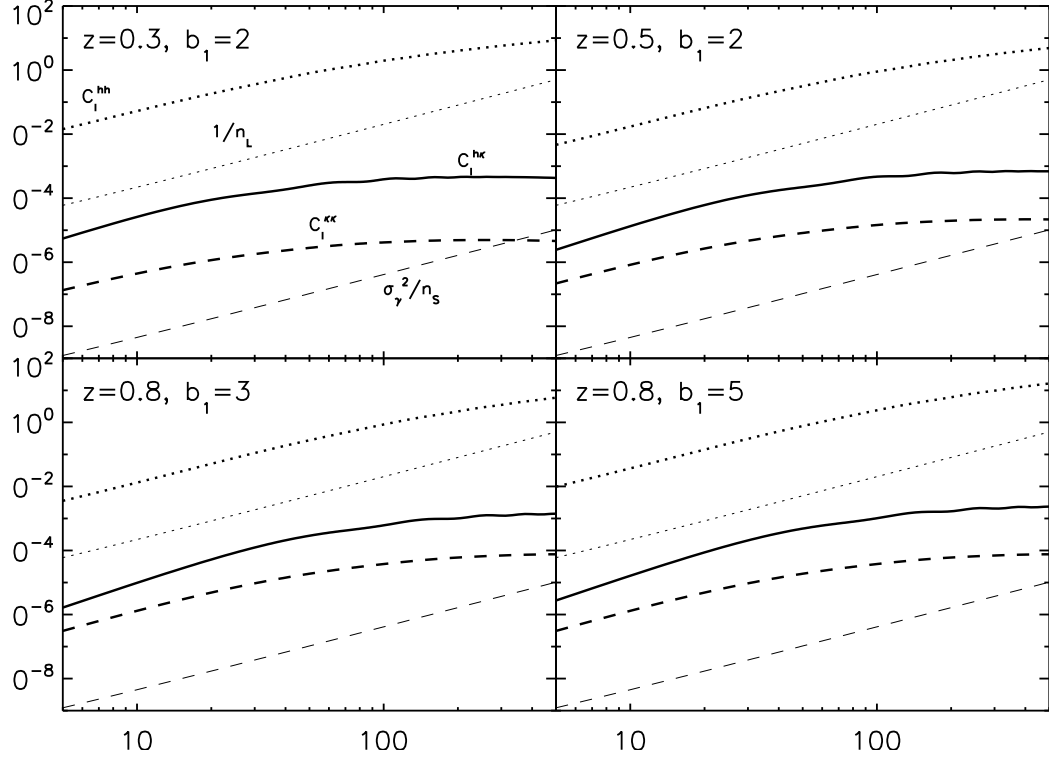


Figure 6.12: Angular power spectra of the galaxy-galaxy correlation, C_l^h (thick dotted lines), the galaxy-convergence cross-correlation, $C_l^{h\kappa}$ (thick solid lines), and the convergence-convergence correlation, C_l^κ (thick dashed lines) for the Gaussian initial condition ($f_{\text{NL}} = 0$). The four panels show the same populations of galaxies and clusters of galaxies as in Fig. 6.10. We also show the galaxy shot noise, $1/n_L$ (thin dotted lines) as well as the source shape noise, σ_γ^2/n_S (thin dashed lines), for $N_L = 10^6$, $\sigma_\gamma = 0.3$, and $n_S = 3.5 \times 10^8 \text{ sr}^{-1}$. We find $1/n_L \ll C_l^h$ and $\sigma_\gamma^2/n_S \ll C_l^\kappa$ for $l \lesssim 100$.

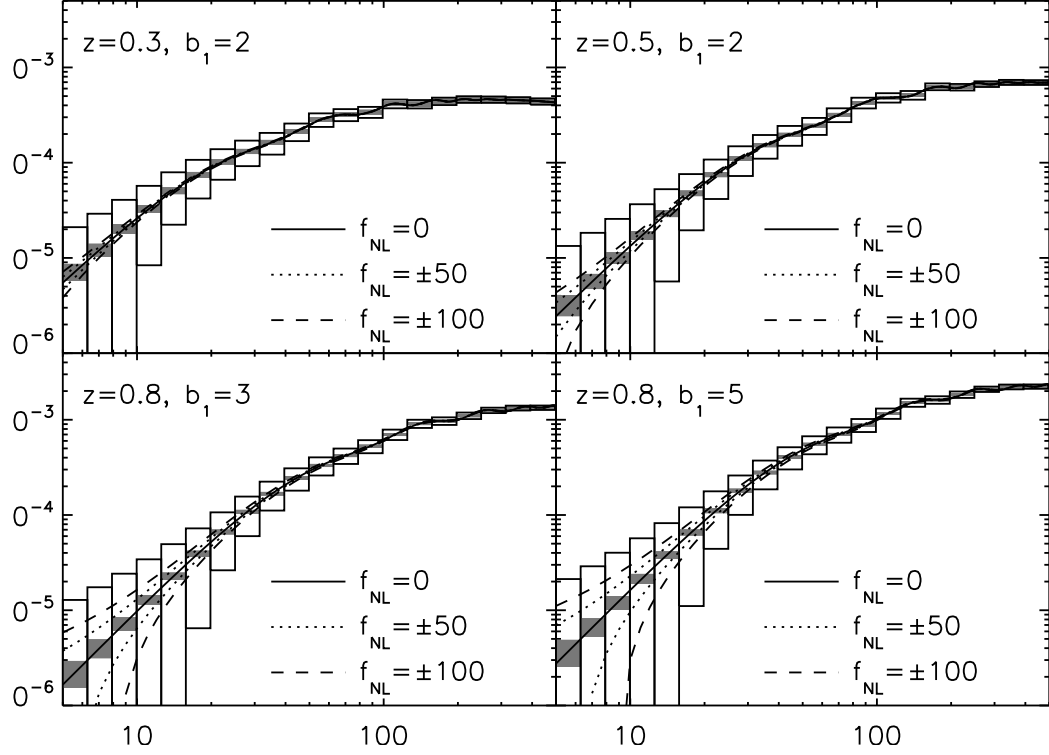


Figure 6.13: Same as Fig. 6.10, with the expected $1\text{-}\sigma$ uncertainties for full-sky lens surveys and a single lens redshift. Adjacent bins are uncorrelated. The open (filled) boxes show the binned uncertainties with (without) the cosmic variance term due to the cosmic shear field included. We used Eq. (6.32) for the open boxes, and Eq. (6.32) with $C_l^{h\kappa} = 0 = C_l^{\kappa}$ for the filled boxes. We use logarithmic bins of $\Delta l = 0.23l$. Note that the uncertainties are calculated for a single lens redshift slice, and thus they will go down as we add more lens redshift slices.

$n_S = 3.5 \times 10^8 \text{ sr}^{-1}$, and $\sigma_\gamma = 0.3$. We find $1/n_L \ll C_l^h$ and $\sigma_\gamma^2/n_S \ll C_l^\kappa$ for the multipoles that we are interested in, i.e., $l \lesssim 100$, and thus we conclude that the uncertainties are totally dominated by the cosmic variance terms. In other words, the size of the uncertainties are insensitive to the exact choices of N_L , σ_γ , or n_S .

We also find that the values of cross correlation coefficients, $r_l \equiv C_l^{h\kappa} / \sqrt{C_l^h C_l^\kappa}$, are small (of order 10–20%): the maximum values are 0.19, 0.15, and 0.13 for $z_L = 0.3, 0.5$, and 0.8 , respectively. This implies that one may ignore the contribution of $C_l^{h\kappa}$ to the covariance matrix, approximating the variance of $C_l^{h\kappa}$ of a single lens redshift slice for a multipole bin of size Δl as:

$$\text{Var}(C_l^{h\kappa}) = \frac{C_l^h C_l^\kappa}{(2l+1)\Delta l f_{\text{sky}}}. \quad (6.35)$$

Therefore, we should be able to measure the galaxy-convergence cross-power spectrum with $C_l^{h\kappa} / \sqrt{\text{Var}(C_l^{h\kappa})} \gtrsim 1$ when the multipoles satisfy

$$l \gtrsim l_{\min} \equiv \frac{1}{r_l \sqrt{2(\Delta l/l) f_{\text{sky}}}}. \quad (6.36)$$

For the galaxy-convergence power spectra in Fig. 6.10 with the full sky coverage ($f_{\text{sky}} = 1$) and $\Delta l/l = 0.23$, we find $l_{\min} = 9.0, 12.1$, and 15.7 for $z_L = 0.3$ ($z_S = 0.65$), 0.5 (1.19), and 0.8 (2.25), respectively.

Similarly, we can estimate the maximum radius below which we can measure the mean tangential shear, $\Delta\Sigma(R)$, as

$$R_{\max} \simeq \frac{\pi d_A(0; z_L)}{l_{\min}}. \quad (6.37)$$

For example, with $\Delta R/R = \Delta l/l = 0.1$, we get $R_{\max} \simeq 215, 260$, and $300 \text{ h}^{-1} \text{ Mpc}$ for $z_L = 0.3, 0.5$ and 0.8 , respectively. These values do give the radii at which the signal-to-noise ratios are roughly unity in Fig. 6.8.

Fig. 6.13 shows the expected 1- σ uncertainties of $C_l^{h\kappa}$ for several populations of lens galaxies. We find that the cosmic variance completely dominates the uncertainties on large scales (low l) where the non-Gaussian effects are the largest. Again, while we find that it would be difficult to measure f_{NL} from a *single lens redshift slice*, combining many redshift slices should help us measure f_{NL} , especially when we can use many slices at moderately high redshifts.

6.3 Halo-mass correlation from galaxy-CMB lensing

6.3.1 Formula

Instead of using the background galaxies for measuring the cosmic shear field due to the intervening mass, one can use the CMB as the background light and measure the shear field of the *CMB lensing* due to the intervening mass between us and the photon decoupling epoch at $z_* \simeq 1089$. See Lewis & Challinor (2006) for a review on the CMB lensing.

The lensing effect makes CMB anisotropies (both temperature and polarization) non-Gaussian by producing a non-vanishing connected four-point function, although it does not produce any non-vanishing three-point function. One can use this property to reconstruct the lensing potential field, hence the projected mass-density field between us and z_* , from the four-point function of CMB Hu & Okamoto (2002); Okamoto & Hu (2003); Hirata & Seljak (2003).

By cross-correlating the halo over-density field, δ_h , at some redshift z_L (measured from spectroscopic observations) and the κ field reconstructed from the CMB lensing, one can measure the halo-convergence angular power spectrum, $C_l^{h\kappa}$.

The angular power spectrum of the galaxy-CMB lensing cross correlation is merely a special case of the galaxy-convergence cross correlation that we have studied in the previous section: all we need to do is to set the source redshift, z_S , to be the redshift of the photon decoupling epoch, $z_* \simeq 1089$, i.e., $z_S = z_*$. Note that for a flat universe $d_A(z_L; z_*) = d_A(0; z_*) - d_A(0; z_L)$ where $d_A(0; z_*) = 9.83 \text{ } h^{-1} \text{ Gpc}$.

Figure 6.14 shows that the CMB lensing at low (high) multipoles are better correlated with low- z (high- z) galaxies. This is due to the shape of the matter power spectrum, as we have explained in the previous section. Note that $C_l^{h\kappa}$ of the CMB lensing for a given multipole decreases more slowly with z_L than that of the galaxy lensing due to the geometrical factor $d_A(z_L; z_S)/d_A(0; z_S)$.

Note that CMB and galaxies at $z \lesssim 1$ are correlated also via the Integrated Sachs-Wolfe (ISW) effect (Boughn & Crittenden, 2004). We shall not include this effect in our cross-correlation calculation for the following reason. We calculate the cross-correlation signal between galaxies and the convergence field reconstructed from CMB. This reconstruction relies on the fact that lensed CMB fluctuations have non-vanishing connected four-point function. On the other hand, the linear ISW effect does not have such a particular form of four-point function induced by lensing, and thus should not contribute to the reconstructed

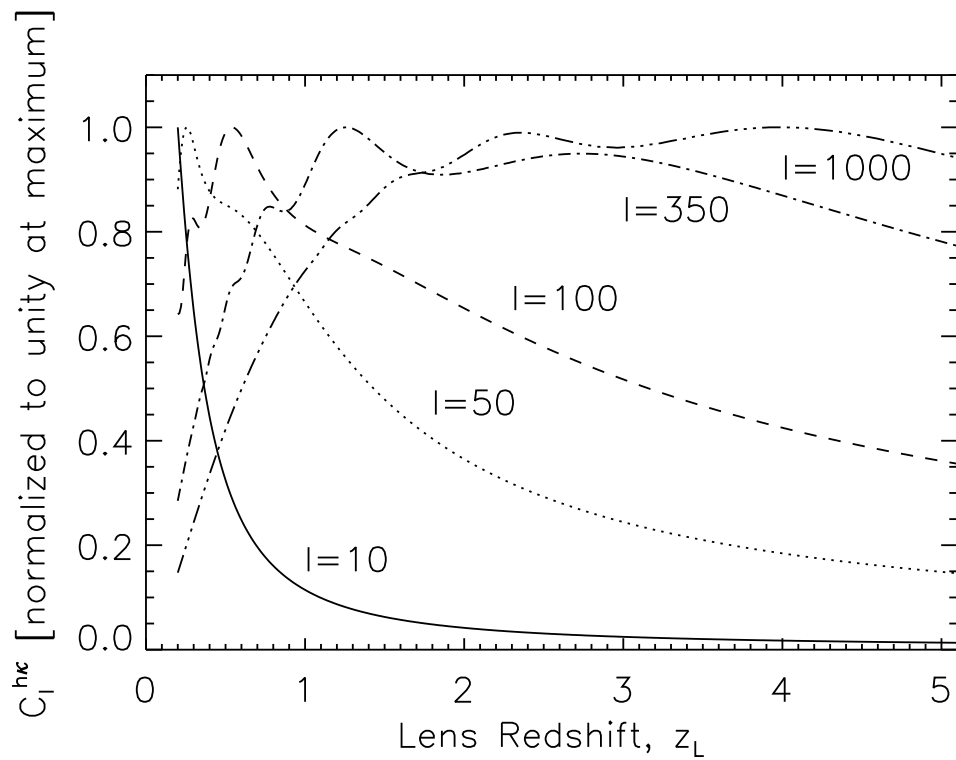


Figure 6.14: Angular power spectrum of the galaxy-CMB lensing, $C_l^{h\kappa}$, at various multipoles as a function of the lens redshift, z_L . We have divided $C_l^{h\kappa}$ by its maximum value. The solid, dotted, dashed, dot-dashed, and triple-dot-dashed lines show $l = 10, 50, 100, 350$, and 1000 , respectively.

convergence field. See Afshordi & Tolley (2008) for the effects of f_{NL} on the galaxy-ISW cross correlation.

6.3.2 Results

We can now calculate $C_l^{h\kappa}$ for various values of f_{NL} . We use

$$\begin{aligned} C_l^{h\kappa}(z_L) &= \frac{4\pi G\rho_0}{c^2}(1+z_L)\frac{d_A(z_L; z_*)}{d_A(0; z_L)d_A(0; z_*)} \\ &\times \left[b_1(z_L) + \Delta b\left(k = \frac{l}{d_A(0; z_L)}, z_L\right) \right] \\ &\times P_m\left[k = \frac{l}{d_A(0; z_L)}, z_L\right], \end{aligned} \quad (6.38)$$

where the scale-dependent bias, $\Delta b(k, z)$, is given by Eq. (6.1).

Figure 6.15 shows $C_l^{h\kappa}(z_L)$ for $f_{\text{NL}} = \pm 50$ and ± 100 for populations of low- z galaxies that we have considered in the previous sections: $b_1 = 2$ at $z_L = 0.3$ (similar to SDSS LRGs, top-left), $b_1 = 2$ at $z_L = 0.5$ (higher- z LRGs, top-right), $b_1 = 2$ at $z_L = 0.8$ (galaxies that can be observed by LSST, bottom-left), and $b_1 = 5$ at $z_L = 0.8$ (clusters of galaxies that can be observed by LSST, bottom-right). The fractional differences between non-Gaussian predictions and the Gaussian prediction ($f_{\text{NL}} = 0$) are exactly the same as those shown in Fig. 6.11: in the limit where Limber's approximation is valid, the galaxy-convergence power spectrum and the galaxy-CMB lensing power spectrum for the same lens galaxies differ only by a constant geometrical factor of $d_A(z_L; z_*)d_A(0; z_S)/d_A(z_L; z_S)d_A(0; z_*)$. Incidentally, for our choice of the source redshifts in the previous section, $2d_A(z_L; z_*)/d_A(0; z_*) = 1.83, 1.73$, and 1.60 for $z_L = 0.3, 0.5$, and 0.8 , respectively.

Therefore, the galaxy-CMB lensing cross correlation would provide a nice cross-check for systematics of the galaxy-convergence cross correlation, and *vice versa*: after all, we are measuring the same quantity, $P_{hm}(k)$, by two different background sources, high- z galaxies and CMB.

In using high- z galaxies as sources, the galaxy-galaxy lensing measurement may be susceptible to systematic errors widely discussed in the lensing literature, namely shear calibration, coherent point spread function (PSF) anisotropy, redshift biases, magnification bias and intrinsic alignments of galaxies. Here we are particularly concerned with errors that affect galaxy-shear cross-correlations by mimicking the angular dependence of the signal due to non-zero f_{NL} . Fortunately most systematic errors that affect shear-shear correlations do not contribute to galaxy-shear cross correlations: for instance, PSF anisotropy affects background galaxy shapes but not foreground galaxy locations (Mandelbaum et al., 2005). With standard lensing data analysis methods, it can be ensured that both the shear calibration

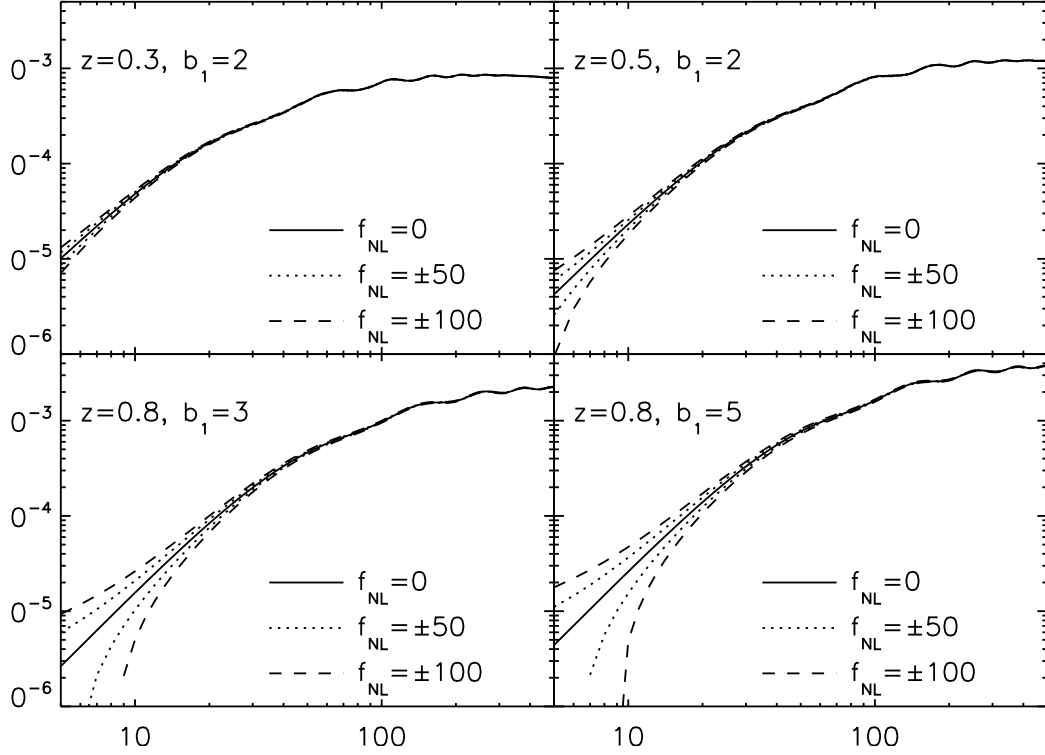


Figure 6.15: Imprints of the local-type primordial non-Gaussianity in the galaxy-CMB lensing power spectrum, $l(l+1)C_l^{h\kappa}/(2\pi)$, for the same populations of lens galaxies as in Fig. 6.3. The solid, dashed, and dotted lines show $f_{\text{NL}} = 0, \pm 50$, and ± 100 , respectively.

and PSF do not contribute a scale dependence to the first order. Biases in the redshift distributions of lens and source galaxies can similarly lead to a mis-estimation of the amplitude of the signal, but not its scale dependence. Thus, to the lowest order, the measurement of f_{NL} via the scale dependence of the galaxy-galaxy lensing signal is robust to the leading systematic errors in weak lensing. But a detailed study of various sources of error is needed given the small signal we are seeking.

Another benefit of using the CMB lensing as a proxy for the intervening matter distribution is that we can probe the galaxy-matter cross correlation at high redshift to which we cannot reach with the galaxy-galaxy lensing method. It is especially useful for probing primordial non-Gaussianity, as the scale dependent bias signal is higher for higher lens redshift: $\Delta b(k, z_L) \propto 1/D(z_L)$ (see Eq. (6.1)). Therefore, we find that even higher-

z populations of galaxies give us a much better chance of detecting the effects of f_{NL} . Figure 6.16 shows $C_l^{h\kappa}(z_L)$ for $f_{\text{NL}} = \pm 50$ and ± 100 for populations of high- z galaxies: $b_1 = 2$ at $z_L = 2$ (top-left), $b_1 = 2.5$ at $z_L = 3$ (top-right), $b_1 = 3$ at $z_L = 4$ (bottom-left), and $b_1 = 3.5$ at $z_L = 5$ (bottom-right). The first one, a spectroscopic galaxy survey at $z_L = 2$ with $b_L = 2$, is within reach by, e.g., the Hobby-Eberly Telescope Dark Energy Experiment (HETDEX) Hill et al. (2004); Hill et al. (2008). There we find, for $f_{\text{NL}} = \pm 50$, $\sim 10\%$ effect at $l \sim 40$, and a factor of two effect at $l \sim 10$ (see Fig. 6.17). The effects grow bigger at higher z : higher- z surveys at $z > 3$ can be done with, e.g., the concept of the Cosmic Inflation Probe (CIP) ⁷. At $z_L = 4$ and 5 (with $b_1 = 3$ and 4, respectively) we find $\sim 10\%$ effect at $l \sim 100$, a factor of two effect at $l \sim 30$, and even bigger effects at $l \lesssim 30$ (see Fig. 6.17).

6.3.3 Covariance matrix of the galaxy-CMB lensing

The covariance matrix of the galaxy-CMB lensing is given by Acquaviva et al. (2008)

$$\begin{aligned} & \langle C_l^{h\kappa} C_{l'}^{h\kappa} \rangle - \langle C_l^{h\kappa} \rangle \langle C_{l'}^{h\kappa} \rangle \\ = & \frac{(C_l^{h\kappa})^2 + (C_l^h + 1/n_L)(C_l^\kappa + N_l^\kappa)}{(2l+1)f_{\text{sky}}} \delta_{ll'}, \end{aligned} \quad (6.39)$$

where N_l^κ is the reconstruction noise from CMB given by Hu & Okamoto (2002). The covariance matrix equation here is the same as Eq. (6.32), except that now the shape noise of source galaxies is replaced by the reconstruction noise of CMB lensing. In what follows, we shall assume a “nearly perfect” CMB experiment considered in Hu & Okamoto (2002), whose Gaussian random detector noise is modeled as (Knox, 1995)

$$\begin{aligned} C_l^T \Big|_{\text{noise}} &= \left(\frac{T_{\text{CMB}}}{\Delta_T} \right)^{-2} e^{l(l+1)\sigma^2/8 \ln 2}, \\ C_l^E \Big|_{\text{noise}} = C_l^B \Big|_{\text{noise}} &= \left(\frac{T_{\text{CMB}}}{\Delta_T} \right)^{-2} e^{l(l+1)\sigma^2/8 \ln 2}, \end{aligned} \quad (6.40)$$

where the white noise level of detectors is $\Delta_T = \Delta_P/\sqrt{2} = 1 \mu\text{K arcmin}$, and the Full-Width-at-Half-Maximum (FWHM) of the beam is $\sigma = 4'$. With these detector parameters and the cosmological parameters of the “WMAP+BAO+SN ML” parameters in Table 1 of Komatsu et al. (2009), we find $N_l^\kappa \simeq 6 \times 10^{-8} \text{ sr}^{-1}$ on large scales, $l < 100$ ⁸.

⁷<http://www.cfa.harvard.edu/cip/>

⁸We summarize the calculation of lensing reconstruction noise of Hu & Okamoto (2002) in Appendix N.

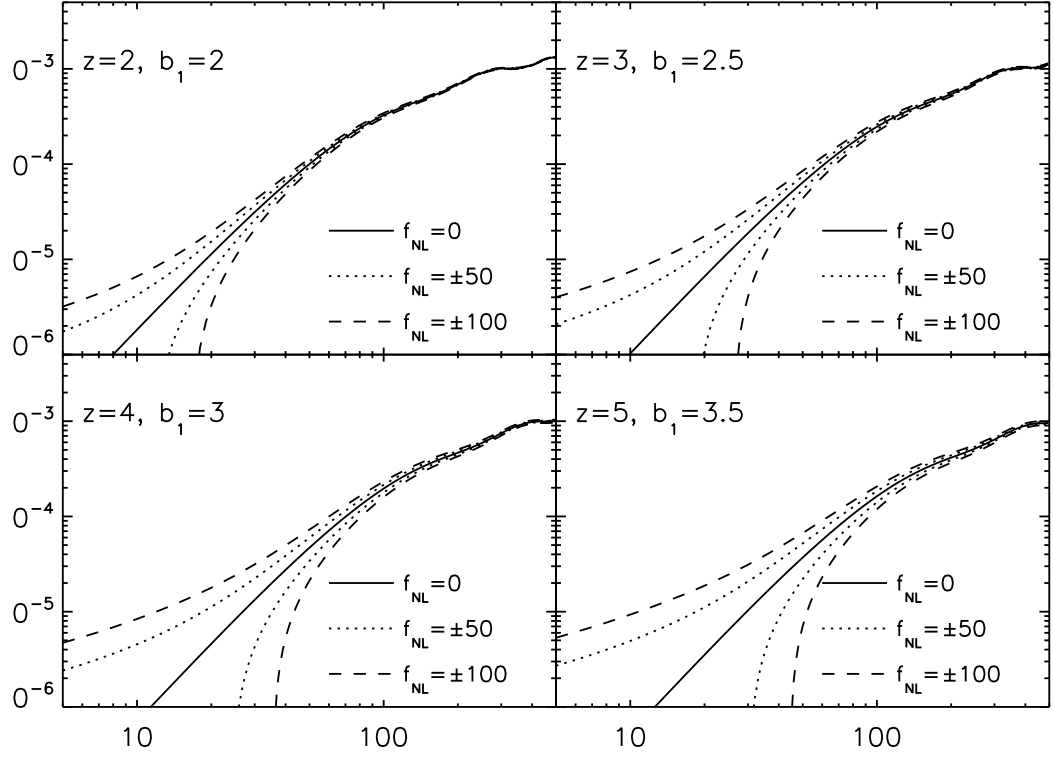


Figure 6.16: Same as Fig. 6.15, but for high- z lens galaxies with $b_1 = 2$ at $z_L = 2$ (top-left), $b_1 = 2.5$ at $z_L = 3$ (top-right), $b_1 = 3$ at $z_L = 4$ (bottom-left), and $b_1 = 3.5$ at $z_L = 5$ (bottom-right).

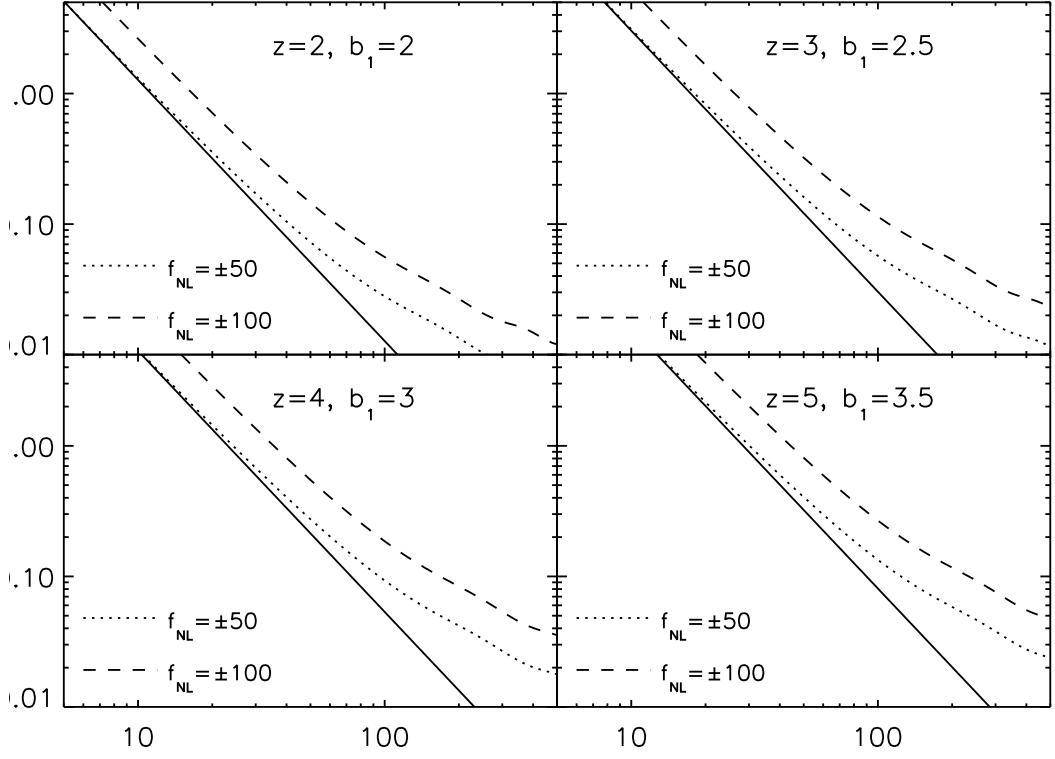


Figure 6.17: Same as Fig. 6.11, but for high- z lens galaxies with $b_1 = 2$ at $z_L = 2$ (top-left), $b_1 = 2.5$ at $z_L = 3$ (top-right), $b_1 = 3$ at $z_L = 4$ (bottom-left), and $b_1 = 3.5$ at $z_L = 5$ (bottom-right).

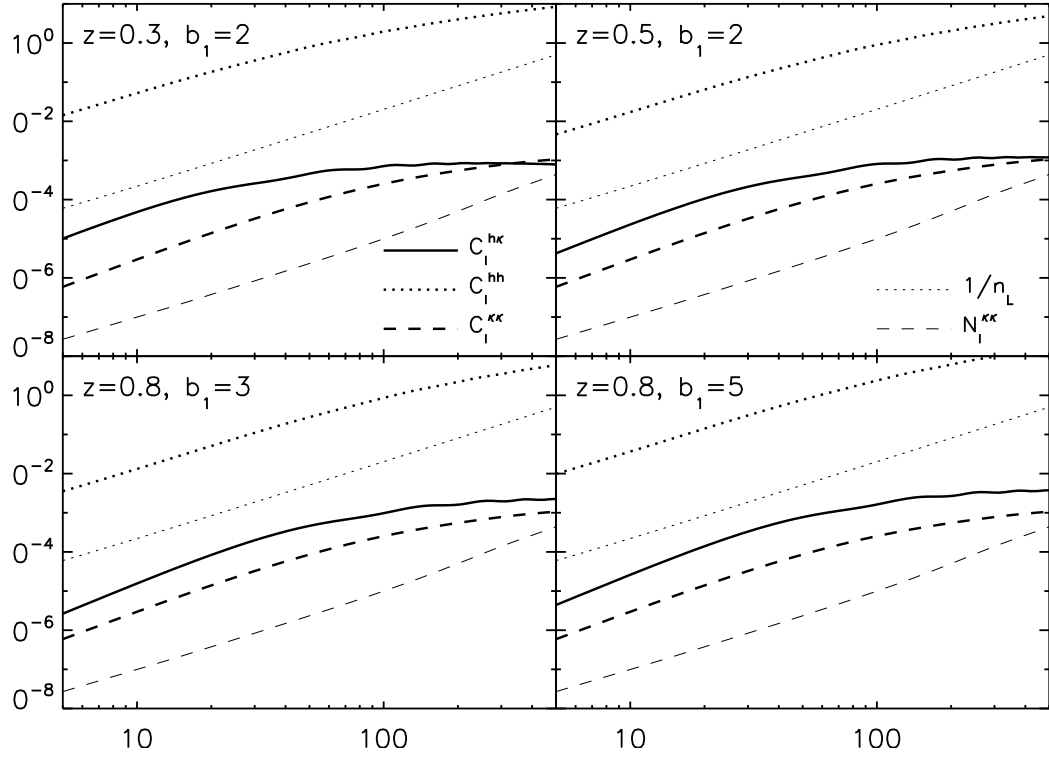


Figure 6.18: Angular power spectra of the galaxy-galaxy correlation, C_l^h (thick dotted lines), the galaxy-convergence cross-correlation, $C_l^{h\kappa}$ (thick solid lines), and the convergence-convergence correlation, C_l^κ (thick dashed lines) for the Gaussian initial condition ($f_{\text{NL}} = 0$). The four panels show the same populations of galaxies and clusters of galaxies as in Fig. 6.15. We also show the galaxy shot noise, $1/n_L$ (thin dotted lines) as well as the lens reconstruction noise, N_l^κ (thin dashed lines), for $N_L = 10^6$ and $N_l^\kappa \simeq 6 \times 10^{-8} \text{ sr}^{-1}$ (for multipoles much smaller than that corresponds to the beam size of $4'$). We find $1/n_L \ll C_l^h$ and $N_l^\kappa \ll C_l^\kappa$ for $l \lesssim 100$.

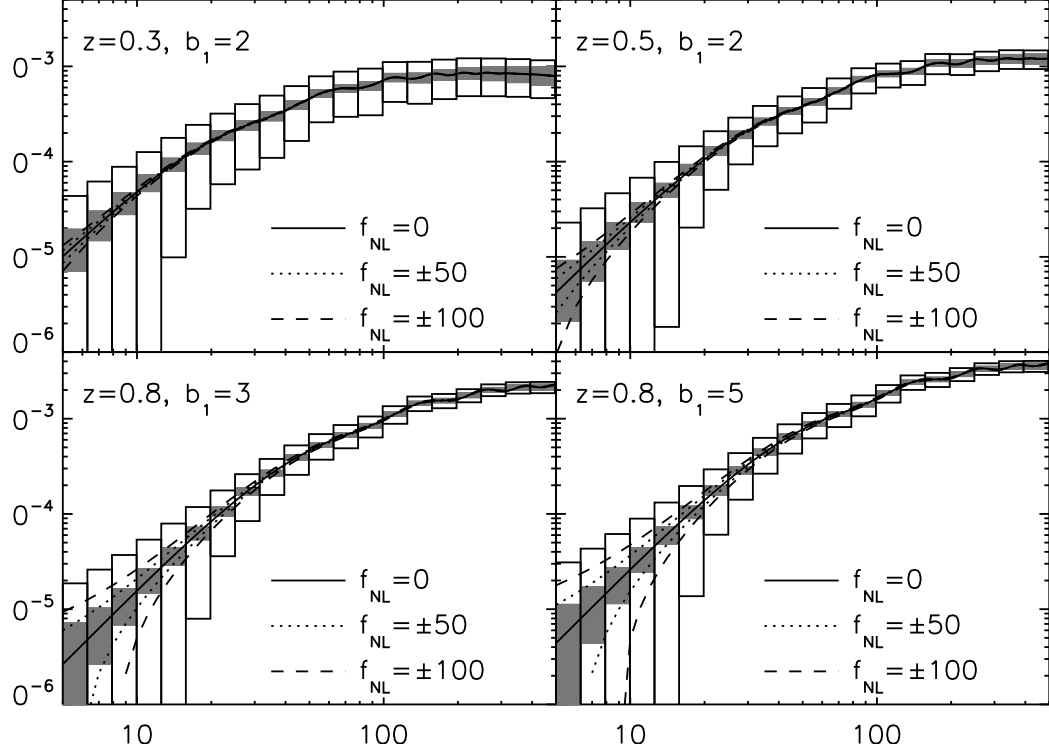


Figure 6.19: Same as Fig. 6.15, but with 1-sigma uncertainty due to the shape noise of source galaxies (filled box, Eq. (6.27)) and full error budget (empty box, diagonal of Eq. (6.25)) including the cosmic variance. We use the multipole bins of size $\Delta l = 0.23l$. For uncertainty of CMB lensing reconstruction, We assume the nearly-perfect reference experiment of Hu & Okamoto (2002): white detector noise $\Delta_T = \Delta_P/\sqrt{2} = 1 \mu K$ arcmin, and FWHM of the beam $\sigma = 4'$.

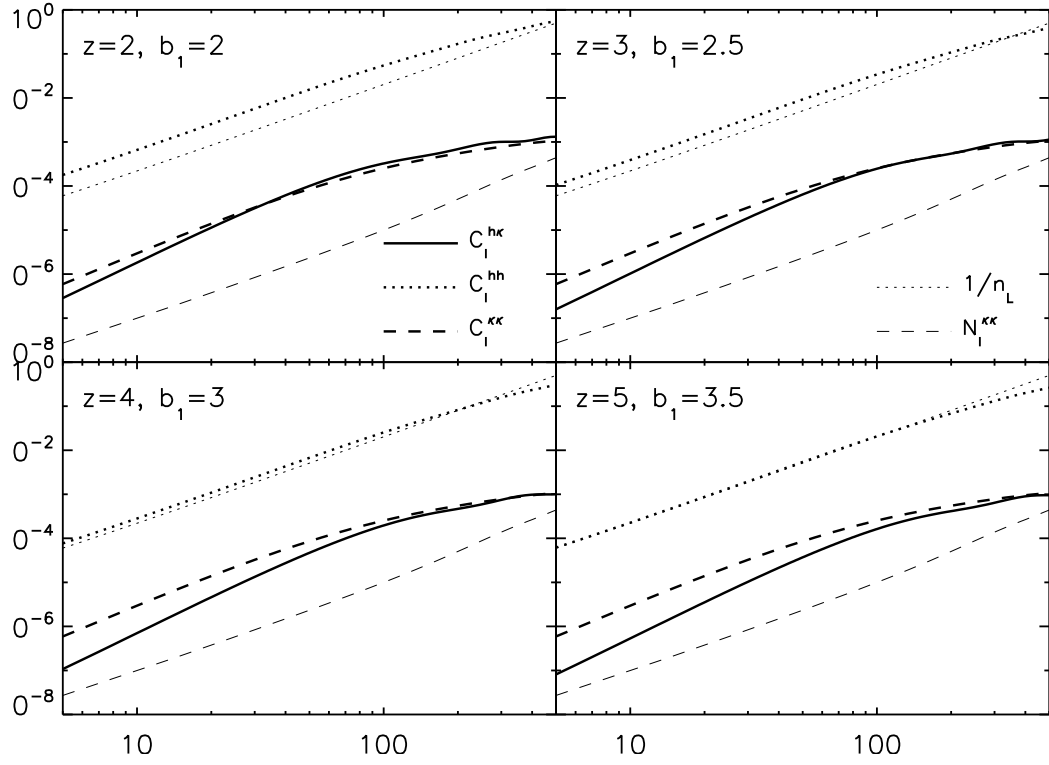


Figure 6.20: Same as Fig. 6.18, but for the high redshift lens galaxies shown in Fig. 6.16. For these populations (and with $N_L = 10^6$), the shot noise is about the same as the galaxy power spectrum, i.e., $C_l^h \simeq 1/n_L$.

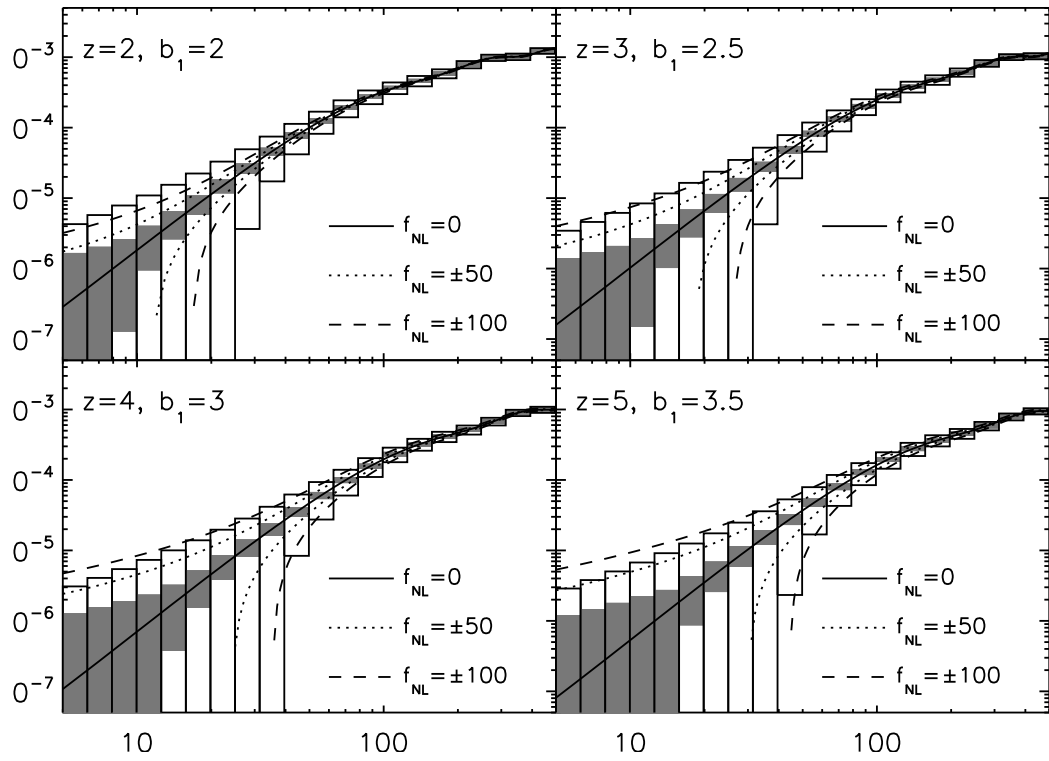


Figure 6.21: Same as Fig. 6.19, but for the high redshift lens galaxies shown in Fig. 6.16.

Fig. 6.18 shows the galaxy-galaxy, galaxy-convergence, convergence-convergence angular power spectra for the Gaussian initial condition ($f_{\text{NL}} = 0$). This figure is qualitatively similar to Fig. 6.12: the galaxy-galaxy correlation is exactly the same, and the galaxy-convergence power spectrum is simply a scaled version of the corresponding curve in Fig. 6.12. The major difference comes from C_l^κ : as the CMB photons travel a longer path than photons from source galaxies, the convergence-convergence power spectrum is higher for the CMB lensing convergence.

On large scales ($l \lesssim 100$), the covariance matrix is dominated by the cosmic variance terms: $1/n_L \ll C_l^h$ and $N_l^\kappa \ll C_l^\kappa$. The cross correlation coefficients are small, of order 10%: the maximum values are 0.12, 0.11, and 0.10 for $z_L = 0.3, 0.5$, and 0.8 , respectively. Therefore, we can again use Eq. (6.35) for estimating the variance, and find l_{min} (Eq. (6.36)) above which we can measure the galaxy-convergence cross correlation with the signal-to-noise ratio greater than unity. For logarithmic bins of $\Delta l/l = 0.23$, we find $l_{\text{min}} = 12.2, 13.5$, and 15.8 for $z_L = 0.3, 0.5$, and 0.8 , respectively. Comparing to the results in Sec 6.2.3, l_{min} is slightly bigger, as C_l^κ (which contributes to the uncertainty) increases more rapidly than $C_l^{h\kappa}$ (the signal we are after) would as the source redshift increases from z_S to z_* .

Fig. 6.19 shows the expected $1\text{-}\sigma$ uncertainties of the angular power spectrum of the galaxy-CMB lensing cross correlation, on top of the predicted Gaussian/non-Gaussian signals with five different values of non-Gaussianity parameters: $f_{\text{NL}} = 0, \pm 50, \pm 100$. We also show the $1\text{-}\sigma$ uncertainties without the cosmic variance due to the cosmic shear. Once again, it would be difficult to measure the effects of f_{NL} from a single lens redshift, but combining many slices would help measure f_{NL} from the galaxy-CMB lensing cross correlation.

What about using even higher- z lens galaxies? As shown in Fig. 6.20, for higher- z populations (with $z_L = 2 - 5$) the galaxy-galaxy power spectra are about the same as the shot noise levels. This is true only for the assumed number of lenses, $N_L = 10^6$ (over the full sky), which is somewhat arbitrary. Increasing N_L will help reduce the noise, but only up to a factor of $\sqrt{2}$. For populations with $C_l^h \simeq 1/n_L$, we can approximate the variance as

$$\text{Var}(C_l^{h\kappa}) = \frac{(C_l^h + 1/n_L)C_l^\kappa}{(2l+1)\Delta l f_{\text{sky}}} \simeq \frac{2C_l^h C_l^\kappa}{(2l+1)\Delta l f_{\text{sky}}}. \quad (6.41)$$

Thus, we find $C_l^{h\kappa}/\sqrt{\text{Var}(C_l^{h\kappa})} \gtrsim 1$ when

$$l \gtrsim \frac{1}{r_l \sqrt{(\Delta l/l) f_{\text{sky}}}}. \quad (6.42)$$

The maximum cross-correlation coefficients are 0.091, 0.084, 0.078, and 0.073 for $z_L = 2, 3, 4$, and 5, respectively. The estimated l_{\max} is then 29 ($z_L = 2$), 34 ($z_L = 3$), 38 ($z_L = 4$) and 42 ($z_L = 5$).

In Fig. 6.21 we compare the expected $1\text{-}\sigma$ uncertainties with the predicted signals from high- z lens galaxies with $f_{\text{NL}} = 0, \pm 50$, and ± 100 . Comparing this result with that in Fig. 6.19, we conclude that higher- z lens populations do provide a better chance of finding the effects of f_{NL} than lower- z lenses, although we would still need to combine many lens redshift slices. In particular, using higher- z lenses, we can find non-Gaussian effects at higher and higher multipoles which are easier to measure; thus, high- z galaxies correlated with CMB lensing offers a yet another nice probe of the local-type primordial non-Gaussianity.

6.4 Discussion and Conclusions

In this paper we have studied the galaxy-galaxy lensing and galaxy-CMB lensing cross-correlation functions. We have focused on large scales, typically larger than 100 Mpc at the lens redshift. While current measurements have high signal-to-noise ratios on much smaller scales, we believe that future surveys will enable detection of interesting physical effects in the large-scale, linear regime.

We derive the full covariance matrix for galaxy-galaxy lensing, including the cosmic variance due to the clustering of lenses and to cosmic shear (Eq. 6.16). We use the linear bias model to provide the halo-mass and halo-halo correlations needed for this calculation. We present results for the covariance of the mean tangential shear measurement as a function of angular separations, as well as for the harmonic space halo-convergence cross-power spectrum. Our calculations show that the errors in $\Delta\Sigma(R)$ are dominated by the cosmic variance term for $R \gtrsim 50 h^{-1}$ Mpc (see Fig. 6.6). Similarly, the errors in the halo-convergence cross power spectra, $C_l^{h\kappa}$, are dominated by the cosmic variance term at $l \lesssim 100$ (see Fig. 6.19).

For Gaussian initial conditions, we show that the baryonic effects in the matter power spectrum (often called Baryon Acoustic Oscillations) produce a “shoulder” in the galaxy-galaxy lensing correlation (i.e., the mean tangential shears), $\Delta\Sigma(R)$, at $R \sim 110 h^{-1}$ Mpc (see Fig. 6.3). This effect should be easy to measure from the next-generation lensing surveys by combining $\Delta\Sigma(R)$ from multiple lens redshift slices.

We consider the prospects of detecting primordial non-Gaussianity of the local-form, characterized by the f_{NL} parameter. We have found that the scale-dependent bias

from the local-form non-Gaussianity with $f_{\text{NL}} = \pm 50$ modifies $\Delta\Sigma(R)$ at the level of 10–20% at $R \sim 300 h^{-1} \text{ Mpc}$ (depending on b_1 and z_L ; see Fig. 6.5) (see Fig. 6.4). The modification grows rapidly toward larger scales, in proportion to R^2 . High- z galaxies at, e.g., $z \gtrsim 2$, cross-correlated with CMB can be used to find the effects of f_{NL} in the galaxy-convergence power spectrum, $C_l^{h\kappa}$. While the effects are probably too small to see from a single lens redshift (see Fig. 6.21), many slices can be combined to beat down the cosmic variance errors. Exactly how many slices are necessary, or what is the optimal strategy to measure f_{NL} from the galaxy-CMB lensing signal requires a more detailed study that incorporates the survey strategy for specific galaxy and lens surveys.

We emphasize that, while the two-point statistics of shear fields are not sensitive to primordial non-Gaussianity, the two-point statistics correlating shear fields with density peaks (i.e., galaxies and clusters of galaxies) are sensitive due to the strong scale-dependence of halo bias on large scales.

Finally, we note that one can also measure the effects of f_{NL} on the halo power spectrum, C_l^h . For example, C_l^h that would be measured from LSST can be used to probe $f_{\text{NL}} \sim 1$ (Carbone et al., 2008); thus, we would expect C_l^h to be more powerful than the lens cross-correlation statistics we studied here. However, a combination of the two measurements would provide useful cross-checks, as galaxy clustering and galaxy-lensing correlations are affected by very different systematics.

Chapter 7

Measuring Power spectrum and Bispectrum

In this chapter, we provide a practical guide for measuring power spectrum and bispectrum by using Fast Fourier Transforms. First, we show the method of calculating the power spectrum and bispectrum from N-body simulations, where input density contrast field is defined in a well defined cubic box. Then, we move to the power spectrum estimation for general cases and show the way we implement the FKP estimator (Feldman et al., 1994) with Fast Fourier Transform. For those two cases, we make special emphases on the normalization coefficients, the shot noise correction and the window function.

7.1 Power spectrum and Bispectrum from N-body simulation

Estimating power spectrum and bispectrum from the N-body simulation data is less complicated as N-body simulations have 1) the cubic box, 2) the constant mean number density. We divide the general procedure of measuring power spectrum and bispectrum from N-body simulation by following five steps:

- (1) Distributing particles onto the regular grid
- (2) Fourier transformation
- (3) Estimating power spectrum and bispectrum
- (4) Deconvolving window function
- (5) Subtracting shot noise

7.1.1 From particle to grid

In order to apply the Fast Fourier Transform technique, we have to assign the density field onto each point in the regular grid. The way we distribute a particle to the nearby grid points is called a ‘particle distribution scheme.’ For a given distribution scheme, we

can define an associated ‘shape function’, which quantifies how a quantity (mass, number, luminosity, etc) of particle is distributed. After this process, the sampling we made from the particle distribution is not a mere sampling of the underlying density field, but a sampling convolved with the ‘window function’ of particle distribution scheme. In this section, we shall review the three distribution schemes which are widely used in practice: Nearest-Grid-Point, Cloud-In-Cloud, Triangular-Shape-Cloud. Although we use the particle number density as a representative example below, one can use the same equation for calculating mass weighted power spectrum, luminosity weighted power spectrum, etc, by multiplying the apropos quantity (mass, luminosity) of each particle.

Let us consider the case with N_p particles (e.g. dark matters, halos, galaxies) in a N-body simulation box. The particle number density is given by

$$n_0(\mathbf{x}) = \sum_{i=1}^{N_p} \delta^D(\mathbf{x} - \mathbf{x}_i), \quad (7.1)$$

where \mathbf{x}_i is the position of i -th particle. Then, the particle number assignment can be formulated by the convolution as

$$n(\mathbf{x}) = \int_V d^3x' n_0(\mathbf{x}') W(\mathbf{x} - \mathbf{x}'), \quad (7.2)$$

where $W(\mathbf{x})$ is the window function which quantifies how much of this particle number density is distributed to a grid point separated by \mathbf{x} . We sample the continuous number density $n(\mathbf{x})$ by the regular grids of size N^3 :

$$n^s(\mathbf{x}_p) \equiv n(\mathbf{x}_p) = \int_V d^3x' n(\mathbf{x}') W(\mathbf{x}_p - \mathbf{x}'). \quad (7.3)$$

Therefore, the sampled density contrast, defined as $\delta^s(\mathbf{x}) \equiv n^s(\mathbf{x})/\bar{n} - 1$ is given by the convolution of the real density contrast and the window function,

$$\delta^m(\mathbf{x}) = [\delta \star W](\mathbf{x}), \quad (7.4)$$

and Fourier transformation of the sampled density contrast is

$$\delta^m(\mathbf{k}) = \delta(\mathbf{k}) W(\mathbf{k}). \quad (7.5)$$

This procedure of convolving with window function can be think of as following. First, we define the cloud shape function (or point spreading function) $S(\mathbf{x}')$ of a particle. The shape function can be uniquely determined for given distribution scheme such that the

fraction of the particle number of the particle at \mathbf{x} assigned to the grid point \mathbf{x}_p is given by integrating the shape function within the cubic cell surrounding \mathbf{x}_p (Hockney & Eastwood, 1988, p. 142). That is, the number of a particle at \mathbf{x} assigned to grid point \mathbf{x}_p $n(\mathbf{x} \rightarrow \mathbf{x}_p)$ is given by

$$n(\mathbf{x} \rightarrow \mathbf{x}_p) = \int_{|\mathbf{x}' - \mathbf{x}_p|_i < H/2} d^3x' S(\mathbf{x}' - \mathbf{x}). \quad (7.6)$$

In one-dimensional case, the window function can be written in terms of the cloud shape function as

$$W(x - x_p) = W_p(x) = \frac{1}{H} \int_{x_p - H/2}^{x_p + H/2} S(x' - x) dx', \quad (7.7)$$

where $H = L/N_{grid}$ is the separation of grids. Using the top-hat function, $\mathcal{T}(x)$,

$$\mathcal{T}(x) = \begin{cases} 1 & \text{if } |x| < 1/2 \\ 1/2 & \text{if } |x| = 1/2 \\ 0 & \text{if otherwise} \end{cases}, \quad (7.8)$$

equation (7.7) can be also written as a convolution of top-hat function and cloud function as:

$$W(x) = \frac{1}{H} \int \mathcal{T}\left(\frac{x'}{H}\right) S(x' - x) dx' \quad (7.9)$$

There are three most widely used distribution (window) functions.

7.1.1.1 NGP

Nearest Grid Point (NGP) scheme assigns particles to their nearest grid points. Therefore, the number density changes discontinuously when particles cross cell boundaries. The one dimensional window function for NGP is proportional to the top-hat function

$$W_{NGP}(x) \equiv \frac{1}{H} \mathcal{T}\left(\frac{x}{H}\right) = \begin{cases} 1/H & \text{if } |x| < H/2 \\ 1/(2H) & \text{if } |x| = H/2 \\ 0 & \text{if otherwise} \end{cases}, \quad (7.10)$$

and its point spreading function is the Dirac delta function as

$$\frac{1}{H} \mathcal{T}\left(\frac{x}{H}\right) = \frac{1}{H} \mathcal{T}\left(\frac{x}{H}\right) \otimes \delta^D(x) = \frac{1}{H} \mathcal{T}\left(\frac{x}{H}\right) \otimes \frac{1}{H} \delta^D\left(\frac{x}{H}\right). \quad (7.11)$$

The Fourier Transformation of the top-hat function is the sinc function.

$$\mathcal{T}(k) = \frac{\sin(k/2)}{k/2} = \text{sinc}\left(\frac{k}{2}\right) \quad (7.12)$$

Proof.

$$\mathcal{T}(k) = \int_{-1/2}^{1/2} e^{-ikx} dx = \frac{e^{-ik/2} - e^{ik/2}}{-ik} = \frac{\sin(k/2)}{k/2} \quad (7.13)$$

□

Therefore, the Fourier Transformation of the NGP window function is

$$W_{NGP}(k) = \text{FT}[\mathcal{T}](Hk) = \text{sinc}\left(\frac{Hk}{2}\right) = \text{sinc}\left(\frac{\pi k}{2k_N}\right)$$

where, $k_N = \pi/H$ is the Nyquist frequency. I use the similarity theorem of the Fourier transformation.

7.1.1.2 CIC

Cloud In Cell (CIC) assignment is the first order distribution scheme which uniformly distributes the particle with top-hat spreading function. In other words, the cloud shape function is given by

$$S_{CIC}(x) = \frac{1}{H} \mathcal{T}\left(\frac{x}{H}\right). \quad (7.14)$$

Therefore, the window function is

$$W_{CIC}(x) = \frac{1}{H} \mathcal{T}\left(\frac{x}{H}\right) \otimes \frac{1}{H} \mathcal{T}\left(\frac{x}{H}\right), \quad (7.15)$$

and its Fourier Transformation can be simply obtained by the convolution theorem:

$$W_{CIC}(k) = W_{NGP}(k)^2 = \text{sinc}^2\left(\frac{\pi k}{2k_N}\right). \quad (7.16)$$

The explicit expression of the CIC window function is

$$W_{CIC}(x) = \frac{1}{H} \begin{cases} 1 - |x|/H & \text{if } |x| < H \\ 0 & \text{otherwise} \end{cases}. \quad (7.17)$$

7.1.1.3 TSC

Triangular Shaped Cloud (TSC) scheme is the second order distribution scheme. Its point spreading function is triangular, as the convolution of the two first order (CIC) functions:

$$S_{TSC}(x) = \frac{1}{H} \mathcal{T}\left(\frac{x}{H}\right) \otimes \frac{1}{H} \mathcal{T}\left(\frac{x}{H}\right). \quad (7.18)$$

Therefore, the window function is

$$W_{TSC}(x) = \frac{1}{H} \mathcal{T}\left(\frac{x}{H}\right) \otimes \frac{1}{H} \mathcal{T}\left(\frac{x}{H}\right) \otimes \frac{1}{H} \mathcal{T}\left(\frac{x}{H}\right), \quad (7.19)$$

and its Fourier Transformation is given by

$$W_{TSC}(k) = W_{NGP}(k)^3 = \text{sinc}^3\left(\frac{\pi k}{2k_N}\right). \quad (7.20)$$

The explicit expression of TSC scheme is

$$W_{TSC}(x) = \frac{1}{H} \begin{cases} \frac{3}{4} - \left(\frac{x}{H}\right)^2 & \text{if } |x| \leq \frac{H}{2} \\ \frac{1}{2} \left(\frac{3}{2} - \frac{|x|}{H}\right)^2 & \text{if } \frac{H}{2} \leq |x| \leq \frac{3H}{2} \\ 0 & \text{otherwise} \end{cases}. \quad (7.21)$$

7.1.1.4 3D window function

As we use the regular cubic grid, the three dimensional window function is simply given as the multiplication of three one dimensional window functions.

$$W(\mathbf{x}) = W(x_1)W(x_2)W(x_3) \quad (7.22)$$

Therefore, its Fourier transformation is

$$W(\mathbf{k}) = \left[\text{sinc}\left(\frac{\pi k_1}{2k_N}\right) \text{sinc}\left(\frac{\pi k_2}{2k_N}\right) \text{sinc}\left(\frac{\pi k_3}{2k_N}\right) \right]^p, \quad (7.23)$$

where $p = 1, 2, 3$ for NGP, CIC and TSC, respectively.

7.1.2 Power spectrum and bispectrum: the estimators

The power spectrum and bispectrum are defined as

$$\langle \delta(\mathbf{k}_1) \delta(\mathbf{k}_2) \rangle = (2\pi)^3 P(k_1) \delta^D(\mathbf{k}_1 + \mathbf{k}_2) \quad (7.24)$$

$$\langle \delta(\mathbf{k}_1) \delta(\mathbf{k}_2) \delta(\mathbf{k}_3) \rangle = (2\pi)^3 B(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) \delta^D(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3). \quad (7.25)$$

Note that both power spectrum and bispectrum are real because of the parity invariance and the reality of the configuration space n -point correlation function. (See, appendix A of Smith et al. (2008b))

With this definition, σ_8 is

$$\sigma_8^2 = \int \frac{d^3k}{(2\pi)^3} P(k) |W(kR)|^2 = \int \frac{dk}{k} \Delta^2(k) |W(kR)|^2$$

with $R = 8\text{Mpc}/h$, and $\Delta^2(k) = P(k)k^3/2\pi^2$ is the dimensionless power spectrum. One can check the normalization of the power spectrum by calculating σ_8 . In some literature, the definition of power spectrum and bispectrum does not contain $(2\pi)^3$, and we have to drop $(2\pi)^3$ in the integration measure of σ_8^2 equation above.

For the normalization of bispectrum, one can check the value of the reduced bispectrum $Q(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)$ which is defined as

$$Q(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = \frac{B(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)}{P(k_1)P(k_2) + P(k_2)P(k_3) + P(k_3)P(k_1)}.$$

On large scales where so called ‘tree-level’ bispectrum model works, the reduced bispectrum is $4/7$ for the equilateral configuration.

In this section, we shall find the proper normalization to the power spectrum and bispectrum estimators which use the *unnormalized* Fast Fourier Transformation (FFT) such as FFTW. For denote the unnormalized discrete Fourier transform result by superscript ‘FTW’.

First, we nondimensionalize the Dirac delta function in the definition of power spectrum and bispectrum. Using the property of delta function,

$$\delta^D(f(x)) = \sum_{x_i \in \text{zeros}} \frac{1}{|f'(x_i)|} \delta^D(x - x_i),$$

we can express the delta function in Fourier space in terms of the Kronecker delta of integer triplet \mathbf{n}_k . We denote the Kronecker delta for such integer triplet as $\delta^K(\mathbf{n}_k)$.

$$\delta^D(\mathbf{k}_p) = \delta^D(k_F \mathbf{n}_k) = \prod_i \delta^D(k_F n_i) = \prod_i \frac{1}{k_F} \delta_{n_i}^K = \frac{1}{k_F^3} \delta^K(\mathbf{n}_k)$$

7.1.2.1 Power spectrum estimator: direct sampling

FTW output is the unnormalized DFT, which is related to the sampled Fourier space density field as

$$\delta^{FTW}(\mathbf{n}_k) = \sum_{\mathbf{n}_r} \delta(\mathbf{n}_r) e^{-i2\pi \mathbf{n}_k \cdot \mathbf{n}_r / N} = \sum_{\mathbf{r}_p} \delta(\mathbf{r}_p) e^{-i\mathbf{k}_p \cdot \mathbf{r}_p} = \frac{\delta(\mathbf{k}_p)}{H^3}. \quad (7.26)$$

From the definition of power spectrum

$$\begin{aligned} \langle \delta(\mathbf{k}_1) \delta(\mathbf{k}_2) \rangle &= H^6 \langle \delta^{FTW}(\mathbf{n}_1) \delta^{FTW}(\mathbf{n}_2) \rangle = (2\pi)^3 P(k_1) \delta^D(\mathbf{k}_1 + \mathbf{k}_2) \\ &= \frac{(2\pi)^3}{k_F^3} \delta^K(\mathbf{n}_1 + \mathbf{n}_2) P(k_F n_1), \end{aligned} \quad (7.27)$$

we find the normalization for the power spectrum from DFT as

$$P(k_F n_1) = \frac{H^6 k_F^3}{(2\pi)^3} \langle \delta^{FFTW}(\mathbf{n}_1) \delta^{FFTW}(-\mathbf{n}_1) \rangle \quad (7.28)$$

where V is the volume of survey, N is number of one-dimensional grid, $H^3 = V/N^3$ and $k_F^3 = (2\pi)^3/V$. Therefore, the final estimator for power spectrum is

$$P(k_F n_1) = \frac{V}{N^6} \langle |\delta^{FFTW}(\mathbf{n}_1)|^2 \rangle = \frac{V}{N^6} \left(\frac{1}{N_k} \sum_{|\mathbf{n}_k - \mathbf{n}_1| \leq \frac{1}{2}} |\delta^{FFTW}(\mathbf{n}_k)|^2 \right), \quad (7.29)$$

where we sum over all Fourier modes within $k_1 - k_F/2 < |\mathbf{k}| < k_1 + k_F/2$ to estimate the power spectrum at $k = k_1 = k_F n_1$.

7.1.2.2 Bispectrum estimator: direct sampling

From the definition of the bispectrum, we also find

$$\begin{aligned} \langle \delta(\mathbf{k}_1) \delta(\mathbf{k}_2) \delta(\mathbf{k}_3) \rangle &= H^9 \langle \delta^{FFTW}(\mathbf{n}_1) \delta^{FFTW}(\mathbf{n}_2) \delta^{FFTW}(\mathbf{n}_3) \rangle \\ &= (2\pi)^3 B(k_1, k_2, k_3) \delta^D(\mathbf{k}_{123}) = \frac{(2\pi)^3}{k_F^3} \delta^D(\mathbf{n}_{123}) B(k_1, k_2, k_3). \end{aligned} \quad (7.30)$$

Therefore, the estimator of bispectrum from DFT is

$$\begin{aligned} &B(k_F n_1, k_F n_2, k_F n_3) \\ &= \frac{H^9 k_F^3}{(2\pi)^3} \langle \delta^{FFTW}(\mathbf{n}_1) \delta^{FFTW}(\mathbf{n}_2) \delta^{FFTW}(\mathbf{n}_3) \rangle \delta^D(\mathbf{n}_{123}) \\ &= \frac{V^2}{N^9} \langle \delta^{FFTW}(\mathbf{n}_1) \delta^{FFTW}(\mathbf{n}_2) \delta^{FFTW}(\mathbf{n}_3) \rangle \delta^D(\mathbf{n}_{123}) \\ &= \frac{V^2}{N^9} \left(\frac{1}{N_{\text{triangle}}} \sum_{\mathbf{m} \in \text{Tri}_{123}} \delta^{FFTW}(\mathbf{m}_1) \delta^{FFTW}(\mathbf{m}_2) \delta^{FFTW}(\mathbf{m}_3) \right), \end{aligned} \quad (7.31)$$

where Tri_{123} is the set of $\{\mathbf{m}_1, \mathbf{m}_2, \mathbf{m}_3\}$ whose magnitude satisfies $|m_i - n_i| \leq 1/2$ and three vectors form a triangle, i.e. $\mathbf{m}_1 + \mathbf{m}_2 + \mathbf{m}_3 = \mathbf{0}$.

7.1.2.3 Estimating power spectrum II

There is another way of estimating the power spectrum. We slightly change the previous estimator as

$$\begin{aligned} P(k_F n_1) &= \frac{V}{N^6} \langle \delta^{FFTW}(\mathbf{n}_1) \delta^{FFTW}(\mathbf{n}_2) \rangle \delta^K(\mathbf{n}_1 + \mathbf{n}_2) \\ &= \frac{V}{N^6} \frac{\sum_{\mathbf{m}_1 \simeq \mathbf{n}_1} \sum_{\mathbf{m}_2 \simeq \mathbf{n}_1} \delta^{FFTW}(\mathbf{m}_1) \delta^{FFTW}(\mathbf{m}_2) \delta^K(\mathbf{m}_1 + \mathbf{m}_2)}{\sum_{\mathbf{m}_1 \simeq \mathbf{n}_1} \sum_{\mathbf{m}_2 \simeq \mathbf{n}_1} \delta^K(\mathbf{m}_1 + \mathbf{m}_2)}. \end{aligned} \quad (7.32)$$

Where, $m_i \simeq n_i$ strictly means that $|m_i - n_i| < s/2$, with $\delta k = sk_F$. Note that

$$\sum_{m_1 \simeq n_1} \sum_{m_2 \simeq n_1} \delta^K(\mathbf{m}_1 + \mathbf{m}_2) = \frac{1}{V_f} \int_{k_1} d^3 q_1 \int_{k_2} d^3 q_2 \delta^D(\mathbf{q}_1 + \mathbf{q}_2) \simeq \frac{4\pi k_1^2 \delta k}{k_F^3} = 4\pi s n_1^2$$

is the total number of independent k -modes inside the spherical shell of radius k and width $\delta k = sk_F$. We shall use the total number of k modes here instead of the actual number of k modes, because when we do the inverse DFT later, both $\delta(\mathbf{k})$ and $\delta(-\mathbf{k})$ shall be summed over.

Using the orthonormality of DFT

$$\delta^K(\mathbf{n}_p + \mathbf{n}_q) = \frac{1}{N^3} \sum_{\mathbf{n}_r} e^{2\pi i \mathbf{n}_r \cdot \mathbf{n}_p / N} e^{2\pi i \mathbf{n}_r \cdot \mathbf{n}_q / N}, \quad (7.33)$$

we transform the estimator as

$$\begin{aligned} P(k_F n_1) &= \frac{V}{N^6} \frac{1}{4\pi s n_1^2} \frac{1}{N^3} \sum_{\mathbf{n}_r} \sum_{m_1 \simeq n_1} \sum_{m_2 \simeq n_1} \delta^{FTW}(\mathbf{m}_1) \delta^{FTW}(\mathbf{m}_2) e^{2\pi i \mathbf{n}_r \cdot \mathbf{m}_1 / N} e^{2\pi i \mathbf{n}_r \cdot \mathbf{m}_2 / N} \\ &= \frac{V}{N^6} \frac{1}{4\pi s n_1^2} \frac{1}{N^3} \sum_{\mathbf{n}_r} \left[\sum_{m_1 \simeq n_1} \delta^{FTW}(\mathbf{m}_1) e^{2\pi i \mathbf{n}_r \cdot \mathbf{m}_1 / N} \right] \left[\sum_{m_2 \simeq n_1} \delta^{FTW}(\mathbf{m}_2) e^{2\pi i \mathbf{n}_r \cdot \mathbf{m}_2 / N} \right] \end{aligned} \quad (7.34)$$

Let's define $\delta_{\mathbf{m}_i}(\mathbf{n}_r)$ as

$$\delta_{\mathbf{m}_i}(\mathbf{n}_r) = \sum_{m \simeq n_i} \delta^{FTW}(\mathbf{m}) e^{2\pi i \mathbf{n}_r \cdot \mathbf{m} / N}$$

In practice, $\delta_{\mathbf{m}_i}(\mathbf{n}_r)$ can be calculated by applying `fftw_dft_c2r` to the array whose values are $\delta^{FTW}(\mathbf{m})$ when $|m - n_i| < s/2$ and otherwise zero. Note that $\delta_{\mathbf{m}_i}(\mathbf{n}_r)$ is real.

Proof. One can break the summation into two part, one with positive m_x index, and another with negative m_x index.

$$\begin{aligned} \delta_{\mathbf{n}_i}(\mathbf{n}_r) &= \sum_{m \simeq n_i} \delta^{FTW}(\mathbf{m}) e^{2\pi i \mathbf{n}_r \cdot \mathbf{m} / N} \\ &= \sum_{m \simeq n_i, m_x > 0} \left[\delta^{FTW}(\mathbf{m}) e^{2\pi i \mathbf{n}_r \cdot \mathbf{m} / N} + \delta^{FTW}(-\mathbf{m}) e^{-2\pi i \mathbf{n}_r \cdot \mathbf{m} / N} \right] \\ &= \sum_{m \simeq n_i, m_x > 0} \left[\delta^{FTW}(\mathbf{m}) e^{2\pi i \mathbf{n}_r \cdot \mathbf{m} / N} + [\delta^*]^{FTW}(\mathbf{m}) e^{-2\pi i \mathbf{n}_r \cdot \mathbf{m} / N} \right] \\ &= \sum_{m \simeq n_i, m_x > 0} \left[\delta^{FTW}(\mathbf{m}) e^{2\pi i \mathbf{n}_r \cdot \mathbf{m} / N} + \text{c.c.} \right] \in \mathbb{R} \end{aligned}$$

In third line, we used the reality of $\delta(\mathbf{r})$, then $\delta(-\mathbf{k}) = \delta^*(\mathbf{k})$. □

Then, we find the second estimator for the power spectrum as

$$P(k_F n_1) = \frac{V}{N^6} \frac{1}{4\pi s n_1^2} \frac{1}{N^3} \sum_{\mathbf{n}_r} \delta_{\mathbf{n}_1}(\mathbf{n}_r)^2. \quad (7.35)$$

This method takes more time than the first estimator which uses the direct sampling method. However, real strength of this method is apparent when we calculate the higher order polyspectra, e.g. Bispectrum and Trispectrum etc, because we do not have to explicitly sum up all the possible triangles (Bispectrum) or rectangles (Trispectrum), etc.

7.1.2.4 Estimating Bispectrum II

We can similarly construct the Bispectrum estimator. The direct sampling estimator we derived before is given by

$$B(k_F n_1, k_F n_2, k_F n_3) = \frac{V^2}{N^9} \langle \delta^{FTW}(\mathbf{n}_1) \delta^{FTW}(\mathbf{n}_2) \delta^{FTW}(\mathbf{n}_3) \rangle \delta^K(\mathbf{n}_1 + \mathbf{n}_2 + \mathbf{n}_3), \quad (7.36)$$

where we estimate the ensemble average by summing over the all possible triangles with side of $k_i - sk_F/2 < q_i < k_i + sk_F/2$. Therefore, the estimator can be recasted as

$$\begin{aligned} & B(k_F n_1, k_F n_2, k_F n_3) \\ &= \frac{V^2}{N^9} \frac{\sum_{m_1 \simeq n_1} \sum_{m_2 \simeq n_2} \sum_{m_3 \simeq n_3} \delta^{FTW}(\mathbf{m}_1) \delta^{FTW}(\mathbf{m}_2) \delta^{FTW}(\mathbf{m}_3) \delta^K(\mathbf{m}_{123})}{\sum_{m_1 \simeq n_1} \sum_{m_2 \simeq n_2} \sum_{m_3 \simeq n_3} \delta^K(\mathbf{m}_{123})}. \end{aligned} \quad (7.37)$$

Again, $m_i \simeq n_i$ means $|m_i - n_i| < s/2$. The denominator is the number of possible triangles with side of (k_1, k_2, k_3) :

$$\begin{aligned} \sum_{m_1 \simeq n_1} \sum_{m_2 \simeq n_2} \sum_{m_3 \simeq n_3} \delta^K(\mathbf{m}_{123}) &= \frac{1}{V_f^2} \int_{k_1} d^3 q_1 \int_{k_2} d^3 q_2 \int_{k_3} d^3 q_3 \delta^K(\mathbf{q}_{123}) \\ &\simeq \frac{8\pi^2 k_1 k_2 k_3 (\delta k)^3}{k_F^6} = 8\pi^2 s^3 n_1 n_2 n_3 \end{aligned} \quad (7.38)$$

By using the normalization of DFT,

$$\delta^D(\mathbf{n}_p + \mathbf{n}_q + \mathbf{n}_s) = \frac{1}{N^3} \sum_{\mathbf{n}_r} e^{2\pi i \mathbf{n}_r \cdot \mathbf{n}_p / N} e^{2\pi i \mathbf{n}_r \cdot \mathbf{n}_q / N} e^{2\pi i \mathbf{n}_r \cdot \mathbf{n}_s / N},$$

and following the exactly same procedure of finding the second estimator for power spectrum, we find the second bispectrum estimator

$$B(k_F n_1, k_F n_2, k_F n_3) = \frac{V^2}{N^9} \left(\frac{1}{8\pi^2 s^3 n_1 n_2 n_3} \right) \frac{1}{N^3} \sum_{\mathbf{n}_r} \delta_{\mathbf{n}_1}(\mathbf{n}_r) \delta_{\mathbf{n}_2}(\mathbf{n}_r) \delta_{\mathbf{n}_3}(\mathbf{n}_r). \quad (7.39)$$

7.1.2.5 Counting the number of triangles

In the previous section, we use the integral approximation to estimate the number of triangles. However, this approximation is broken down when three vectors are parallel to each other, i.e. when

$$\mathbf{k}_1 = \alpha \mathbf{k}_2 = \beta \mathbf{k}_3.$$

Also, it is not an accurate approximation for the triangle including the large scale modes. Therefore, in this section, we present the way we calculate the actual number of triangles for given triangular configurations.

We shall use the exactly same trick as you used in the previous section for estimating bispectrum. Since it is a trivial normalization factor, *the number of triangles are the same as the bispectrum for the unit density contrast.* i.e.

$$\begin{aligned} N_{tri}(n_1, n_2, n_3) &= \sum_{m_1 \simeq n_1} \sum_{m_2 \simeq n_2} \sum_{m_3 \simeq n_3} \delta^K(\mathbf{m}_1 + \mathbf{m}_2 + \mathbf{m}_3) \\ &= \frac{1}{N^3} \sum_{\mathbf{n}_r} \sum_{m_1 \simeq n_1} \sum_{m_2 \simeq n_2} \sum_{m_3 \simeq n_3} e^{2\pi i \mathbf{n}_r \cdot \mathbf{m}_1 / N} e^{2\pi i \mathbf{n}_r \cdot \mathbf{m}_2 / N} e^{2\pi i \mathbf{n}_r \cdot \mathbf{m}_3 / N} \\ &= \frac{1}{N^3} \sum_{\mathbf{n}_r} \left[\sum_{m_1 \simeq n_1} e^{2\pi i \mathbf{n}_r \cdot \mathbf{m}_1 / N} \right] \left[\sum_{m_2 \simeq n_2} e^{2\pi i \mathbf{n}_r \cdot \mathbf{m}_2 / N} \right] \left[\sum_{m_3 \simeq n_3} e^{2\pi i \mathbf{n}_r \cdot \mathbf{m}_3 / N} \right] \\ &\equiv \frac{1}{N^3} \sum_{\mathbf{n}_r} I_{n_1}(\mathbf{n}_r) I_{n_2}(\mathbf{n}_r) I_{n_3}(\mathbf{n}_r), \end{aligned}$$

where

$$I_{n_i}(\mathbf{n}_r) \equiv \sum_{m_i \simeq n_i} e^{2\pi i \mathbf{n}_r \cdot \mathbf{m}_i / N}.$$

The function $I_{n_i}(\mathbf{n}_r)$ is the inverse Fourier transformation of the function in k space, which has the value unity within the shell of $|m_i - n_i| < s/2$, otherwise zero.

By calculating the number of triangles, and compare the result with the exact calculation, we find that the logical size of the 1D Fourier grid has to be at least three times as large as that of the maximum wavenumber for which we want to estimate the bispectrum. That is, the array size of calculating $I_{n_i}(\mathbf{n}_r)$, **bnmesh**, has to satisfy

$$\mathbf{bnmesh} > 3(\mathbf{s} \times \mathbf{nkmax}), \quad (7.40)$$

where \mathbf{s} and \mathbf{nkmax} are the bin size, and maximum wavenumber in the unit of the fundamental frequency. This is to avoid the fictitious increasing of the number of triangles (when calculating the number of triangles) or power (when calculating bispectrum) due to the aliasing effect.

7.1.3 Deconvolution

Now, we have the estimators for the power spectrum and the bispectrum. However, as we have employed the distribution scheme, the power spectrum and the bispectrum we would measure with those estimators are not the same as the power/bi-spectrum of the ‘real’ density contrast, but the power/bi-spectrum of density contrast convolved with the window function. Therefore, the power spectrum and bispectrum we estimate will show the artificial power suppression on small scales. Therefore, we have to deconvolve the window function due to the particle distribution scheme in order to estimate the power spectrum and bispectrum of the *true* density contrast.

7.1.3.1 Deconvolving only window function

First, as we know the exact shape of the window function in Fourier space, we can simply divide the resulting density contrast in Fourier space by the window function. That is, we deconvolve each k mode of density contrast as

$$\delta(\mathbf{k}) = \frac{\delta^m(\mathbf{k})}{W(\mathbf{k})}, \quad (7.41)$$

or, deconvolve the estimated power spectrum by

$$P(\mathbf{k}) = \left| \frac{\delta^m(\mathbf{k})}{W(\mathbf{k})} \right|^2 = P^m(k_1, k_2, k_3) \left[\text{sinc} \left(\frac{\pi k_1}{2k_N} \right) \text{sinc} \left(\frac{\pi k_2}{2k_N} \right) \text{sinc} \left(\frac{\pi k_3}{2k_N} \right) \right]^{-2p}, \quad (7.42)$$

for $k < k_N$. Again, $p = 1, 2, 3$ for NGP, CIC and TSC scheme, respectively. Here, superscript m denote the measured quantity.

7.1.3.2 Deconvolving window function and aliasing

When we want to extract the power spectrum for $k \simeq k_N$ ¹, we have to take the alias effect into account. As we show in Appendix A.2.1, the Fourier counterpart of the sampled data are the aliased sum of Fourier transformation of the continuous function. i.e.

$$P^m(\mathbf{k}) = \sum_{\mathbf{n}} |W(\mathbf{k} + 2k_N \mathbf{n})|^2 P(\mathbf{k} + 2k_N \mathbf{n}) \quad (7.43)$$

For the case of uniformly distributed random particles, the power spectrum is simply a constant (Poisson shot noise), $1/n$, where n is the number density. In that case, the

¹Actually, it may not happen for most of N-body simulations because of the Force/Mass resolution is usually too poor to use the power spectrum for such a high wavenumber.

convolved power spectrum becomes

$$P^m(\mathbf{k}) = \frac{1}{n} \sum_n |W(\mathbf{k} + 2k_N \mathbf{n})|^2. \quad (7.44)$$

Let us calculate the convolved power spectrum analytically for $p = 1, 2, 3$ cases:

$$\begin{aligned} \sum_n |W(\mathbf{k} + 2k_N \mathbf{n})|^2 &= \prod_{i=1}^3 \left[\sum_{n=-\infty}^{\infty} W(k_i + 2k_N n)^2 \right] \\ &= \prod_{i=1}^3 \left[\sum_{n=-\infty}^{\infty} \frac{\sin^{2p} \left(\frac{\pi}{2k_N} (k_i + 2k_N n) \right)}{\left(\frac{\pi}{2k_N} (k_i + 2k_N n) \right)^{2p}} \right] \\ &= \prod_{i=1}^3 \left[\sin^{2p} \left(\frac{\pi k_i}{2k_N} \right) \sum_{n=-\infty}^{\infty} \frac{1}{\left(\frac{\pi k_i}{2k_N} + \pi n \right)^{2p}} \right] \end{aligned} \quad (7.45)$$

Here, the infinite summation can be calculated by using following identity (Jing, 2005). (so called *Glaisher's series* ²)

$$\sum_{M=-\infty}^{\infty} \frac{1}{(a + Md)^2} = \left[\frac{\pi}{d} \csc \frac{a\pi}{d} \right]^2 \quad (7.46)$$

NGP (p=1)

Let's put $d = \pi$.

$$\sum_{n=-\infty}^{\infty} \frac{1}{(a + n\pi)^2} = \csc^2 a \quad (7.47)$$

Therefore, for $p = 1$,

$$\sum_n |W(\mathbf{k} + 2k_N \mathbf{n})|^2 = \prod_{i=1}^3 \sin^2 \left(\frac{\pi k_i}{2k_N} \right) \csc^2 \left(\frac{\pi k_i}{2k_N} \right) = 1, \quad (7.48)$$

and sum over infinite aliasing effect of power suppression ends up giving rise to the original power.

CIC (p=2)

Differentiate both side with respect to a twice, we get

$$\sum_{n=-\infty}^{\infty} \frac{6}{(a + n\pi)^4} = 2(3 - 2\sin^2 a) \csc^4 a \quad (7.49)$$

²F. Soddy, Proceedings of the Royal Society of London, Series A, Mathematical and Physical Sciences, Vol. 182, No. 989. (Dec. 16.1943), pp. 113-129

Therefore, the aliased sum becomes

$$\begin{aligned}\sum_{\mathbf{n}} |W(\mathbf{k} + 2k_N \mathbf{n})|^2 &= \prod_{i=1}^3 \sin^4 \left(\frac{\pi k_i}{2k_N} \right) \left[1 - \frac{2}{3} \sin^2 \left(\frac{\pi k_i}{2k_N} \right) \right] \csc^4 \left(\frac{\pi k_i}{2k_N} \right) \\ &= \prod_{i=1}^3 \left[1 - \frac{2}{3} \sin^2 \left(\frac{\pi k_i}{2k_N} \right) \right].\end{aligned}\quad (7.50)$$

TSC (p=3)

Differentiate the both side with a four times, we get

$$\sum_{n=-\infty}^{\infty} \frac{120}{(a + n\pi)^6} = (120 - 120 \sin^2 a + 16 \sin^4 a) \csc^6 a \quad (7.51)$$

Therefore,

$$\begin{aligned}\sum_{\mathbf{n}} |W(\mathbf{k} + 2k_N \mathbf{n})|^2 &= \prod_{i=1}^3 \sin^6 \left(\frac{\pi k_i}{2k_N} \right) \left[1 - \sin^2 \left(\frac{\pi k_i}{2k_N} \right) + \frac{2}{15} \sin^4 \left(\frac{\pi k_i}{2k_N} \right) \right] \csc^6 \left(\frac{\pi k_i}{2k_N} \right) \\ &= \prod_{i=1}^3 \left[1 - \sin^2 \left(\frac{\pi k_i}{2k_N} \right) + \frac{2}{15} \sin^4 \left(\frac{\pi k_i}{2k_N} \right) \right].\end{aligned}\quad (7.52)$$

When the measured power spectrum include the both *true* power spectrum and Poisson shot noise, we have to cure it iteratively, See, e.g. Jing (2005). However, as the nonlinear galaxy power spectrum is dominated by the constant term P_0 and large k plateau of nonlinear bias terms, P_{b2} , P_{b22} , we may simply apply the same aliased window function as for the constant power spectrum. In practice, we calculated both window-corrected and window-alias-corrected power spectrum, and only use the wavenumber ranges where those two power spectra agree with each other.

7.1.4 Poisson shot noise

Finally, we have to subtract the Poisson shot noise from the deconvolved power spectrum and bispectrum. As a general terminology, ‘shot noise’ refers to the self-particle contribution to the statistics. We call it Poisson shot noise, as it appears when we think of a realization of galaxy (matter) distribution as a Poisson sampling of underlying smooth galaxy (matter) density contrast field. Poisson shot noise appears whenever we calculate the n -point function from a discrete set of objects, like a galaxy, dark matter, etc.

7.1.4.1 Poisson sampling and underlying density field

Suppose we have a point process $n(\mathbf{r})$ which is a “Poisson sample” of some continuous stochastic field $1 + \delta(\mathbf{r})$. That is, the probability that an infinitesimal volume element δV contains an object is $\bar{n}(\mathbf{r}) [1 + \delta(\mathbf{r})] \delta V$.

Following Peebles (1980, §36) we describe the process by dividing the space into the infinitesimal micro-cells of volume δV which has a occupation numbers $n_i = 0$ or 1. That is, the statistical average of self-correlator for a given cell is

$$\langle n_i^3 \rangle = \langle n_i^2 \rangle = \langle n_i \rangle = \bar{n}(\mathbf{r}_i) \delta V_i, \quad (7.53)$$

and the correlator for different cells are given by the underlying density contrast as

$$\begin{aligned} & \langle n_i n_j \rangle_{i \neq j} \\ &= \bar{n}(\mathbf{r}_i) \bar{n}(\mathbf{r}_j) \delta V_i \delta V_j [1 + \langle \delta(\mathbf{r}_i) \delta(\mathbf{r}_j) \rangle] \end{aligned} \quad (7.54)$$

$$\begin{aligned} & \langle n_i n_j n_k \rangle_{i \neq j \neq k} \\ &= \bar{n}(\mathbf{r}_i) \bar{n}(\mathbf{r}_j) \bar{n}(\mathbf{r}_k) \delta V_i \delta V_j \delta V_k \\ & \times [1 + \langle \delta(\mathbf{r}_i) \delta(\mathbf{r}_j) \rangle + \langle \delta(\mathbf{r}_j) \delta(\mathbf{r}_k) \rangle + \langle \delta(\mathbf{r}_i) \delta(\mathbf{r}_k) \rangle + \langle \delta(\mathbf{r}_i) \delta(\mathbf{r}_j) \delta(\mathbf{r}_k) \rangle]. \end{aligned} \quad (7.55)$$

7.1.4.2 Power spectrum and Bispectrum of discrete particles

We follow the derivation given in Feldman et al. (1994). Consider the expectation value of

$$\int d^3 r \int d^3 r' g(\mathbf{r}, \mathbf{r}') n(\mathbf{r}) n(\mathbf{r}'),$$

for an arbitrary function $g(\mathbf{r}, \mathbf{r}')$. By using the infinitesimal micro-cells, the expectation value becomes

$$\begin{aligned} & \left\langle \int d^3 r \int d^3 r' g(\mathbf{r}, \mathbf{r}') n(\mathbf{r}) n(\mathbf{r}') \right\rangle \\ &= \int d^3 r \int d^3 r' g(\mathbf{r}, \mathbf{r}') \langle n(\mathbf{r}) n(\mathbf{r}') \rangle \\ &= \sum_{i,j} g(\mathbf{r}_i, \mathbf{r}_j) \langle n_i n_j \rangle \\ &= \sum_{i \neq j} \delta V_i \delta V_j g(\mathbf{r}_i, \mathbf{r}_j) \bar{n}(\mathbf{r}_i) \bar{n}(\mathbf{r}_j) [1 + \langle \delta(\mathbf{r}_i) \delta(\mathbf{r}_j) \rangle] + \sum_{i=j} \delta V_i g(\mathbf{r}_i, \mathbf{r}_i) \bar{n}(\mathbf{r}_i) \\ &= \int d^3 r \int d^3 r' g(\mathbf{r}, \mathbf{r}') \bar{n}(\mathbf{r}) \bar{n}(\mathbf{r}') [1 + \langle \delta(\mathbf{r}) \delta(\mathbf{r}') \rangle] + \int d^3 r g(\mathbf{r}, \mathbf{r}) \bar{n}(\mathbf{r}) \\ &= \int d^3 r \int d^3 r' g(\mathbf{r}, \mathbf{r}') \{ \bar{n}(\mathbf{r}) \bar{n}(\mathbf{r}') [1 + \langle \delta(\mathbf{r}) \delta(\mathbf{r}') \rangle] + \bar{n}(\mathbf{r}) \delta^D(\mathbf{r} - \mathbf{r}') \}. \end{aligned} \quad (7.56)$$

As this equation holds for an arbitrary function $g(\mathbf{r}, \mathbf{r}')$, comparing the first and last line of the equation above, we find the relation between the two point correlator of the discrete number density and that of underlying density contrast:

$$\langle n(\mathbf{r})n(\mathbf{r}') \rangle = \bar{n}(\mathbf{r})\bar{n}(\mathbf{r}') [1 + \langle \delta(\mathbf{r})\delta(\mathbf{r}') \rangle] + \bar{n}(\mathbf{r})\delta^D(\mathbf{r} - \mathbf{r}'). \quad (7.57)$$

The 2nd term in Equation (7.57) is called a *Poisson shot noise*, as the Dirac delta function manifests its identity as a self-particle contribution. The shot noise term $\bar{n}(\mathbf{r})\delta^D(\mathbf{r} - \mathbf{r}')$ can also be understood as follows. Because the presence of galaxies obeys the Poisson statistics, for a single position \mathbf{r} with the mean number density $\bar{n}(\mathbf{r})$, its variance is $\sigma^2[n(\mathbf{r})] \equiv \langle n(\mathbf{r})n(\mathbf{r}) \rangle - \bar{n}^2(\mathbf{r}) = \bar{n}(\mathbf{r})$.

By using equation (7.57), the correlation function of galaxy density contrast measured from the ‘discrete’ samples of galaxies is given by

$$\begin{aligned} \xi_n(\mathbf{r} - \mathbf{r}') &\equiv \langle \delta_n(\mathbf{r})\delta_n(\mathbf{r}') \rangle \\ &= \left\langle \left(\frac{n(\mathbf{r}) - \bar{n}(\mathbf{r})}{\bar{n}(\mathbf{r})} \right) \left(\frac{n(\mathbf{r}') - \bar{n}(\mathbf{r}')}{\bar{n}(\mathbf{r}')} \right) \right\rangle \\ &= \frac{\langle n(\mathbf{r})n(\mathbf{r}') \rangle - \bar{n}(\mathbf{r})\bar{n}(\mathbf{r}')}{\bar{n}(\mathbf{r})\bar{n}(\mathbf{r}')} \\ &= \langle \delta(\mathbf{r})\delta(\mathbf{r}') \rangle + \frac{\delta^D(\mathbf{r} - \mathbf{r}')}{\bar{n}(\mathbf{r})}, \end{aligned} \quad (7.58)$$

where we denote the density contrast of the discrete Poisson sample as

$$\delta_n(\mathbf{r}) \equiv \frac{n(\mathbf{r}) - \bar{n}(\mathbf{r})}{\bar{n}(\mathbf{r})}.$$

The Fourier transform of Equation (7.58) yields following relation in Fourier space:

$$\langle \delta_n(\mathbf{k})\delta_n(\mathbf{k}') \rangle = \langle \delta(\mathbf{k})\delta(\mathbf{k}') \rangle + \int d^3r \frac{1}{\bar{n}(\mathbf{r})} e^{-i\mathbf{r} \cdot (\mathbf{k} + \mathbf{k}')}. \quad (7.59)$$

For N-body simulation, as the mean number density over the simulation box are constant, $\bar{n}(\mathbf{r}) \equiv \bar{n}$; the equation (7.59) reduces to

$$\langle \delta_n(\mathbf{k})\delta_n(\mathbf{k}') \rangle = \langle \delta(\mathbf{k})\delta(\mathbf{k}') \rangle + \frac{(2\pi)^3}{\bar{n}} \delta^D(\mathbf{k}' + \mathbf{k}). \quad (7.60)$$

Therefore, we get the formula for Poisson shot noise.

$$P_n(\mathbf{k}) = P(\mathbf{k}) + \frac{1}{\bar{n}}. \quad (7.61)$$

7.1.4.3 Bispectrum of discrete particles

Similarly, we calculate the bispectrum of discrete particles from the expectation value of

$$\int d^3 r_1 \int d^3 r_2 \int d^3 r_3 g(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) n(\mathbf{r}_1) n(\mathbf{r}_2) n(\mathbf{r}_3)$$

for an arbitrary function $g(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)$. By using the infinitesimal micro-cells, the expectation value becomes

$$\begin{aligned} & \left\langle \int d^3 r_1 \int d^3 r_2 \int d^3 r_3 g(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) n(\mathbf{r}_1) n(\mathbf{r}_2) n(\mathbf{r}_3) \right\rangle \\ &= \int d^3 r_1 \int d^3 r_2 \int d^3 r_3 g(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) \langle n(\mathbf{r}_1) n(\mathbf{r}_2) n(\mathbf{r}_3) \rangle \\ &= \sum_{i,j,k} g(\mathbf{r}_i, \mathbf{r}_j, \mathbf{r}_k) \langle n_i n_j n_k \rangle \\ &= \sum_{i \neq j \neq k} \delta V_i \delta V_j \delta V_k g_{ijk} \bar{n}_i \bar{n}_j \bar{n}_k [1 + \langle \delta_i \delta_j \rangle + \langle \delta_j \delta_k \rangle + \langle \delta_k \delta_i \rangle + \langle \delta_i \delta_j \delta_k \rangle] \\ &\quad + \sum_{i=j \neq k} \delta V_i \delta V_k g_{iik} \bar{n}_i \bar{n}_k [1 + \langle \delta_i \delta_k \rangle] + (2 \text{ cyclic}) + \sum_{i=j=k} \delta V_i g_{iii} \bar{n}_i \\ &= \int d^3 r_1 \int d^3 r_2 \int d^3 r_3 g(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) \\ &\quad \times \left\{ \bar{n}(\mathbf{r}_1) \bar{n}(\mathbf{r}_2) \bar{n}(\mathbf{r}_3) [1 + \langle \delta(\mathbf{r}_1) \delta(\mathbf{r}_2) \delta(\mathbf{r}_3) \rangle + \langle \delta(\mathbf{r}_1) \delta(\mathbf{r}_2) \rangle + (2 \text{ cyclic})] \right. \\ &\quad + \bar{n}(\mathbf{r}_1) \bar{n}(\mathbf{r}_2) [1 + \langle \delta(\mathbf{r}_1) \delta(\mathbf{r}_2) \rangle] \delta^D(\mathbf{r}_1 - \mathbf{r}_3) + (2 \text{ cyclic}) \\ &\quad \left. + \bar{n}(\mathbf{r}_1) \delta^D(\mathbf{r}_1 - \mathbf{r}_2) \delta^D(\mathbf{r}_1 - \mathbf{r}_3) \right\}, \end{aligned} \quad (7.62)$$

where we use subscript to mark the coordinate in the fourth and fifth line. For example, $g_{ijk} \equiv g(\mathbf{r}_i, \mathbf{r}_j, \mathbf{r}_k)$. Again, as the equation (7.62) has to hold for arbitrary $g(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)$, we find

$$\begin{aligned} & \langle n(\mathbf{r}_1) n(\mathbf{r}_2) n(\mathbf{r}_3) \rangle \\ &= \bar{n}(\mathbf{r}_1) \bar{n}(\mathbf{r}_2) \bar{n}(\mathbf{r}_3) [1 + \langle \delta(\mathbf{r}_1) \delta(\mathbf{r}_2) \delta(\mathbf{r}_3) \rangle + \langle \delta(\mathbf{r}_1) \delta(\mathbf{r}_2) \rangle + (2 \text{ cyclic})] \\ &\quad + \bar{n}(\mathbf{r}_1) \bar{n}(\mathbf{r}_2) [1 + \langle \delta(\mathbf{r}_1) \delta(\mathbf{r}_2) \rangle] \delta^D(\mathbf{r}_1 - \mathbf{r}_3) + (2 \text{ cyclic}) \\ &\quad + \bar{n}(\mathbf{r}_1) \delta^D(\mathbf{r}_1 - \mathbf{r}_2) \delta^D(\mathbf{r}_1 - \mathbf{r}_3). \end{aligned} \quad (7.63)$$

We calculate the three-point correlation function of discrete particles by

$$\begin{aligned} \langle \delta_n(\mathbf{r}_1) \delta_n(\mathbf{r}_2) \delta_n(\mathbf{r}_3) \rangle &= \left\langle \left(\frac{n(\mathbf{r}_1)}{\bar{n}(\mathbf{r}_1)} - 1 \right) \left(\frac{n(\mathbf{r}_2)}{\bar{n}(\mathbf{r}_2)} - 1 \right) \left(\frac{n(\mathbf{r}_3)}{\bar{n}(\mathbf{r}_3)} - 1 \right) \right\rangle \\ &= \frac{\langle n(\mathbf{r}_1) n(\mathbf{r}_2) n(\mathbf{r}_3) \rangle}{\bar{n}(\mathbf{r}_1) \bar{n}(\mathbf{r}_2) \bar{n}(\mathbf{r}_3)} - \frac{\langle n(\mathbf{r}_1) n(\mathbf{r}_2) \rangle}{\bar{n}(\mathbf{r}_1) \bar{n}(\mathbf{r}_2)} + (2 \text{ cyclic}) + 2. \end{aligned} \quad (7.64)$$

Using equation (7.57) and equation (7.63), we find

$$\begin{aligned}
& \langle \delta_n(\mathbf{r}_1) \delta_n(\mathbf{r}_2) \delta_n(\mathbf{r}_3) \rangle \\
&= \langle \delta(\mathbf{r}_1) \delta(\mathbf{r}_2) \delta(\mathbf{r}_3) \rangle + \frac{\delta^D(\mathbf{r}_1 - \mathbf{r}_3)}{\bar{n}(\mathbf{r}_3)} \langle \delta(\mathbf{r}_1) \delta(\mathbf{r}_2) \rangle + (2 \text{ cyclic}) \\
&+ \frac{1}{\bar{n}(\mathbf{r}_2) \bar{n}(\mathbf{r}_3)} \delta^D(\mathbf{r}_1 - \mathbf{r}_2) \delta^D(\mathbf{r}_1 - \mathbf{r}_3), \tag{7.65}
\end{aligned}$$

and the Fourier transform of equation (7.65) yields,

$$\begin{aligned}
\langle \delta_n(\mathbf{k}_1) \delta_n(\mathbf{k}_2) \delta_n(\mathbf{k}_3) \rangle &= \langle \delta(\mathbf{k}_1) \delta(\mathbf{k}_2) \delta(\mathbf{k}_3) \rangle \\
&+ \int d^3 r_1 \int d^3 r_2 \frac{1}{\bar{n}(\mathbf{r}_1)} \langle \delta(\mathbf{r}_1) \delta(\mathbf{r}_2) \rangle e^{-i\mathbf{r}_1 \cdot (\mathbf{k}_1 + \mathbf{k}_3) - i\mathbf{r}_2 \cdot \mathbf{k}_2} \\
&+ \int d^3 r_1 \frac{1}{\bar{n}^2(\mathbf{r}_1)} e^{-i\mathbf{r}_1 \cdot \mathbf{k}_{123}}. \tag{7.66}
\end{aligned}$$

Especially, when the mean number density does not vary in time, the three point function in k space becomes

$$\begin{aligned}
& \langle \delta_n(\mathbf{k}_1) \delta_n(\mathbf{k}_2) \delta_n(\mathbf{k}_3) \rangle \\
&= \langle \delta(\mathbf{k}_1) \delta(\mathbf{k}_2) \delta(\mathbf{k}_3) \rangle + \frac{1}{\bar{n}} \langle \delta(\mathbf{k}_{13}) \delta(\mathbf{k}_2) \rangle + (2 \text{ cyclic}) + \frac{(2\pi)^3}{\bar{n}^2} \delta^D(\mathbf{k}_{123}), \tag{7.67}
\end{aligned}$$

and, the bispectrum of discrete Poisson samples is reduced to

$$B_n(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = B(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) + \frac{1}{\bar{n}} \left\{ P(k_1) + P(k_2) + P(k_3) \right\} + \frac{1}{\bar{n}^2}. \tag{7.68}$$

7.2 Power spectrum from Galaxy surveys

For the real galaxy survey, we have to take into account the spatial variance of the mean number density. In that case, we have to weight galaxies differently depending on the mean number density in order to optimize the variance of power spectrum. The most popular weighting is given from Feldman et al. (1994, FKP). In this section, we first review the FKP estimator, and derive the optimal weighting function for power spectrum measurement. Then, we show the practical implementation of FKP estimator for galaxy surveys by using Discrete Fourier Transformation.

7.2.1 The FKP estimator

7.2.1.1 The power spectrum with weighting function $w(\mathbf{r})$

Let us denote the weighted overdensity $F(\mathbf{r})$ as a ‘weighted’ overdensity of a galaxies at position \mathbf{r} in the survey:

$$F(\mathbf{r}) \equiv w(\mathbf{r}) [n(\mathbf{r}) - \bar{n}(\mathbf{r})] = w(\mathbf{r}) \bar{n}(\mathbf{r}) \delta_n(\mathbf{r}) \equiv W(\mathbf{r}) \delta_n(\mathbf{r}). \quad (7.69)$$

Here, $\bar{n}(\mathbf{r})$ is the mean number density of galaxies expected at position \mathbf{r} , and $w(\mathbf{r})$ is a weighting function which optimizes the variance of the estimated power spectrum. Combining the two effect of the selection function and the weighting defines the window function, $W(\mathbf{r})$.

By convolution theorem, the Fourier mode of the weighted density field becomes

$$F(\mathbf{k}) = \int \frac{d^3 q}{(2\pi)^3} W(\mathbf{k} - \mathbf{q}) \delta_n(\mathbf{q}). \quad (7.70)$$

The two point correlation function of the weighted density field is

$$\begin{aligned} \langle F(\mathbf{k}) F(\mathbf{k}') \rangle &= \int \frac{d^3 q}{(2\pi)^3} \int \frac{d^3 q'}{(2\pi)^3} W(\mathbf{k} - \mathbf{q}) W(\mathbf{k}' - \mathbf{q}') \langle \delta_n(\mathbf{q}) \delta_n(\mathbf{q}') \rangle \\ &= \int \frac{d^3 q}{(2\pi)^3} W(\mathbf{k} - \mathbf{q}) W(\mathbf{k}' + \mathbf{q}) P(\mathbf{q}) \\ &\quad + \int \frac{d^3 q}{(2\pi)^3} \int \frac{d^3 q'}{(2\pi)^3} W(\mathbf{k} - \mathbf{q}) W(\mathbf{k}' - \mathbf{q}') \int d^3 r \frac{1}{\bar{n}(\mathbf{r})} e^{-i\mathbf{r} \cdot (\mathbf{q} + \mathbf{q}')}. \end{aligned}$$

The second term is further simplifies to

$$\begin{aligned} &\int \frac{d^3 q}{(2\pi)^3} \int \frac{d^3 q'}{(2\pi)^3} W(\mathbf{k} - \mathbf{q}) W(\mathbf{k}' - \mathbf{q}') \int d^3 r \frac{1}{\bar{n}(\mathbf{r})} e^{-i\mathbf{r} \cdot (\mathbf{q} + \mathbf{q}')} \\ &= \int d^3 r W^2(\mathbf{r}) \frac{1}{\bar{n}(\mathbf{r})} e^{-i\mathbf{r} \cdot (\mathbf{k} + \mathbf{k}')} = \int d^3 r w^2(\mathbf{r}) \bar{n}(\mathbf{r}) e^{-i\mathbf{r} \cdot (\mathbf{k} + \mathbf{k}')}. \end{aligned}$$

Therefore, the power spectrum of the weighted overdensity can be calculated from

$$\langle F(\mathbf{k}) F(\mathbf{k}') \rangle = \int \frac{d^3 q}{(2\pi)^3} W(\mathbf{k} - \mathbf{q}) W(\mathbf{k}' + \mathbf{q}) P(\mathbf{q}) + \int d^3 r w^2(\mathbf{r}) \bar{n}(\mathbf{r}) e^{-i\mathbf{r} \cdot (\mathbf{k} + \mathbf{k}')}, \quad (7.71)$$

that is,

$$\langle F(\mathbf{k}) F(-\mathbf{k}) \rangle = \int \frac{d^3 q}{(2\pi)^3} |W(\mathbf{k} - \mathbf{q})|^2 P(\mathbf{q}) + \int d^3 r w^2(\mathbf{r}) \bar{n}(\mathbf{r}). \quad (7.72)$$

In Feldman, Kaiser & Peacock (1994, FKP), they first subtract the synthetic catalog, $n_s(\mathbf{r})$, generated by the number density of $\bar{n}(\mathbf{r})/\alpha$. Also, they divide the ‘magnitude’

of the window function

$$W \equiv \left[\int d^3r W^2(\mathbf{r}) \right]^{1/2} = \left[\int \frac{d^3k}{(2\pi)^3} |W(\mathbf{k})|^2 \right]^{1/2}$$

when defining the ‘weighted’ density field. That is, the weighted density field in FKP is

$$F(\mathbf{r}) \equiv \frac{w(\mathbf{r}) [n(\mathbf{r}) - \alpha n_s(\mathbf{r})]}{W} = \frac{W(\mathbf{r})}{W} [\delta_n(\mathbf{r}) - \delta_s(\mathbf{r})], \quad (7.73)$$

where

$$\delta_s(\mathbf{r}) \equiv \frac{n_s(\mathbf{r}) - \bar{n}(\mathbf{r})/\alpha}{\bar{n}(\mathbf{r})/\alpha}.$$

Assuming that the synthetic catalog is completely random, and independent of the galaxy distribution, i.e.,

$$\langle \delta_n(\mathbf{r}) \delta_s(\mathbf{r}') \rangle = 0 \quad (7.74)$$

$$\langle \delta_s(\mathbf{r}) \delta_s(\mathbf{r}') \rangle = \alpha \frac{\delta^D(\mathbf{r} - \mathbf{r}')}{\bar{n}(\mathbf{r})}, \quad (7.75)$$

the power spectrum of the weighted density field of FKP is

$$\tilde{P}(\mathbf{k}) \equiv \langle |F(\mathbf{k})|^2 \rangle = \int \frac{d^3q}{(2\pi)^3} \frac{|W(\mathbf{k} - \mathbf{q})|^2}{W^2} P(\mathbf{q}) + \frac{1 + \alpha}{W^2} \int d^3r w^2(\mathbf{r}) \bar{n}(\mathbf{r}). \quad (7.76)$$

Note that in FKP, they refer $W(\mathbf{k})/W$ as $G(\mathbf{k})$:

$$G(\mathbf{k}) \equiv \frac{\int d^3r \bar{n}(\mathbf{r}) w(\mathbf{r}) e^{-i\mathbf{k}\cdot\mathbf{r}}}{[\int d^3r \bar{n}^2(\mathbf{r}) w^2(\mathbf{r})]^{1/2}}.$$

If survey has a typical size of $D \simeq V^{1/3}$, then the window function $W(\mathbf{k})/W$ will be a rather compact function with width $\delta k \sim 1/D$. Therefore, for high enough wavenumber, $|\mathbf{k}| \gg 1/D$, we can approximate the power spectrum of weighted overdensity as

$$\tilde{P}(\mathbf{k}) \simeq P(\mathbf{k}) + P_{\text{shot}}$$

provided that $P(\mathbf{k})$ is a smooth function of \mathbf{k} . Here,

$$P_{\text{shot}} \equiv \frac{(1 + \alpha) \int d^3r \bar{n}(\mathbf{r}) w^2(\mathbf{r})}{\int d^3r \bar{n}^2(\mathbf{r}) w^2(\mathbf{r})}.$$

is the shot noise power spectrum in the presence of the weighting function. We estimate the angular averaged (monopole) power spectrum estimation by

$$\hat{P}(k) \equiv \frac{1}{V_k} \int_{V_k} d^3q [\tilde{P}(\mathbf{q}) - P_{\text{shot}}], \quad (7.77)$$

where V_k is the volume of the shell in Fourier space.

7.2.2 The variance of the power spectrum

We estimate the variance of the power spectrum by

$$\sigma_p^2 \equiv \left\langle \left[\hat{P}(k) - P(k) \right]^2 \right\rangle = \frac{1}{V_k^2} \int_{V_k} d^3q \int_{V_k} d^3q' \left\langle \delta \hat{P}(\mathbf{q}) \delta \hat{P}(\mathbf{q}') \right\rangle \quad (7.78)$$

with $\delta \hat{P}(\mathbf{q}) = \hat{P}(\mathbf{q}) - P(\mathbf{q})$. If $F(\mathbf{k})$ obeys Gaussian statistics, then

$$\left\langle \delta \hat{P}(\mathbf{q}) \delta \hat{P}(\mathbf{q}') \right\rangle = |\langle F(\mathbf{q}) F^*(\mathbf{q}') \rangle|^2$$

With Equation (7.71), we write the right hand side as

$$\begin{aligned} & \langle F(\mathbf{q}) F^*(\mathbf{q}') \rangle \\ &= \int \frac{d^3p}{(2\pi)^3} \frac{W(\mathbf{q}-\mathbf{p}) W^*(\mathbf{q}'-\mathbf{p})}{W^2} P(\mathbf{p}) + \frac{1+\alpha}{W^2} \int d^3r w^2(\mathbf{r}) \bar{n}(\mathbf{r}) e^{-i\mathbf{r} \cdot (\mathbf{q}-\mathbf{q}')}. \end{aligned} \quad (7.79)$$

In FKP paper, the second term is referred as $S(\mathbf{k})$:

$$S(\mathbf{k}) \equiv \frac{(1+\alpha) \int d^3r w^2(\mathbf{r}) \bar{n}(\mathbf{r}) e^{-i\mathbf{k} \cdot \mathbf{r}}}{\int d^3r w^2(\mathbf{r}) \bar{n}^2(\mathbf{r})}$$

As Equation (7.78) integrate over the same spherical shell, we consider the case when $\mathbf{q}-\mathbf{q}' = \delta \mathbf{q}$. For $|\mathbf{q}| \gg 1/D$, we approximate Equation (7.79) as

$$\begin{aligned} & \langle F(\mathbf{q}) F^*(\mathbf{q}') \rangle \\ & \simeq \frac{P(\mathbf{q})}{W^2} \int \frac{d^3p}{(2\pi)^3} W(\mathbf{q}-\mathbf{p}) W^*(\mathbf{q}'-\mathbf{p}) + \frac{1+\alpha}{W^2} \int d^3r w^2(\mathbf{r}) \bar{n}(\mathbf{r}) e^{-i\mathbf{r} \cdot (\mathbf{q}-\mathbf{q}')}, \end{aligned} \quad (7.80)$$

where the first term becomes

$$\begin{aligned} & \int \frac{d^3p}{(2\pi)^3} W(\mathbf{q}-\mathbf{p}) W^*(\mathbf{q}'-\mathbf{p}) \\ &= \int \frac{d^3p}{(2\pi)^3} \int d^3r \int d^3r' W(\mathbf{r}) W(\mathbf{r}') e^{-i\mathbf{r} \cdot (\mathbf{q}-\mathbf{p})} e^{i\mathbf{r}' \cdot (\mathbf{q}'-\mathbf{p})} \\ &= \int d^3r \int d^3r' W(\mathbf{r}) W(\mathbf{r}') \delta^D(\mathbf{r}-\mathbf{r}') e^{-i\mathbf{r} \cdot \mathbf{q}} e^{i\mathbf{r}' \cdot \mathbf{q}'} \\ &= \int d^3r W^2(\mathbf{r}) e^{-i\mathbf{r} \cdot (\mathbf{q}-\mathbf{q}')}. \end{aligned}$$

In FKP, this function is called $Q(\mathbf{k})$:

$$Q(\mathbf{k}) \equiv \frac{\int d^3r w^2(\mathbf{r}) \bar{n}^2(\mathbf{r}) e^{-i\mathbf{k} \cdot \mathbf{r}}}{\int d^3r w^2(\mathbf{r}) \bar{n}^2(\mathbf{r})}. \quad (7.81)$$

Therefore, the variance of the power spectrum is

$$\begin{aligned} \left\langle \delta \hat{P}(\mathbf{q}) \delta \hat{P}(\mathbf{q}') \right\rangle &= |\langle F(\mathbf{q}) F^*(\mathbf{q}') \rangle|^2 \\ &= |P(\mathbf{q}) Q(\delta \mathbf{q}) + S(\delta \mathbf{q})|^2. \end{aligned} \quad (7.82)$$

7.2.3 Optimal weighting

When the spherical shell is larger than the coherence length ($\sim 1/D$), the double integration in Equation (7.78) reduces to

$$\sigma_p^2(k) = \frac{1}{V_k} \int d^3q |P(k)Q(\mathbf{q}) + S(\mathbf{q})|^2.$$

Therefore, the fractional variance of the power is

$$\begin{aligned} \frac{\sigma_p^2(k)}{P^2(k)} &= \frac{1}{V_k} \int d^3q \left| Q(\mathbf{q}) + \frac{S(\mathbf{q})}{P(k)} \right|^2 \\ &= \frac{1}{V_k W^2} \int d^3q \left| \int d^3r w^2(\mathbf{r}) \bar{n}^2(\mathbf{r}) \left(1 + \frac{1+\alpha}{\bar{n}(\mathbf{r})P(k)} \right) e^{-i\mathbf{q}\cdot\mathbf{r}} \right|^2. \end{aligned} \quad (7.83)$$

Using Parseval's theorem³, the equation is further simplified as

$$\frac{\sigma_p^2(k)}{P^2(k)} = \frac{(2\pi)^3}{V_k W^2} \int d^3r w^4(\mathbf{r}) \bar{n}^4(\mathbf{r}) \left(1 + \frac{1+\alpha}{\bar{n}(\mathbf{r})P(k)} \right)^2. \quad (7.84)$$

For optimal weighting, we choose the weighting function $w(\mathbf{r})$ which minimizes the variance in Equation (7.84). First, let's abbreviate the equation as

$$\frac{\sigma_p^2(k)}{P^2(k)} = \frac{\int d^3r w^4(\mathbf{r}) f(\mathbf{r})}{[\int d^3r w^2(\mathbf{r}) g(\mathbf{r})]^2}$$

That is,

$$\begin{aligned} f(\mathbf{r}) &= \bar{n}^4(\mathbf{r}) \left(1 + \frac{1+\alpha}{\bar{n}(\mathbf{r})P(k)} \right)^2 \\ g(\mathbf{r}) &= \bar{n}^2(\mathbf{r}). \end{aligned}$$

When we take the variation of $w(\mathbf{r}) = w_0(\mathbf{r}) + \delta w(\mathbf{r})$, Equation (7.84) becomes

$$\begin{aligned} \frac{\sigma_p^2(k)}{P^2(k)} &= \frac{\int d^3r w_0^4 [1 + 4\delta w/w_0] f}{[\int d^3r w_0^2 (1 + 3\delta w/w_0) g]^2} \\ &= \frac{\int d^3r w_0^4 f}{[\int d^3r w_0^2 g]^2} \left\{ 1 + 4 \left(\frac{\int d^3r w_0^3 \delta w f}{\int d^3r w_0^4 f} - \frac{\int d^3r w_0 \delta w g}{\int d^3r w_0^2 g} \right) \right\} + \dots \end{aligned} \quad (7.85)$$

³The proof of Parseval's theorem is given as following.

$$\begin{aligned} \int d^3q \left| \int d^3r f(\mathbf{r}) e^{-i\mathbf{q}\cdot\mathbf{r}} \right|^2 &= \int d^3q \int d^3r f(\mathbf{r}) e^{-i\mathbf{q}\cdot\mathbf{r}} \int d^3r' f(\mathbf{r}') e^{i\mathbf{q}\cdot\mathbf{r}'} \\ &= (2\pi)^3 \int d^3r \int d^3r' f(\mathbf{r}) f(\mathbf{r}') \delta^D(\mathbf{r} - \mathbf{r}') = (2\pi)^3 \int d^3r f^2(\mathbf{r}) \end{aligned}$$

Therefore, the optimal weighting function has to satisfy

$$\frac{\int d^3r w_0^3 \delta w f}{\int d^3r w_0^4 f} = \frac{\int d^3r w_0 \delta w g}{\int d^3r w_0^2 g}, \quad (7.86)$$

whose solution is given by

$$w_0^2 \propto g/f. \quad (7.87)$$

In terms of the mean number density and the power spectrum, the optimal weighting has to satisfy

$$w(\mathbf{r}) \propto \frac{1}{\bar{n}(\mathbf{r}) + (1 + \alpha)/P(k)}. \quad (7.88)$$

Finally, the dimensionless optimal weighting function is

$$w(\mathbf{r}) = \frac{1}{1 + \alpha + P(k)\bar{n}(\mathbf{r})}. \quad (7.89)$$

When we choose a large number of synthetic sample ($\alpha \ll 1$), we recover the result of the FKP optimal weighting function:

$$w(\mathbf{r}) = \frac{1}{1 + P(k)\bar{n}(\mathbf{r})}. \quad (7.90)$$

7.3 Implementing the FKP estimator

In this section, we show the Discrete Fourier Transform implementation of the FKP estimator. Let us consider the case when the mean number density depends on the position: $\bar{n}(\mathbf{r})$. In the previous section, we have shown that the optimal weighting function is given by

$$w(\mathbf{r}) = \frac{1}{1 + \bar{n}(\mathbf{r})P(k)}. \quad (7.91)$$

We define the weighted overdensity in the discrete grid point \mathbf{n}_r as

$$\begin{aligned} F(\mathbf{n}_r) &= \frac{w(\mathbf{n}_r)}{W} [N_g(\mathbf{n}_r) - \alpha N_s(\mathbf{n}_r)] = \frac{w(\mathbf{n}_r)\bar{N}(\mathbf{n}_r)}{W} [\delta_n(\mathbf{n}_r) - \delta_s(\mathbf{n}_r)] \\ &= \frac{W(\mathbf{n}_r)H^3}{W} [\delta_n(\mathbf{n}_r) - \delta_s(\mathbf{n}_r)]. \end{aligned} \quad (7.92)$$

Here, $N_i(\mathbf{n}_r)$ denote the number density of galaxies ($i = g$), and synthetic random samples ($i = s$) assigned to the grid \mathbf{n}_r ,

$$N_i(\mathbf{n}_r) \equiv \int_{\mathbf{n}_r} d^3r n_i(\mathbf{r}) W^s(\mathbf{r} - \mathbf{n}_r), \quad (7.93)$$

where $W^s(\mathbf{x})$ is the sampling window function we discussed in Section 7.1.1. Similarly, $\bar{N}(\mathbf{n}_r)$ is the mean number density of galaxies assigned to the grid \mathbf{n}_r :

$$\bar{N}(\mathbf{n}_r) \equiv \int_{\mathbf{n}_r} d^3r \bar{n}(\mathbf{r}) W^s(\mathbf{r} - \mathbf{n}_r). \quad (7.94)$$

As synthetic random samples are generated from the mean number density rescaled by $1/\alpha$, they are related by

$$\bar{N}(\mathbf{n}_r) = \langle N_g(\mathbf{n}_r) \rangle = \alpha \langle N_s(\mathbf{n}_r) \rangle. \quad (7.95)$$

Also we define the survey window function on the grid point \mathbf{n}_r as

$$W(\mathbf{n}_r) \equiv \frac{w(\mathbf{n}_r) \bar{N}(\mathbf{n}_r)}{H^3}. \quad (7.96)$$

Fourier transform the weighted overdensity above yields⁴

$$\begin{aligned} F_{DFT}(\mathbf{n}_k) \\ = \sum_{\mathbf{n}_r} F(\mathbf{n}_r) e^{-i2\pi \mathbf{n}_k \cdot \mathbf{n}_r / N} = \frac{H^3}{W} \sum_{\mathbf{n}_r} W(\mathbf{n}_r) [\delta_n(\mathbf{n}_r) - \delta_s(\mathbf{n}_r)] e^{-i2\pi \mathbf{n}_k \cdot \mathbf{n}_r / N}. \end{aligned} \quad (7.97)$$

That is, the Fourier transform is given by the convolution of $W(\mathbf{n}_k)$ and $\delta_n(\mathbf{n}_k) - \delta_s(\mathbf{n}_k)$. Let us be explicit about the convolution in DFT. For discrete sampling of $A(\mathbf{n}_r)$ and $B(\mathbf{n}_r)$, the Fourier transform of its multiplication $C(\mathbf{n}_r) \equiv A(\mathbf{n}_r)B(\mathbf{n}_r)$ is given by

$$\begin{aligned} C_{DFT}(\mathbf{n}_k) \\ &= \sum_{\mathbf{n}_r} A(\mathbf{n}_r) B(\mathbf{n}_r) e^{-i2\pi \mathbf{n}_k \cdot \mathbf{n}_r / N} \\ &= \sum_{\mathbf{n}_r} \left[\frac{1}{V} \sum_{\mathbf{n}_p} A(\mathbf{n}_p) e^{i2\pi \mathbf{n}_p \cdot \mathbf{n}_r / N} \right] \left[\frac{1}{V} \sum_{\mathbf{n}_q} B(\mathbf{n}_q) e^{i2\pi \mathbf{n}_q \cdot \mathbf{n}_r / N} \right] e^{-i2\pi \mathbf{n}_k \cdot \mathbf{n}_r / N} \\ &= \frac{1}{V^2} \sum_{\mathbf{n}_p} \sum_{\mathbf{n}_q} A(\mathbf{n}_p) B(\mathbf{n}_q) \left[\sum_{\mathbf{n}_r} e^{i2\pi (\mathbf{n}_p + \mathbf{n}_q) \cdot \mathbf{n}_r / N} e^{-i2\pi \mathbf{n}_k \cdot \mathbf{n}_r / N} \right] \\ &= \frac{N^3}{V^2} \sum_{\mathbf{n}_p} \sum_{\mathbf{n}_q} A(\mathbf{n}_p) B(\mathbf{n}_q) \delta_{\mathbf{n}_k \pm mN, \mathbf{n}_p + \mathbf{n}_q} \\ &= \frac{N^3}{V^2} \sum_{\mathbf{n}_q} A((\mathbf{n}_k - \mathbf{n}_q)_N) B(\mathbf{n}_q). \end{aligned} \quad (7.98)$$

⁴For the normalization of the Discrete Fourier Transform, see Appendix A.

Here, m can be any integer, and $A((\mathbf{n}_k)_N)$ denotes the integer triplet modulated by N . Using the convolution in DFT, we find the DTF of F becomes

$$F_{DFT}(\mathbf{n}_k) = \frac{1}{WV} \sum_{\mathbf{n}_q} W((\mathbf{n}_k - \mathbf{n}_q)_N) [\delta_n(\mathbf{n}_q) - \delta_s(\mathbf{n}_q)]. \quad (7.99)$$

Note that when $W(\mathbf{n}_r) \equiv w\bar{n}$ (constant), then,

$$W(\mathbf{n}_k) = H^3 \sum_{\mathbf{n}_r} w\bar{n} e^{-i2\pi \mathbf{n}_k \cdot \mathbf{n}_r / N} = H^3 w\bar{n} N^3 \delta_{\mathbf{n}_k, \mathbf{0}}$$

therefore, we reproduce the normalization of the previous section:

$$F_{DFT}(\mathbf{n}_k) = \frac{w\bar{n}}{W} [\delta_n(\mathbf{n}_k) - \delta_s(\mathbf{n}_k)].$$

We estimate the power spectrum by calculating the ensemble average of

$$\begin{aligned} \langle F_{DFT}(\mathbf{n}_k) F_{DFT}(\mathbf{n}_{k'}) \rangle &= \frac{1}{W^2 V^2} \sum_{\mathbf{n}_q} \sum_{\mathbf{n}_{q'}} W((\mathbf{n}_k - \mathbf{n}_q)_N) W((\mathbf{n}_{k'} - \mathbf{n}_{q'})_N) \\ &\quad \times \langle [\delta_n(\mathbf{n}_q) - \delta_s(\mathbf{n}_q)] [\delta_n(\mathbf{n}_{q'}) - \delta_s(\mathbf{n}_{q'})] \rangle. \end{aligned} \quad (7.100)$$

First, let us evaluate the two point correlation function without the window function for the discrete grid

$$\begin{aligned} &\langle \delta_n(\mathbf{n}_q) \delta_n(\mathbf{n}_{q'}) \rangle \\ &= H^6 \sum_{\mathbf{n}_r} \sum_{\mathbf{n}_{r'}} \langle \delta_n(\mathbf{n}_r) \delta_n(\mathbf{n}_{r'}) \rangle e^{-i2\pi \mathbf{n}_r \cdot \mathbf{n}_q / N} e^{-i2\pi \mathbf{n}_{r'} \cdot \mathbf{n}_{q'} / N} \\ &= H^6 \sum_{\mathbf{n}_r} \sum_{\mathbf{n}_{r'}} \left[\langle \delta(\mathbf{n}_r) \delta(\mathbf{n}_{r'}) \rangle + \frac{\delta_{\mathbf{n}_r, \mathbf{n}_{r'}}^K}{H^3 \bar{n}(\mathbf{n}_r)} \right] e^{-i2\pi \mathbf{n}_r \cdot \mathbf{n}_q / N} e^{-i2\pi \mathbf{n}_{r'} \cdot \mathbf{n}_{q'} / N} \\ &= \langle \delta(\mathbf{n}_q) \delta(\mathbf{n}_{q'}) \rangle + H^3 \sum_{\mathbf{n}_r} \frac{1}{\bar{n}(\mathbf{n}_r)} e^{-i2\pi \mathbf{n}_r \cdot (\mathbf{n}_q + \mathbf{n}_{q'}) / N}, \end{aligned} \quad (7.101)$$

where in the second line, we changes the Dirac delta function to the Kronecker delta by explicitly factoring out the dimensionality ($1/H^3$). From the same procedure, we can also calculate

$$\langle \delta_n(\mathbf{n}_q) \delta_s(\mathbf{n}_{q'}) \rangle = 0, \quad (7.102)$$

$$\langle \delta_s(\mathbf{n}_q) \delta_s(\mathbf{n}_{q'}) \rangle = H^3 \sum_{\mathbf{n}_r} \frac{\alpha}{\bar{n}(\mathbf{n}_r)} e^{-i2\pi \mathbf{n}_r \cdot (\mathbf{n}_q + \mathbf{n}_{q'}) / N}. \quad (7.103)$$

Putting the result above all together, we finally have

$$\begin{aligned}
& \langle F_{DFT}(\mathbf{n}_k) F_{DFT}(\mathbf{n}_{k'}) \rangle \\
&= \frac{1}{W^2 V^2} \sum_{\mathbf{n}_q} \sum_{\mathbf{n}_{q'}} W((\mathbf{n}_k - \mathbf{n}_q)_N) W((\mathbf{n}_{k'} - \mathbf{n}_{q'})_N) \langle \delta(\mathbf{n}_q) \delta(\mathbf{n}_{q'}) \rangle \\
&+ H^3 \frac{1+\alpha}{W^2 V^2} \sum_{\mathbf{n}_q} \sum_{\mathbf{n}_{q'}} W((\mathbf{n}_k - \mathbf{n}_q)_N) W((\mathbf{n}_{k'} - \mathbf{n}_{q'})_N) \sum_{\mathbf{n}_r} \frac{1}{\bar{n}(\mathbf{n}_r)} e^{-i2\pi \mathbf{n}_r \cdot (\mathbf{n}_q + \mathbf{n}_{q'})/N}.
\end{aligned} \tag{7.104}$$

The first term is the Window function convolved galaxy power spectrum:

$$\begin{aligned}
& \frac{1}{W^2 V^2} \sum_{\mathbf{n}_q} \sum_{\mathbf{n}_{q'}} W((\mathbf{n}_k - \mathbf{n}_q)_N) W((\mathbf{n}_{k'} - \mathbf{n}_{q'})_N) \frac{(2\pi)^3}{k_{F1} k_{F2} k_{F3}} P(\mathbf{n}_q) \delta_{\mathbf{n}_q, -\mathbf{n}_{q'}}^D \\
&= \frac{1}{W^2 V} \sum_{\mathbf{n}_q} W((\mathbf{n}_k - \mathbf{n}_q)_N) W((\mathbf{n}_{k'} + \mathbf{n}_q)_N) P(\mathbf{n}_q),
\end{aligned} \tag{7.105}$$

and the second term is the shot-noise term:

$$\begin{aligned}
& H^9 \frac{1+\alpha}{W^2 V^2} \sum_{\mathbf{n}_q} \sum_{\mathbf{n}_{q'}} \sum_{\mathbf{n}_{r1}} W(\mathbf{n}_{r1}) e^{-i2\pi \mathbf{n}_{r1} \cdot (\mathbf{n}_k - \mathbf{n}_q)/N} \\
& \times \sum_{\mathbf{n}_{r2}} W(\mathbf{n}_{r2}) e^{-i2\pi \mathbf{n}_{r2} \cdot (\mathbf{n}_{k'} - \mathbf{n}_{q'})/N} \sum_{\mathbf{n}_r} \frac{1}{\bar{n}(\mathbf{n}_r)} e^{-i2\pi \mathbf{n}_r \cdot (\mathbf{n}_q + \mathbf{n}_{q'})/N} \\
&= H^9 \frac{1+\alpha}{W^2 V^2} \sum_{\mathbf{n}_{r1}} \sum_{\mathbf{n}_{r2}} \sum_{\mathbf{n}_r} \frac{W(\mathbf{n}_{r1}) W(\mathbf{n}_{r2})}{\bar{n}(\mathbf{n}_r)} e^{-i2\pi (\mathbf{n}_{r1} \cdot \mathbf{n}_k + \mathbf{n}_{r2} \cdot \mathbf{n}_{k'})/N} N^3 \delta_{\mathbf{n}_{r1}, \mathbf{n}_r}^D N^3 \delta_{\mathbf{n}_{r2}, \mathbf{n}_r}^D \\
&= N^6 H^9 \frac{1+\alpha}{W^2 V^2} \sum_{\mathbf{n}_r} \frac{W^2(\mathbf{n}_r)}{\bar{n}(\mathbf{n}_r)} e^{-i2\pi \mathbf{n}_r \cdot (\mathbf{n}_k + \mathbf{n}_{k'})/N} \\
&= H^3 \frac{1+\alpha}{W^2} \sum_{\mathbf{n}_r} \frac{W^2(\mathbf{n}_r)}{\bar{n}(\mathbf{n}_r)} e^{-i2\pi \mathbf{n}_r \cdot (\mathbf{n}_k + \mathbf{n}_{k'})/N}.
\end{aligned} \tag{7.106}$$

Adding up the results, we find that

$$\langle |F_{DFT}(\mathbf{n}_k)|^2 \rangle = \frac{1}{W^2 V} \sum_{\mathbf{n}_q} |W((\mathbf{n}_k - \mathbf{n}_q)_N)|^2 P(\mathbf{n}_q) + H^3 \frac{1+\alpha}{W^2} \sum_{\mathbf{n}_r} \frac{W^2(\mathbf{n}_r)}{\bar{n}(\mathbf{n}_r)}. \tag{7.107}$$

The normalization factor W can be calculated as

$$W \equiv \left[\int d^3 r W^2(\mathbf{r}) \right]^{1/2} = \left[\sum_{\mathbf{n}_r} \frac{w^2(\mathbf{n}_r) \bar{N}^2(\mathbf{n}_r)}{H^3} \right]^{1/2}. \tag{7.108}$$

Let us check the limiting case when $\bar{n}(\mathbf{n}_r) = \bar{n}$. For that case, the weighting function

is also a constant, $w(\mathbf{n}_r) = w$. Then, the normalization factor becomes

$$W = \left[\sum_{\mathbf{n}_r} w^2 \bar{n}^2 H^3 \right]^{1/2} = [N^3 w^2 \bar{n}^2 H^3]^{1/2} = \sqrt{V} w \bar{n}, \quad (7.109)$$

and power spectrum can be estimated by

$$\begin{aligned} & \left\langle |F_{DFT}(\mathbf{n}_k)|^2 \right\rangle \\ &= \frac{1}{V^2 w^2 \bar{n}^2} \sum_{\mathbf{n}_q} V^2 w^2 \bar{n}^2 \delta_{\mathbf{n}_k, \mathbf{n}_q} P(\mathbf{n}_q) + H^3 \frac{1 + \alpha}{V w^2 \bar{n}^2} \sum_{\mathbf{n}_r} \frac{w^2 \bar{n}^2}{\bar{n}} \\ &= P(\mathbf{n}_k) + \frac{1 + \alpha}{\bar{n}}, \end{aligned} \quad (7.110)$$

which are what we expected from the calculation of the previous section.

Finally, we have to correct for the window function due to the number density distribution, by following the method described in Section 7.1.3.

7.3.1 The estimator

In summary, we estimate the power spectrum as follows.

7.3.1.1 Constant weighting

When we do not employ the weighting function, first calculate

$$F(\mathbf{n}_r) = N_g(\mathbf{n}_r) - \alpha N_s(\mathbf{n}_r) \quad (7.111)$$

and Fourier transform it. Then, the square of the Fourier transform becomes

$$\begin{aligned} & \frac{1}{W^2} \left\langle |F_{DFT}(\mathbf{n}_k)|^2 \right\rangle \\ &= \frac{1}{W^2 V} \sum_{\mathbf{n}_q} |W(\mathbf{n}_k - \mathbf{n}_q)|^2 \hat{P}(\mathbf{n}_q) + H^3 \frac{1 + \alpha}{W^2} \sum_{\mathbf{n}_r} \frac{W^2(\mathbf{n}_r)}{\bar{n}(\mathbf{n}_r)}, \end{aligned} \quad (7.112)$$

where the window function is given by

$$W(\mathbf{n}_r) = \frac{\bar{N}(\mathbf{n}_r)}{H^3} \simeq \frac{\alpha N_s(\mathbf{n}_r)}{H^3}, \quad (7.113)$$

and its Fourier transform is

$$W(\mathbf{n}_k) = H^3 \sum_{\mathbf{n}_r} W(\mathbf{n}_r) e^{-i2\pi \mathbf{n}_k \cdot \mathbf{n}_r / N}. \quad (7.114)$$

As we distribute particle numbers to regular grid points, $F_{DFT}(\mathbf{n}_k)$ has to be deconvolved by the method we described in 7.1.3.

The normalization factor W^2 is given by

$$W^2 = \frac{1}{H^3} \sum_{\mathbf{n}_r} \bar{N}^2(\mathbf{n}_r) \simeq \frac{\alpha^2}{H^3} \sum_{\mathbf{n}_r} N_s^2(\mathbf{n}_r). \quad (7.115)$$

Note that we can approximate the average number of galaxies $\bar{N}(\mathbf{n}_r)$ by

$$\bar{N}(\mathbf{n}_r) \equiv \int_{H^3} d^3r \bar{n}(r) \simeq \alpha N_s(\mathbf{n}_r). \quad (7.116)$$

With this approximation, and $\bar{n}(\mathbf{n}_r) \simeq \alpha N_s(\mathbf{n}_r)/H^3$ the shot noise term may be approximated as

$$H^3 \frac{1+\alpha}{W^2} \sum_{\mathbf{n}_r} \frac{W^2(\mathbf{n}_r)}{\bar{n}(\mathbf{n}_r)} \simeq H^3 \left(\frac{1+\alpha}{\alpha} \right) \frac{\sum N_s(\mathbf{n}_r)}{\sum N_s^2(\mathbf{n}_r)}. \quad (7.117)$$

7.3.1.2 FKP optimal weighting

When estimating power spectrum with a optimal weighting function

$$w(\mathbf{n}_r) = \frac{1}{1 + \bar{n}(\mathbf{r})P(k)} \simeq \frac{1}{1 + \alpha N_s(\mathbf{r})P(k)/H^3} = \frac{H^3}{H^3 + \alpha N_s(\mathbf{r})P(k)}, \quad (7.118)$$

first calculate

$$F(\mathbf{n}_r) = w(\mathbf{n}_r) [N_g(\mathbf{n}_r) - \alpha N_s(\mathbf{n}_r)] \quad (7.119)$$

and Fourier transform it. Then, the square of the Fourier transform becomes

$$\begin{aligned} & \frac{1}{W^2} \left\langle |F_{DFT}(\mathbf{n}_k)|^2 \right\rangle \\ &= \frac{1}{W^2 V} \sum_{\mathbf{n}_q} |W(\mathbf{n}_k - \mathbf{n}_q)|^2 P(\mathbf{n}_q) + H^3 \frac{1+\alpha}{W^2} \sum_{\mathbf{n}_r} \frac{W^2(\mathbf{n}_r)}{\bar{n}(\mathbf{n}_r)}, \end{aligned} \quad (7.120)$$

where the window function is given by

$$W(\mathbf{n}_r) = \frac{w(\mathbf{n}_r)\bar{N}(\mathbf{n}_r)}{H^3} \simeq \frac{\alpha w(\mathbf{n}_r)N_s(\mathbf{n}_r)}{H^3} = \frac{\alpha N_s(\mathbf{n}_r)}{H^3 + \alpha N_s(\mathbf{n}_r)P(k)}, \quad (7.121)$$

and its Fourier transform is

$$W(\mathbf{n}_k) = H^3 \sum_{\mathbf{n}_r} W(\mathbf{n}_r) e^{-i2\pi \mathbf{n}_k \cdot \mathbf{n}_r / N}. \quad (7.122)$$

Then, according to the number distribution scheme, $F_{DFT}(\mathbf{n}_k)$ has to be deconvolved by the method we described in 7.1.3. The normalization factor W^2 is

$$W^2 = \frac{1}{H^3} \sum_{\mathbf{n}_r} w^2(\mathbf{n}_r) \bar{N}^2(\mathbf{n}_r) \simeq H^3 \sum_{\mathbf{n}_r} \left[\frac{\alpha N_s(\mathbf{n}_r)}{H^3 + \alpha N_s(\mathbf{n}_r) P(k)} \right]^2, \quad (7.123)$$

and the shot noise term may be approximated as

$$\begin{aligned} & H^3 \frac{1 + \alpha}{W^2} \sum_{\mathbf{n}_r} \frac{W^2(\mathbf{n}_r)}{\bar{n}(\mathbf{n}_r)} \\ & \simeq H^3 \left(\frac{1 + \alpha}{\alpha} \right) \left(\sum_{\mathbf{n}_r} \frac{N_s(\mathbf{n}_r)}{(H^3 + \alpha N_s(\mathbf{n}_r) P(k))^2} \right) \left(\sum_{\mathbf{n}_r} \frac{N_s^2(\mathbf{n}_r)}{(H^3 + \alpha N_s(\mathbf{n}_r) P(k))^2} \right)^{-1}. \end{aligned} \quad (7.124)$$

Appendices

Appendix A

Fourier transforms

In this Appendix, we present the Fourier convention we adopt throughout this dissertation. We define the Fourier transform, Fourier series and Discrete Fourier Transform in a consistence manner.

A.1 Continuous Fourier transform

In our convention, the Fourier transform of function $f(\mathbf{x})$ is defined as

$$f(\mathbf{k}) = \int d^3x f(\mathbf{x}) e^{-i\mathbf{k}\cdot\mathbf{x}}, \quad (\text{A.1})$$

and the inverse Fourier transform is defined as

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

where the integration extended to the entire space. Note that, unless it causes come confusion, we use the same notation for the functions in real space and Fourier space, as we consider them as different *representations* of a function defined in the Hilbert space. We shall distinguish them by the explicit argument, but, when this notation can cause confusion, we shall explicitly indicate the Fourier transform by FT.

A.1.1 Delta function

The Fourier transform [Eq. (A.1)] of a function $f(\mathbf{x})$ followed by the inverse Fourier transform [Eq. (A.2)] has to be identical to the original function:

$$\begin{aligned} f(\mathbf{x}) &= \int_{-\infty}^{\infty} \frac{d^3k}{(2\pi)^3} \left(\int_{-\infty}^{\infty} d^3x' f(\mathbf{x}') e^{-i\mathbf{k}\cdot\mathbf{x}'} \right) e^{i\mathbf{k}\cdot\mathbf{x}} \\ &= \int_{-\infty}^{\infty} d^3x' \left(\int_{-\infty}^{\infty} \frac{d^3k}{(2\pi)^3} e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')} \right) f(\mathbf{x}'), \end{aligned} \quad (\text{A.3})$$

or

$$\begin{aligned} f(\mathbf{k}) &= \int_{-\infty}^{\infty} d^3x \left(\int_{-\infty}^{\infty} \frac{d^3k'}{(2\pi)^3} f(\mathbf{k}') e^{i\mathbf{k}' \cdot \mathbf{x}} \right) e^{-i\mathbf{k} \cdot \mathbf{x}} \\ &= \int_{-\infty}^{\infty} d^3k' \left(\int_{-\infty}^{\infty} \frac{d^3x}{(2\pi)^3} e^{i\mathbf{x} \cdot (\mathbf{k}' - \mathbf{k})} \right) f(\mathbf{k}'). \end{aligned} \quad (\text{A.4})$$

The relations above motivate us to define Dirac delta function as

$$\delta^D(\mathbf{x} - \mathbf{x}') \equiv \int_{-\infty}^{\infty} \frac{d^3k}{(2\pi)^3} e^{i\mathbf{k} \cdot (\mathbf{x} - \mathbf{x}')}, \quad (\text{A.5})$$

which satisfies

$$f(\mathbf{x}) = \int_{-\infty}^{\infty} d^3x' \delta^D(\mathbf{x} - \mathbf{x}') f(\mathbf{x}'). \quad (\text{A.6})$$

A.1.2 Convolution theorem

In this convention, the Fourier transformation of convolution of two functions $f(\mathbf{x})$ and $g(\mathbf{x})$, which is defined as

$$h(\mathbf{x}) = [f \star g](\mathbf{x}) \equiv \int d^3x_1 \int d^3x_2 f(\mathbf{x}_1) g(\mathbf{x}_2) \delta^D(\mathbf{x} - \mathbf{x}_1 - \mathbf{x}_2), \quad (\text{A.7})$$

can be calculated as

$$\begin{aligned} h(\mathbf{k}) &= \int d^3x \left[\int d^3x_1 \int d^3x_2 f(\mathbf{x}_1) g(\mathbf{x}_2) \delta^D(\mathbf{x} - \mathbf{x}_1 - \mathbf{x}_2) \right] e^{-i\mathbf{k} \cdot \mathbf{x}} \\ &= \int d^3x \left[\int d^3x_1 \int d^3x_2 f(\mathbf{x}_1) g(\mathbf{x}_2) \int \frac{d^3k'}{(2\pi)^3} e^{i\mathbf{k}' \cdot (\mathbf{x} - \mathbf{x}_1 - \mathbf{x}_2)} \right] e^{-i\mathbf{k} \cdot \mathbf{x}} \\ &= \int \frac{d^3k'}{(2\pi)^3} \int d^3x e^{i\mathbf{x} \cdot (\mathbf{k}' - \mathbf{k})} \left[\int d^3x_1 f(\mathbf{x}_1) e^{-i\mathbf{k}' \cdot \mathbf{x}_1} \right] \left[\int d^3x_2 g(\mathbf{x}_2) e^{-i\mathbf{k}' \cdot \mathbf{x}_2} \right] \\ &= \int d^3k' \delta^D(\mathbf{k}' - \mathbf{k}) f(\mathbf{k}') g(\mathbf{k}') \\ &= f(\mathbf{k}) g(\mathbf{k}). \end{aligned} \quad (\text{A.8})$$

This property can be trivially generalized to the convolution of arbitrary number of functions f_1, f_2, \dots, f_n . For their convolution $h(\mathbf{x})$,

$$\begin{aligned} h(\mathbf{x}) &= [f_1 \star \dots \star f_n](\mathbf{x}) \\ &\equiv \int d^3x_1 \dots \int d^3x_n f_1(\mathbf{x}_1) \dots f_n(\mathbf{x}_n) \delta^D(\mathbf{x} - \sum_{i=1}^n \mathbf{x}_i), \end{aligned} \quad (\text{A.9})$$

its Fourier transformation is given by

$$h(\mathbf{k}) = f_1(\mathbf{k}) f_2(\mathbf{k}) \dots f_n(\mathbf{k}) \quad (\text{A.10})$$

Conversely, for the product of two functions

$$h(\mathbf{x}) = f(\mathbf{x})g(\mathbf{x}), \quad (\text{A.11})$$

its Fourier transformation is given by the convolution of their Fourier transforms:

$$\begin{aligned} h(\mathbf{k}) &= \int d^3x f(\mathbf{x})g(\mathbf{x})e^{-i\mathbf{k}\cdot\mathbf{x}} \\ &= \int d^3x \left(\int \frac{d^3k_1}{(2\pi)^3} f(\mathbf{k}_1)e^{i\mathbf{k}_1\cdot\mathbf{x}} \right) \left(\int \frac{d^3k_2}{(2\pi)^3} g(\mathbf{k}_2)e^{i\mathbf{k}_2\cdot\mathbf{x}} \right) e^{-i\mathbf{k}\cdot\mathbf{x}} \\ &= \int \frac{d^3k_1}{(2\pi)^3} \int d^3k_2 f(\mathbf{k}_1)g(\mathbf{k}_2) \left(\int \frac{d^3x}{(2\pi)^3} e^{i\mathbf{x}\cdot(\mathbf{k}_1+\mathbf{k}_2-\mathbf{k})} \right) \\ &= \int \frac{d^3k_1}{(2\pi)^3} \int d^3k_2 f(\mathbf{k}_1)g(\mathbf{k}_2)\delta^D(\mathbf{k}_1+\mathbf{k}_2-\mathbf{k}) \end{aligned} \quad (\text{A.12})$$

Again, one can extend to n functions $f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_n(\mathbf{x})$. The Fourier transformation of the product of n functions

$$h(\mathbf{x}) = f_1(\mathbf{x}) \cdots f_n(\mathbf{x}) \quad (\text{A.13})$$

is given by the convolution in Fourier space.

$$\begin{aligned} h(\mathbf{k}) &= \int \frac{d^3k_1}{(2\pi)^3} \int \frac{d^3k_2}{(2\pi)^3} \cdots \int d^3k_n f_1(\mathbf{k}_1) \cdots f_n(\mathbf{k}_n) \delta^D(\mathbf{k} - \sum_{i=1}^n \mathbf{k}_i) \\ &= \frac{1}{(2\pi)^{3(n-1)}} [f_1 \star f_2 \star \cdots \star f_n](\mathbf{k}) \end{aligned} \quad (\text{A.14})$$

Note that we have $(2\pi)^{3(n-1)}$ in denominator, as Dirac delta function absorbs one factor of $(2\pi)^3$.

A.2 From Fourier Transform to Fourier Series

Let us consider the three dimensional cube of volume L^3 , and a periodic function $f(\mathbf{x})$ which has a period of L for all three directions. The periodicity implies that its Fourier transform $f(\mathbf{k})$ is discretized with the *fundamental frequency* $k_F = 2\pi/L$ as an interval,

$$f(\mathbf{k}) = \begin{cases} f(k_F \mathbf{n}_k) & \mathbf{n}_k = (i, j, k) \in Z^3 \\ 0 & \text{otherwise} \end{cases}, \quad (\text{A.15})$$

and the function $f(\mathbf{x})$ can be expressed as

$$f(\mathbf{x}) = \frac{1}{L^3} \sum_{\mathbf{n}_k} f(k_F \mathbf{n}_k) e^{ik_F \mathbf{n}_k \cdot \mathbf{x}}. \quad (\text{A.16})$$

Proof. As $f(\mathbf{x})$ has a periodicity \mathbf{L} ($\mathbf{L} = L(i, j, k)$, where i, j, k are integers), $f(\mathbf{x}) = f(\mathbf{x} + \mathbf{L})$, Fourier transform the both side leads $\int \frac{d^3k}{(2\pi)^3} f(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}} = \int \frac{d^3k}{(2\pi)^3} f(\mathbf{k}) e^{i\mathbf{k}\cdot(\mathbf{x}+\mathbf{L})}$. Therefore, the periodicity is maintained *if and only if* $\mathbf{k} \cdot \mathbf{L} = 2m\pi$ for some integer m . Since we assume the periodicity for all three spatial directions with same length L , three dimensional wave-vector has to be $\mathbf{k} = k_F \mathbf{n}_k$, with $k_F = 2\pi/L$ and some integer triplet \mathbf{n}_k . Then, the Fourier transformation now becomes a summation:

$$f(\mathbf{x}) = \int \frac{d^3k}{(2\pi)^3} f(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}} = \frac{\delta k^3}{(2\pi)^3} \sum_{\mathbf{n}_k} f(k_F \mathbf{n}_k) e^{ik_F \mathbf{n}_k \cdot \mathbf{x}} = \frac{1}{L^3} \sum_{\mathbf{n}_k} f(k_F \mathbf{n}_k) e^{ik_F \mathbf{n}_k \cdot \mathbf{x}}$$

Here, we use $\delta k^3 = k_F^3$. □

Note that when periods are different for all three directions: L_1 , L_2 and L_3 , we have to use three different fundamental frequencies $k_{Fi} \equiv 2\pi/L_i$ but $f(\mathbf{k})$ is still discretized with k_{Fi} intervals.

The equation (A.16) is simply a Fourier Series. Using the orthonormality condition in the periodic interval,

$$\int_0^{2\pi} dx e^{-inx} e^{imx} = 2\pi \delta_{m,n}, \quad (\text{A.17})$$

we can easily find the inverse :

$$f(\mathbf{k}) = \int_V d^3x f(\mathbf{x}) e^{-i\mathbf{k}\cdot\mathbf{x}} \quad (\text{A.18})$$

where, V stands for the cube of side L .

Similarly, if we consider a function defined at discrete points $f(\mathbf{x}_r)$, where $\mathbf{x}_r = H\mathbf{n}_r$ with a integer triplet \mathbf{n}_r , its Fourier transform becomes

$$f(\mathbf{k}) = \int d^3x f(\mathbf{x}_r) e^{-i\mathbf{k}\cdot\mathbf{x}_r} = H^3 \sum_{\mathbf{n}_r} f(H\mathbf{n}_r) e^{-iH\mathbf{k}\cdot\mathbf{n}_r}. \quad (\text{A.19})$$

It is also easy to see that $f(\mathbf{k})$ is the periodic function with periodicity of $k_g \equiv 2\pi/H$:

$$f(\mathbf{k}) = f(\mathbf{k} + k_g \mathbf{m}), \quad (\text{A.20})$$

for every integer triplet \mathbf{m} . Finally, using the orthonormal condition, we get the inverse transformation.

$$f(\mathbf{x}) = \int_{V_k} \frac{d^3k}{(2\pi)^3} f(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}} \quad (\text{A.21})$$

where V_k is the volume of the Fourier space cube of size k_g^3 . Therefore, the periodicity of a function implies the discreteness of its Fourier counterpart and the discreteness of a function implies the periodicity of its Fourier counterpart.

A.2.1 Sampling and aliasing

Let us consider the continuous underlying density field $\delta(\mathbf{x})$ and its Fourier transformation $\hat{\delta}(\mathbf{k})$, and sample the density field at discrete points $\mathbf{x}_p = H\mathbf{n}_p$ (\mathbf{n}_p is the integer triplet): $\hat{\delta}(\mathbf{x}_p) = \delta(\mathbf{x}_p)$. As we sample a function at discrete points, its Fourier pair is given by equation (A.19) and equation (A.21):

$$\hat{\delta}(\mathbf{k}) = H^3 \sum_{\mathbf{n}_p} \hat{\delta}(\mathbf{x}_p) e^{-i\mathbf{k} \cdot \mathbf{x}_p} \quad (\text{A.22})$$

$$\hat{\delta}(\mathbf{x}_p) = \int_{V_k} \frac{d^3k}{(2\pi)^3} \hat{\delta}(\mathbf{k}) e^{i\mathbf{k} \cdot \mathbf{x}_p}. \quad (\text{A.23})$$

At the same time, we can also relate the sampled density field to the underlying density field as

$$\hat{\delta}(\mathbf{x}_p) = \delta(\mathbf{x}_p) = \int \frac{d^3k}{(2\pi)^3} \delta(\mathbf{k}) e^{i\mathbf{k} \cdot \mathbf{x}_p} = \int_{V_k} \frac{d^3k'}{(2\pi)^3} \sum_{\mathbf{n}_a} \delta(\mathbf{k}' - k_g \mathbf{n}_a) e^{i\mathbf{k}' \cdot \mathbf{x}_p}, \quad (\text{A.24})$$

where in the second equality we divide the Fourier space into infinitely many cubes of k_g^3 , and, then, move those cubes into the V_k region by coordinate translation. By comparing equation (A.23) and equation (A.24), we find

$$\hat{\delta}(\mathbf{k}) = \sum_{\mathbf{n}_a} \delta(\mathbf{k} - k_g \mathbf{n}_a). \quad (\text{A.25})$$

This equation tells that the Fourier transform of the discrete-sampling of a function is equal to the *aliased sum* of infinitely many copies of Fourier transform of underlying function.

We can formulate it as follow. Let us define the one-dimensional sampling function

$$\Pi(x) = \sum_{n=-\infty}^{\infty} \delta_{x,n}, \quad (\text{A.26})$$

where $\delta_{x,n}$ is the Kronecker delta, which is 1 when $x = n$, and zero otherwise. Then, its Fourier transform is given by

$$\Pi(k) = 2\pi \sum_{m=-\infty}^{\infty} \delta^D(k - 2\pi m). \quad (\text{A.27})$$

Proof. The Fourier transform of the sampling function is

$$\Pi(k) = \int dx \Pi(x) e^{-ikx} = \sum_{n=-\infty}^{\infty} \int dx \delta_{x,n} e^{-ikx} = \lim_{N \rightarrow \infty} \sum_{n=-N}^N e^{-ikn}, \quad (\text{A.28})$$

where the sum of geometric series can be further simplified as,

$$\sum_{n=-N}^N e^{-ikn} = \frac{e^{ikN} - e^{ik(N+1)}}{1 - e^{-ik}} = (2N+1) \frac{\text{sinc}[k(2N+1)/2]}{\text{sinc}(k/2)}. \quad (\text{A.29})$$

This function peaks at $k = 2m\pi$ for integer m . As $N \rightarrow \infty$, the height of the peaks $(2N+1)$ goes to infinity, but the width of the peaks decreases as $1/(2N+1)$ at the same time so that the total area inside of peaks remains constant. Indeed, the function is proportional to the sum of infinitely many Dirac delta functions:

$$\Pi(k) = \lim_{N \rightarrow \infty} (2N+1) \frac{\text{sinc}[k(N+1/2)]}{\text{sinc}(k/2)} = 2\pi \sum_{m=-\infty}^{\infty} \delta^D(k - 2m\pi). \quad (\text{A.30})$$

□

Using the sampling function, the procedure of sampling can be think of as multiplying an underlying (continuous) density field to the sampling function.

$$\hat{\delta}(\mathbf{x}) \equiv \Pi\left(\frac{\mathbf{x}}{H}\right) \delta(\mathbf{x}), \quad (\text{A.31})$$

where the three dimensional sampling function is defined as $\Pi(\mathbf{x}) \equiv \Pi(x)\Pi(y)\Pi(z)$. By using the similarity theorem¹,

$$\text{FT} \left[\Pi\left(\frac{\mathbf{x}}{H}\right) \right] = 2\pi H \sum_{m=-\infty}^{\infty} \delta^D(kH - 2\pi m) = 2\pi \sum_{m=-\infty}^{\infty} \delta^D\left(k - \frac{2\pi}{H}m\right), \quad (\text{A.32})$$

we calculate Fourier Transform of the sampled function $\hat{\delta}(\mathbf{x})$ as from the convolution theorem:

$$\hat{\delta}(\mathbf{k}) = (2\pi)^3 \int \frac{d^3q}{(2\pi)^3} \sum_{\mathbf{n}_a} \delta^D\left(\mathbf{q} - \frac{2\pi}{H}\mathbf{n}_a\right) \delta(\mathbf{k} - \mathbf{q}) = \sum_{\mathbf{n}_a} \delta(\mathbf{k} - k_g \mathbf{n}_a), \quad (\text{A.33})$$

where $k_g = 2\pi/H$. Therefore, one can think of the Fourier Series as the Fourier Transformation of the sampled function with equal interval H .

Related to equation (A.33), there is a Nyquist sampling theorem. The theorem states that *aliasing can be avoided if the Nyquist frequency $k_g/2$ is greater than the maximum*

¹The similarity theorem states that the Fourier transformation of $f\left(\frac{x}{a}\right)$ is the same as $|a|f(ka)$ in one dimension. It is obvious that in three dimension, the pre-factor becomes $|a|^3$.

Proof. Changing of variable from x to $x' = x/a$ in the following integral $f(k) = \int dx f(x/a) e^{-ikx}$ leads $f(k) = |a| \int dx' f(x') e^{-ikx'a}$. Finally, change k to ka , we get $\text{FT}[f(x/a)] = |a|f(ka)$. □

component frequency. That is, as aliasing will mix the wave-modes coming from $\pm \mathbf{n}_r k_g$, if the power spectrum is truncated to be zero for $|k| > k_{Ny} \equiv k_g/2$, then there should be no aliasing effect. However, that's not the case for cosmological density power spectrum, and even worse, as shown in Chapter 2, the galaxy power spectrum contains a constant term which dominates on small scales. In that case, we can calculate the infinite sum of aliasing effect analytically. See Section 7.1.3.2 for detail.

In summary, sampling in the x space leads the periodicity in Fourier space. In addition to that, the relation between sampled x space function and its Fourier counterpart is given by the Fourier Series :

$$\delta(\mathbf{x}_p) = \int_{V_k} \frac{d^3 k}{(2\pi)^3} \delta(\mathbf{k}) e^{i\mathbf{k} \cdot \mathbf{x}_p} \quad (\text{A.34})$$

$$\delta(\mathbf{k}) = H^3 \sum_{\mathbf{x}_p} \delta(\mathbf{x}_p) e^{-i\mathbf{k} \cdot \mathbf{x}_p} \quad (\text{A.35})$$

Similarly, if we have the finite sampling in k space, it leads opposite pair of Fourier Series :

$$\delta(\mathbf{x}) = \frac{1}{L^3} \sum_{\mathbf{k}_p} \delta(\mathbf{k}_p) e^{i\mathbf{x} \cdot \mathbf{k}_p} \quad (\text{A.36})$$

$$\delta(\mathbf{k}_p) = \int_V d^3 x \delta(\mathbf{x}) e^{-i\mathbf{x} \cdot \mathbf{k}_p} \quad (\text{A.37})$$

Here, $\mathbf{x}_p = H \mathbf{n}_r$ and $\mathbf{k}_p = k_F \mathbf{n}_r$. The sampling interval is related to the periodicity in Fourier counter part as $k_g = 2\pi/H$ and $L = 2\pi/k_F$. The periodic volume is $V = L^3$ and $V_k = k_g^3$ for configuration space and Fourier space, respectively.

A.3 Fourier Series to Discrete Fourier Transform

Suppose we have a periodic function $\delta(\mathbf{x})$ which has a periodicity of L for all three directions. Consider the case when we sample the function $\delta(\mathbf{x})$ with constant interval $H = L/N$. In the previous section, we find that the sampled series of $\delta(\mathbf{x}_p)$ forms a Fourier Series of periodic function in \mathbf{k} space with following relation:

$$\delta(\mathbf{x}_p) = \int_{V_k} \frac{d^3 k}{(2\pi)^3} \delta(\mathbf{k}) e^{i\mathbf{k} \cdot \mathbf{x}_p} \quad (\text{A.38})$$

$$\delta(\mathbf{k}) = H^3 \sum_{\mathbf{x}_p} \delta(\mathbf{x}_p) e^{-i\mathbf{k} \cdot \mathbf{x}_p}. \quad (\text{A.39})$$

Now we add up the periodicity of $\delta(\mathbf{x})$, which implies that the Fourier counterpart $\delta(\mathbf{k})$ is also non-zero only for the case when $\mathbf{k} = k_F \mathbf{n}_k$, ($k_F = 2\pi/L$ is the fundamental frequency,

\mathbf{n}_k is an integer triplet). Therefore, the Fourier space integration becomes a finite sum as

$$\delta(\mathbf{x}_p) = \frac{1}{L^3} \sum_{\mathbf{k}_p \in V_k} \delta(\mathbf{k}_p) e^{i\mathbf{k}_p \cdot \mathbf{x}_p}. \quad (\text{A.40})$$

Inverse transformation can be found by using the orthogonality condition of the Discrete Fourier Transformation. For one dimensional DFT, the orthogonality condition is

$$\sum_{k=0}^{N-1} e^{-2\pi i k p / N} e^{2\pi i k q / N} = N \delta_{p, q \pm mN}^K \quad (\text{A.41})$$

where, p, q and m are integers, and N is the size of the DFT array. We can extend it to the n -dimension :

$$\sum_{\mathbf{n}_k} e^{-2\pi i \mathbf{n}_k \cdot \mathbf{n}_p / N} e^{2\pi i \mathbf{n}_k \cdot \mathbf{n}_q / N} = N^n \delta_{\mathbf{n}_p, \mathbf{n}_q \pm N \mathbf{n}_r}^K \quad (\text{A.42})$$

where, $\mathbf{n}_p, \mathbf{n}_q$ and \mathbf{n}_r is the n -dimensional triplet of integers.

Using integer triplets, the DFT becomes

$$\delta(H \mathbf{n}_r) = \frac{1}{L^3} \sum_{\mathbf{n}_k} \delta(k_F \mathbf{n}_k) e^{i \frac{2\pi}{N} \mathbf{n}_k \cdot \mathbf{n}_r} \quad (\text{A.43})$$

and using the identity above, the inverse transformation is:

$$\delta(k_F \mathbf{n}_q) = H^3 \sum_{\mathbf{n}_r} \delta(H \mathbf{n}_r) e^{-i \frac{2\pi}{N} \mathbf{n}_q \cdot \mathbf{n}_r}. \quad (\text{A.44})$$

Proof.

$$\begin{aligned} \sum_{\mathbf{n}_r} \delta(H \mathbf{n}_r) e^{-i \frac{2\pi}{N} \mathbf{n}_q \cdot \mathbf{n}_r} &= \frac{1}{L^3} \sum_{\mathbf{n}_k} \delta(k_F \mathbf{n}_k) \sum_{\mathbf{n}_r} e^{i \frac{2\pi}{N} \mathbf{n}_k \cdot \mathbf{n}_r} e^{-i \frac{2\pi}{N} \mathbf{n}_q \cdot \mathbf{n}_r} \\ &= \frac{1}{L^3} \sum_{\mathbf{n}_k} \delta(k_F \mathbf{n}_k) N^3 \delta_{\mathbf{n}_k, \mathbf{n}_q}^K = \frac{1}{H^3} \delta(k_F \mathbf{n}_q) \end{aligned}$$

□

Therefore, when we sample a periodic function $\delta(\mathbf{x})$ with data points \mathbf{x}_p , the DFT pair of the sampled list and its DFT dual, the sampling of $\delta(\mathbf{k})$ at \mathbf{k}_p , is given by

$$\delta(\mathbf{r}_p) = \frac{1}{L^3} \sum_{\mathbf{k}_p \in V_k} \delta(\mathbf{k}_p) e^{i\mathbf{k}_p \cdot \mathbf{x}_p} \quad (\text{A.45})$$

$$\delta(\mathbf{k}_p) = H^3 \sum_{\mathbf{x}_p \in V} \delta(\mathbf{x}_p) e^{-i\mathbf{k}_p \cdot \mathbf{x}_p}. \quad (\text{A.46})$$

A.4 Discrete Fourier Transform with FFTW

FFTW², the Fastest Fourier Transform in the West, is a publically available routine for the Fast Fourier Transform. In this section, we provide a practical guide on calculating Fourier Transform of a real field. For estimating the power spectrum and the bispectrum from N-body simulation and realistic galaxy survey, see Chapter 7, and for generating initial condition for N-body simulation, see Appendix E.

First, let us summarize two properties of the Discrete Fourier Transform of a real field. The Fourier transform $\delta_k(\mathbf{n}_k)$ of a regularly sampled (with N^d sampling points in d -dimensional space) real field $\delta_r(\mathbf{n}_r)$ satisfies

(P1) The Hermitianity condition:

$$\delta_k(-\mathbf{n}_k) = \delta_k^*(\mathbf{n}_k), \quad (\text{A.47})$$

where superscript $*$ denotes the complex conjugate.

(P2) The periodicity:

$$\delta_k(\mathbf{n}_k) = \delta_k(\mathbf{n}_k + N\mathbf{m}), \quad (\text{A.48})$$

for general integer vector \mathbf{m} .

As we often Fourier transform the real density field in three dimension, let us use the Fourier transform of $128 \times 128 \times 128$ array as an example. Following lists are the simplest sequence of FFTW in C/C++ and FORTRAN.

C/C++

```
fftw_complex *deltk;
double *deltar;
fftw_plan plan;
int nn = 128;
int cn2= nn/2+1;
int N = nn*nn*nn;
int Nc = nn*nn*cn2;

deltak = (fftw_complex*) fftw_malloc(sizeof(fftw_complex) * Nc);
deltar = (double *) fftw_malloc(N * sizeof(double));
plan = fftw_plan_dft_r2c_3d(nn, nn, nn, deltar, deltak);
<<--- code to calculate deltar --->
fftw_execute(plan);
fftw_destroy_plan(plan);
fftw_free(deltak);
fftw_free(deltar);
```

²<http://www.fftw.org>

FORTRAN

```

integer, parameter :: DOUBLE = KIND(0.0D0)
integer, parameter :: nn = 128
integer, parameter :: nnm = nn-1
integer, parameter :: cn = nn/2
integer*8 :: plan
real(KIND=DOUBLE), allocatable, dimension(:, :, :) :: deltar
complex(KIND=DOUBLE), allocatable, dimension(:, :, :) :: deltak

allocate(deltar(0:nnm,0:nnm,0:nnm),deltak(0:cn,0:nnm,0:nnm))
call dfftw_plan_dft_r2c_3d(plan,nn,nn,nn,deltar,deltak)
<<--- code to calculate deltar --->>
call dfftw_execute(plan)
call dfftw_destroy_plan(plan)
deallocate(dk)
deallocate(rhomesh)

```

For both cases, we declare two arrays: `deltar` and `deltak`, and `plan` for executing FFTW. For FORTRAN, `plan` has to be defined as `integer*8` type. Then, we allocate the arrays. For C/C++, FFTW has its own memory allocation subroutine (`fftw_malloc`), but in FORTRAN, we have to manually allocate the array.

Within the periodicity in Fourier space, $k_g = 2\pi/H$, there are $(k_g/k_F)^3 = (L/H)^3$ wave-modes. However, due to the Hermitianity condition (P1), the half of the modes inside of the periodic box are not real degrees of freedom. Therefore, `fftw_dft_r2c` routine of FFTW cut the half of the Fourier space cube. It is done for the fastest varying index; the first index for FORTRAN, and the last index for C/C++. Therefore, for one dimensional sampling size of `nn`, we have to set the input(real) array as

`deltar[0:nn-1,0:nn-1,0:nn-1]`

and output(complex) array as

`deltak[0:N/2,0:N-1,0:N-1],`

for FORTRAN, and

`deltak[0:N-1,0:N-1,0:N/2],`

for C/C++. After executing Discrete Fourier Transform by `fftw_execute`, we destroy the plan (`fftw_destroy`), and finally deallocate the arrays (`fftw_free`).

Note that the output array elements are stored *in order*, with the zero wavenumber (DC) component stored in `dk(0,0,0)`. All Fourier modes have positive wave numbers, and we can change the wave numbers by using the Periodicity (P2) condition as

`deltak[l,m,n] = deltak[(l,m,n)+N*(i,j,k)],`

where (i, j, k) is a integer triplet.

The output of **FFTW** is not normalized. That is, for a Fourier transform and inverse Fourier transform, it calculates

$$\delta^{\text{FFTW}}(\mathbf{n}_k) = \sum_{\mathbf{n}_r} \delta(\mathbf{n}_r) e^{-i2\pi \mathbf{n}_k \cdot \mathbf{n}_r} \equiv \frac{1}{H^3} \delta(\mathbf{n}_k), \quad (\text{A.49})$$

$$\delta^{\text{FFTW}}(\mathbf{n}_r) = \sum_{\mathbf{n}_k} \delta(\mathbf{n}_k) e^{i2\pi \mathbf{n}_k \cdot \mathbf{n}_r} \equiv L^3 \delta(\mathbf{n}_r), \quad (\text{A.50})$$

respectively. Here, we denote δ^{FFTW} as a **FFTW** result.

Let us look more closely on how output result of **FFTW** is stored. We first start from the simplest one-dimensional case, and expand to two and three dimensions. For this section, we shall uses the **FORTRAN** convention, and one can simply change the order of array indices for **C/C++** convention.

A.4.1 Output array of **dft_r2c_1d**

Let's start from the DFT of an array with N real numbers : $\mathbf{a}[\mathbf{n}]$ with $n = 0, \dots, N - 1$. The DFT of this real array is an array of N complex numbers : $\mathbf{Fa}[\mathbf{n}]$. Let's call $N/2$ (round down, or $(N - 1)/2$ if N is an odd integer) **cnmesh**, which corresponds to the *Nyquist frequency*.

Because of the periodicity of \mathbf{Fa} and the reality of \mathbf{a} , we have

$$\mathbf{Fa}[\mathbf{N}-\mathbf{n}] = \mathbf{Fa}[-\mathbf{n}] = \text{conjg}(\mathbf{Fa}[\mathbf{n}]).$$

That is, for every $\mathbf{n1} > \mathbf{cnmesh}$, there exists a $\mathbf{n2} = \mathbf{N} - \mathbf{n1} < \mathbf{cnmesh}$ such that

$$\mathbf{Fa}[\mathbf{n1}] = \text{conjg}(\mathbf{Fa}[\mathbf{n2}]).$$

Therefore, it is unnecessary to store $\mathbf{Fa}[\mathbf{n}]$ for $\mathbf{n} > \mathbf{cnmesh}$. Also, when $\mathbf{n} = 0$, we have

$$\mathbf{Fa}[0] = \text{conjg}(\mathbf{Fa}[0]),$$

therefore $\mathbf{Fa}[0]$ is a real number. If N is an even number, $\mathbf{Fa}[\mathbf{cnmesh}]$ is also real number as

$$\mathbf{Fa}[\mathbf{cnmesh}] = \mathbf{Fa}[\mathbf{N}-\mathbf{cnmesh}] = \mathbf{Fa}[-\mathbf{cnmesh}] = \text{conjg}(\mathbf{Fa}[\mathbf{cnmesh}]).$$

Combining all the result above, **FFTW** store the result of **dft_r2c_1d** to the complex array of $\mathbf{Fa}[0:\mathbf{cnmesh}]$, which includes $N/2 + 1$ (rounded down if N is an odd number)

complex numbers. Note that the number of degrees of freedom is conserved on this output array because one (for odd N) or two (for even N) elements are actually real. For even N , we have $2(N/2 + 1) - 2 = N$ real d.o.f., and for odd N , we have $2[(N - 1)/2 + 1] - 1 = N$ d.o.f in the output array.

Inversely, if one wants to construct the DFT component of a real array of size N , say, from a give power spectrum, one has to follow following procedure:

- (1) Prepare an array `Fa[0:cnmesh]`, where `cnmesh` = $N/2$, rounded down if N is odd.
- (2)-1 If N is even, assign a complex number to `Fa[1:cnmesh-1]`, and a real number to `Fa[0]=cmplx(#,0.d0)` and `Fa[cnmesh]=cmplx(#,0.d0)`
- (2)-2 If N is odd, assign a complex number to `Fa[1:cnmesh]`, and a real number to `Fa[0]=cmplx(#,0.d0)`

A.4.2 Output array of `dft_r2c_2d`

For two dimensional real array `a[0:N-1,0:N-1]` and its DFT complex array `Fa[0:N-1,0:N-1]`, we have the similar property :

$$\text{Fa}[N-n1, N-n2] = \text{Fa}[-n1, -n2] = \text{conj}(\text{Fa}[n1, n2]).$$

With this condition, again, we can reduce the number of elements in complex array `Fa` by roughly factor of two, as for every grid point `(n1,n2)` with `n1 > cnmesh`, there exists a grid point `(m1,m2)` whose `Fa` value is simply a complex conjugate of the original point. That is,

$$\text{Fa}[n1, n2] = \text{conj}(\text{Fa}[m1, m2])$$

if `m1=N-n1` and `m2=N-n2`. Therefore, it is sufficient to store a half of 2×2 plane, and conventionally FFTW reduces the direction of the fastest varying index by about half. That is, we need only `Fa[0:cnmesh,0:N-1]`. Using the condition above, it is also easy to show that DC component (`Fa[0,0]`), and the Nyquist components (`Fa[cnmesh,cnmesh]`, `Fa[0,cnmesh]`, `Fa[cnmesh,0]`, when N is even) are also real numbers.

One has to be careful about the boundaries with `n1=0` or `n1=cnmesh` (with even N). It is because in those boundaries, Fourier modes are not independent from each other. That is,

$$\text{Fa}[0, n2] = \text{conj}(\text{Fa}[N, N-n2]) = \text{conj}(\text{Fa}[0, N-n2])$$

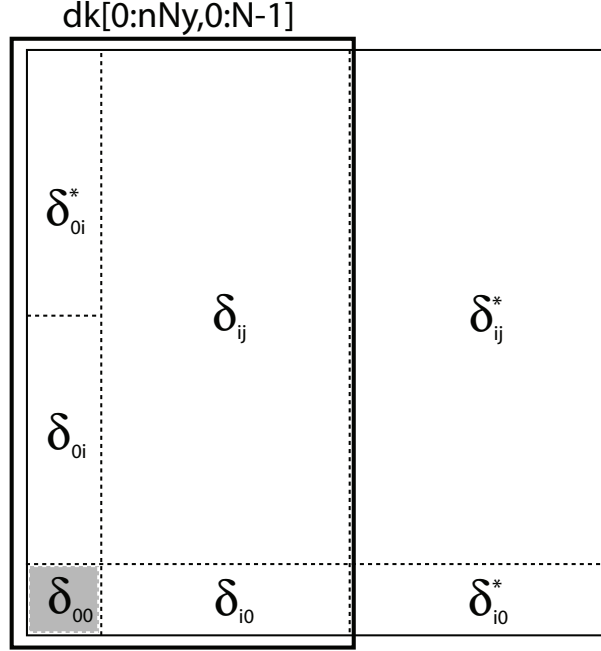


Figure A.1: Result of FFTW stored in the complex array with half size. Shaded region shows DC component which has a real value. This diagram is for an odd N .

for both even and odd N and

$$Fa[N/2, n2] = \text{conjg}(Fa[-N/2, -n2]) = \text{conjg}(Fa[N/2, N-n2])$$

only for even N .

Let us calculate the real degrees of freedom (d.o.f). For even N , we have $2N(N/2 + 1) = N^2 + 2N$ d.o.f. in the reduced **Fa**. Among them, four points (DC plus three Nyquist frequencies) have real values, and we have $(N - 2)/2$ pairs of complex-conjugated points on each boundary (**n1=0** and **n1=cnmesh**). Therefore, total real d.o.f. is $N^2 + 2N - 4 - 2(N - 2) = N^2$.

For odd N , we have total $2N((N - 1)/2 + 1) = N^2 + N$ d.o.f.s in the reduced **Fa**. For **n1=0** boundary, there are $(N - 1)/2$ of complex-conjugated pairs and one DC (real) point. In sum, total real d.o.f. is $N^2 + N - (N - 1) - 1 = N^2$.

In summary, here is a prescription of generating the DFT component of a real array of size $N \times N$.

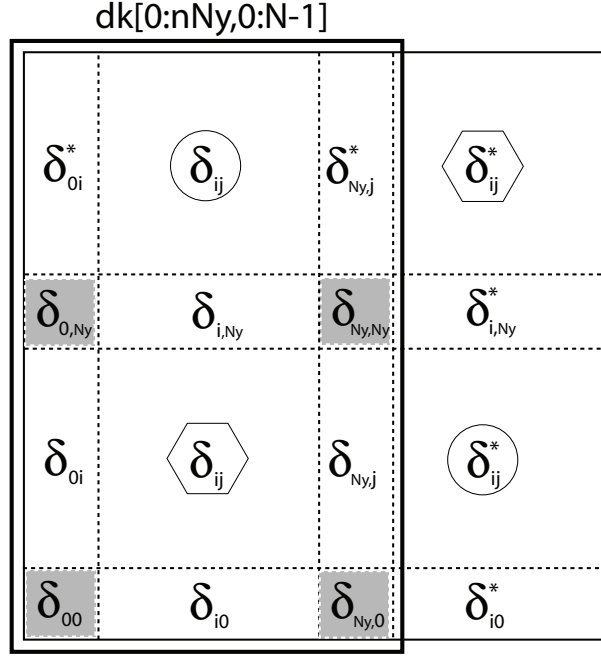


Figure A.2: Same as Figure A.1, but for an even N . DC components and Nyquist components are shaded with grey.

- (1) Prepare an array $Fa[0:cnmesh,0:N-1]$, where $cnmesh = N/2$, rounded down if N is odd.
- (2)-1 If N is even, assign a complex number to $Fa[1:cnmesh-1,0:N-1]$, and a real number to DC, $Fa[0,0]$, and Nyquist modes : $Fa[0,cnmesh]$ $Fa[cnmesh,0]$ $Fa[cnmesh,cnmesh]$. Then, assign complex numbers to a half row of $Fa[0,1:cnmesh-1]$ and $Fa[cnmesh,1:cnmesh-1]$, and its complex conjugate to another half row of $Fa[0,cnmesh+1:N-1]$ and $Fa[cnmesh,cnmesh+1:N-1]$.
- (2)-2 If N is odd, assign a complex number to $Fa[1:cnmesh,0:N-1]$, and a real number to $Fa[0,0]$. Then, assign complex numbers to a half row of $Fa[0,1:cnmesh]$, and its complex conjugate to another half row of $Fa[0,cnmesh+1:N-1]$.

The data structure and symmetric properties are shown in Figure A.1 and A.2.

A.4.3 Output array of `dft_r2c_3d`

Extend to three-dimensional array is trivial. With the general property

$$\text{Fa}[N-n1, N-n2, N-n3] = \text{Fa}[-n1, -n2, -n3] = \text{conjg}(\text{Fa}[n1, n2, n3]).$$

With the same argument, we can reduce the number of elements in complex array **Fa** by roughly factor of two. Therefore, it is sufficient to store only a half of a cube, and conventionally FFTW reduces the direction of the fastest varying index by about half. That is, we need only `Fa[0:cnmesh, 0:N-1, 0:N-1]`. Using the condition above, it is also easy to show that DC component (`Fa[0,0,0]`), and the Nyquist components (`Fa[cnmesh, cnmesh, cnmesh]`, `Fa[0, cnmesh, cnmesh]`, `Fa[cnmesh, 0, cnmesh]`, `Fa[cnmesh, cnmesh, 0]`, `Fa[0, 0, cnmesh]`, `Fa[0, cnmesh, 0]`, `Fa[0, 0, cnmesh]`, when N is even) are also real numbers.

One has to be careful about the boundaries with `n1=0` or `n1=cnmesh` (with even N). It is because in those boundaries, Fourier modes are not independent from each other. That is,

$$\text{Fa}[0, n2, n3] = \text{conjg}(\text{Fa}[N, N-n2, N-n3]) = \text{conjg}(\text{Fa}[0, N-n2, N-n3])$$

for both even and odd N and

$$\text{Fa}[N/2, n2, n3] = \text{conjg}(\text{Fa}[-N/2, -n2, -n3]) = \text{conjg}(\text{Fa}[N/2, N-n2, N-n3])$$

only for even N .

Let's calculate the real d.o.f.. For even N , we have $2N^2(N/2+1) = N^3 + 2N^2$ d.o.f. in the reduced **Fa**. Among them, eight points (DC plus seven Nyquist frequencies) have real values, and we have $(N^2 - 4)/2$ pairs of complex-conjugated points on each boundary (`n1=0` and `n1=cnmesh`). Therefore, total real d.o.f. is $N^3 + 2N^2 - 8 - 2(N^2 - 4) = N^3$. For odd N , we have total $2N^2((N-1)/2+1) = N^3 + N^2$ d.o.f.s in the reduced **Fa**. For `n1=0` boundary, there are $(N^2 - 1)/2$ of complex-conjugated pairs and one DC (real) point. In sum, total real d.o.f. is $N^3 + N^2 - (N^2 - 1) - 1 = N^3$.

In summary, here is the prescription to generate the DFT component of a real array of size $N \times N \times N$.

(1) Prepare an array $\text{Fa}[0:\text{cnmesh}, 0:N-1, 0:N-1]$, where $\text{cnmesh} = N/2$, rounded down if N is odd.

(2)-1 If N is even, assign a complex number to $\text{Fa}[1:\text{cnmesh}-1, 0:N-1, 0:N-1]$, and a real number to DC, $\text{Fa}[0, 0, 0]$, and Nyquest modes :

$\text{Fa}[\text{cnmesh}, \text{cnmesh}, \text{cnmesh}]$, $\text{Fa}[0, \text{cnmesh}, \text{cnmesh}]$,
 $\text{Fa}[\text{cnmesh}, 0, \text{cnmesh}]$, $\text{Fa}[\text{cnmesh}, \text{cnmesh}, 0]$, $\text{Fa}[0, 0, \text{cnmesh}]$,
 $\text{Fa}[0, \text{cnmesh}, 0]$, $\text{Fa}[0, 0, \text{cnmesh}]$.

Then, assign complex numbers to a half row of $\text{Fa}[0, 1:\text{cnmesh}, 0:N-1]$ and $\text{Fa}[\text{cnmesh}, 1:\text{cnmesh}, 0:N-1]$, and its complex conjugate to another half row of $\text{Fa}[0, \text{cnmesh}+1:N-1, 0:N-1]$ and $\text{Fa}[\text{cnmesh}, \text{cnmesh}+1:N-1, 0:N-1]$, respectively. Also, assign the complex conjugate pair to

$\text{Fa}[0, (/0, \text{cnmesh}/), 1:N-1]$ and $\text{Fa}[\text{cnmesh}, (/0, \text{cnmesh}/), 0:N-1]$. See, Figure A.2.

(2)-2 If N is odd, assign a complex number to $\text{Fa}[1:\text{cnmesh}, 0:N-1, 0:N-1]$, and a real number to $\text{Fa}[0, 0, 0]$.

Then, assign complex numbers to a half row of $\text{Fa}[0, 1:\text{cnmesh}, 0:N-1]$, and its complex conjugate to another half row of $\text{Fa}[0, \text{cnmesh}+1:N-1, 0:N-1]$. See, Figure A.1.

Appendix B

Derivation of one-loop power spectrum

In this appendix, we derive the one-loop (next-to-leading order) power spectrum of the quantity $X(\mathbf{k}, \tau)$ whose general perturbative expansion is given by equation (2.55).

B.1 Gaussian linear density field

We calculate $P_{X,13}$ and $P_{X,22}$ when $\delta_1(\mathbf{k}, \tau)$ follows the Gaussian distribution. As these functions involve the ensemble average of four δ_1 s, we shall use following identity from Wick's theorem¹:

$$\begin{aligned} \langle \delta_1(\mathbf{k}_1) \delta_1(\mathbf{k}_3) \delta_1(\mathbf{k}_3) \delta_1(\mathbf{k}_4) \rangle &= (2\pi)^6 \left[P_L(k_1) P_L(k_3) \delta^D(\mathbf{k}_1 + \mathbf{k}_2) \delta^D(\mathbf{k}_3 + \mathbf{k}_4) \right. \\ &\quad + P_L(k_3) P_L(k_2) \delta^D(\mathbf{k}_1 + \mathbf{k}_3) \delta^D(\mathbf{k}_2 + \mathbf{k}_4) \\ &\quad \left. + P_L(k_2) P_L(k_1) \delta^D(\mathbf{k}_1 + \mathbf{k}_4) \delta^D(\mathbf{k}_2 + \mathbf{k}_3) \right]. \end{aligned} \quad (\text{B.1})$$

Furthermore, by requiring that the ensemble average of $X(\mathbf{k}, \tau)$ has to vanish, we have following constraint

$$K_2^{(s)}(\mathbf{k}, -\mathbf{k}) = 0.$$

Let's first calculate $P_{X,22}(\mathbf{k})$, which is defined as

$$\left\langle X^{(2)}(\mathbf{k}) X^{(2)}(\mathbf{k}') \right\rangle \equiv (2\pi)^3 P_{X,22}(\mathbf{k}) \delta^D(\mathbf{k} + \mathbf{k}'), \quad (\text{B.2})$$

¹For a rigorous treatment of the Gaussian random field and the derivation of Wick's theorem, see Section 2.2.

where we omit the explicit time dependency for brevity. We evaluate the left hand side of the definition

$$\begin{aligned}
& \langle X^{(2)}(\mathbf{k})X^{(2)}(\mathbf{k}') \rangle \\
&= \int \frac{d^3 q_1}{(2\pi)^3} \int \frac{d^3 q_2}{(2\pi)^3} K_2^{(s)}(\mathbf{q}_1, \mathbf{k} - \mathbf{q}_1) K_2^{(s)}(\mathbf{q}_2, \mathbf{k}' - \mathbf{q}_2) \\
&\quad \times \langle \delta_1(\mathbf{q}_1) \delta_1(\mathbf{k} - \mathbf{q}_1) \delta_1(\mathbf{q}_2) \delta_1(\mathbf{k}' - \mathbf{q}_2) \rangle \\
&= \int d^3 q_1 \int d^3 q_2 K_2^{(s)}(\mathbf{q}_1, \mathbf{k} - \mathbf{q}_1) K_2^{(s)}(\mathbf{q}_2, \mathbf{k}' - \mathbf{q}_2) \\
&\quad \times \left[P_L(\mathbf{q}_1) P_L(\mathbf{k} - \mathbf{q}_1) \delta^D(\mathbf{q}_1 + \mathbf{q}_2) \delta^D(\mathbf{k} + \mathbf{k}') \right. \\
&\quad \left. + P_L(\mathbf{q}_1) P_L(\mathbf{q}_2) \delta^D(\mathbf{k}' + \mathbf{q}_1 - \mathbf{q}_2) \delta^D(\mathbf{k} + \mathbf{q}_2 - \mathbf{q}_1) \right] \\
&= (2\pi)^3 \left\{ 2 \int \frac{d^3 q}{(2\pi)^3} P_L(q) P_L(|\mathbf{k} - \mathbf{q}|) \left[K_2^{(s)}(\mathbf{q}, \mathbf{k} - \mathbf{q}) \right]^2 \right\} \delta^D(\mathbf{k} + \mathbf{k}'),
\end{aligned}$$

therefore

$$P_{X,22}(\mathbf{k}) = 2 \int \frac{d^3 q}{(2\pi)^3} P_L(q) P_L(|\mathbf{k} - \mathbf{q}|) \left[K_2^{(s)}(\mathbf{q}, \mathbf{k} - \mathbf{q}) \right]^2. \quad (\text{B.3})$$

We define $P_{X,13}(\mathbf{k})$ in the similar manner:

$$\langle X^{(1)}(\mathbf{k}') X^{(3)}(\mathbf{k}) \rangle \equiv (2\pi)^3 P_{X,13}(\mathbf{k}) \delta^D(\mathbf{k} + \mathbf{k}'). \quad (\text{B.4})$$

We calculate the left hand side as

$$\begin{aligned}
& \langle \delta_1(\mathbf{k}) \delta_3(\mathbf{k}') \rangle \\
&= K_1^{(s)}(\mathbf{k}') \int \frac{d^3 q_1}{(2\pi)^3} \int \frac{d^3 q_2}{(2\pi)^3} \int d^3 q_3 \delta^D(\mathbf{k}' - \mathbf{q}_1 - \mathbf{q}_2 - \mathbf{q}_3) K_3^{(s)}(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) \\
&\quad \times \langle \delta_1(\mathbf{k}') \delta_1(\mathbf{q}_1) \delta_1(\mathbf{q}_2) \delta_1(\mathbf{q}_3) \rangle \\
&= 3K_1^{(s)}(\mathbf{k}') \int d^3 q_1 \int d^3 q_2 \int d^3 q_3 \delta^D(\mathbf{k} - \mathbf{q}_1 - \mathbf{q}_2 - \mathbf{q}_3) K_3^{(s)}(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) \\
&\quad \times P_L(\mathbf{k}') P_L(\mathbf{q}_2) \delta^D(\mathbf{k}' + \mathbf{q}_1) \delta^D(\mathbf{q}_2 + \mathbf{q}_3) \\
&= (2\pi)^3 \left[3K_1^{(s)}(-\mathbf{k}) P_L(k) \int \frac{d^3 q}{(2\pi)^3} P_L(q) K_3^{(s)}(\mathbf{k}, \mathbf{q}, -\mathbf{q}) \right] \delta^D(\mathbf{k} + \mathbf{k}').
\end{aligned}$$

For the galaxy power spectrum $K_1^{(s)}(\mathbf{k})$ is a real function, and from the Hermitianity of $X^{(1)}(-\mathbf{k}) = [X^{(1)}(\mathbf{k})]^*$, one can show that $K_1^{(s)}(-\mathbf{k}) = K_1^{(s)}(\mathbf{k})$. Therefore,

$$P_{X,13}(\mathbf{k}) = 3K_1^{(s)}(\mathbf{k}) P_L(k) \int \frac{d^3 q}{(2\pi)^3} P_L(q) K_3^{(s)}(\mathbf{k}, \mathbf{q}, -\mathbf{q}). \quad (\text{B.5})$$

B.2 Non-Gaussian linear density field

When linear density field follows non-Gaussian distribution, effect on the power spectrum comes through the higher-order correlation function such as bispectrum, $B_L(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)$, and trispectrum $T_L(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4)^2$.

First, we calculate the leading order non-Gaussian correction term which comes from δ_1^3 term in equation (2.57):

$$\begin{aligned} & \langle X^{(1)}(\mathbf{k}') X^{(2)}(\mathbf{k}) \rangle \\ &= K_1^{(s)}(\mathbf{k}') \int \frac{d^3 q}{(2\pi)^3} K_2^{(s)}(\mathbf{q}, \mathbf{k} - \mathbf{q}) \langle \delta_1(\mathbf{k}') \delta_1(\mathbf{q}) \delta_1(\mathbf{k} - \mathbf{q}) \rangle \\ &= (2\pi)^3 \left[K_1^{(s)}(\mathbf{k}') \int \frac{d^3 q}{(2\pi)^3} K_2^{(s)}(\mathbf{q}, \mathbf{k} - \mathbf{q}) B_L(\mathbf{k}', \mathbf{q}, \mathbf{k} - \mathbf{q}) \right] \delta^D(\mathbf{k} + \mathbf{k}'). \end{aligned}$$

Therefore, the leading order non-Gaussian term is

$$\begin{aligned} \Delta P_{X,NG}(\mathbf{k}) &= 2P_{X,12}(\mathbf{k}) \\ &= 2K_1^{(s)}(\mathbf{k}) \int \frac{d^3 q}{(2\pi)^3} K_2^{(s)}(\mathbf{q}, \mathbf{k} - \mathbf{q}) B_L(-\mathbf{k}, \mathbf{q}, \mathbf{k} - \mathbf{q}). \end{aligned} \quad (\text{B.8})$$

Another non-Gaussian correction term comes from $P_{X,22}$ and $P_{X,13}$. For non-Gaussian linear density field, the four point correlator becomes

$$\begin{aligned} \langle \delta_1(\mathbf{k}_1) \delta_1(\mathbf{k}_2) \delta_1(\mathbf{k}_3) \delta_1(\mathbf{k}_4) \rangle &= (2\pi)^6 \left[P_L(k_1) P_L(k_3) \delta^D(\mathbf{k}_1 + \mathbf{k}_2) \delta^D(\mathbf{k}_3 + \mathbf{k}_4) \right. \\ &\quad + P_L(k_3) P_L(k_2) \delta^D(\mathbf{k}_1 + \mathbf{k}_3) \delta^D(\mathbf{k}_2 + \mathbf{k}_4) \\ &\quad + P_L(k_2) P_L(k_1) \delta^D(\mathbf{k}_1 + \mathbf{k}_4) \delta^D(\mathbf{k}_2 + \mathbf{k}_3) \left. \right] \\ &\quad + (2\pi)^3 T_L(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) \delta^D(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3 + \mathbf{k}_4), \end{aligned}$$

instead of equation (B.1), which is valid for Gaussian case. The new term in the four point correlator which is proportional to the linear trispectrum generates yet another class of

²Bispectrum and Trispectrum are defined as

$$\langle \delta_1(\mathbf{k}_1) \delta_1(\mathbf{k}_2) \delta_1(\mathbf{k}_3) \rangle = (2\pi)^3 B_L(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) \delta^D(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3), \quad (\text{B.6})$$

and

$$\langle \delta_1(\mathbf{k}_1) \delta_1(\mathbf{k}_2) \delta_1(\mathbf{k}_3) \delta_1(\mathbf{k}_4) \rangle_c = (2\pi)^3 T_L(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) \delta^D(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3 + \mathbf{k}_4), \quad (\text{B.7})$$

where $\langle \dots \rangle_c$ denotes the *connected* part of correlation function.

non-Gaussian correction terms:

$$\begin{aligned}\Delta P_{X,22}(\mathbf{k}) &= \int \frac{d^3 q_1}{(2\pi)^3} \int \frac{d^3 q_2}{(2\pi)^3} K_2^{(s)}(\mathbf{q}_1, \mathbf{k} - \mathbf{q}_1) K_2^{(s)}(\mathbf{q}_2, -\mathbf{k} - \mathbf{q}_2) \\ &\quad \times T_L(\mathbf{q}_1, \mathbf{k} - \mathbf{q}_1, \mathbf{q}_2, -\mathbf{k} - \mathbf{q}_2)\end{aligned}\tag{B.9}$$

$$\begin{aligned}\Delta P_{X,13}(\mathbf{k}) &= K_1^{(s)}(\mathbf{k}) \int \frac{d^3 q_1}{(2\pi)^3} \int \frac{d^3 q_2}{(2\pi)^3} K_3^{(s)}(\mathbf{q}_1, \mathbf{q}_2, \mathbf{k} - \mathbf{q}_1 - \mathbf{q}_2) \\ &\quad \times T_L(-\mathbf{k}, \mathbf{q}_1, \mathbf{q}_2, \mathbf{k} - \mathbf{q}_1 - \mathbf{q}_2).\end{aligned}\tag{B.10}$$

As linear trispectrum is generated from two-loop order in the curvature perturbation, ζ , these terms are important on even larger scales than scales where equation (B.8) is important.

Appendix C

Derivation of the primordial bispectrum and trispectrum from local non-Gaussianity

In this appendix, we derive the bispectrum and the trispectrum of primordial curvature perturbation in the presence of local type non-Gaussianity. For local type non-Gaussianity, primordial curvature perturbation, $\Phi(\mathbf{x})$, is given by

$$\Phi(\mathbf{x}) = \phi(\mathbf{x}) + f_{\text{NL}} [\phi^2(\mathbf{x}) - \langle \phi^2 \rangle] + g_{\text{NL}} \phi^3(\mathbf{x}) + \dots \quad (\text{C.1})$$

In Fourier space, by using the convolution theorem, we can recast the equation as

$$\begin{aligned} \Phi(\mathbf{k}) &\equiv \phi^{(1)}(\mathbf{k}) + \phi^{(2)}(\mathbf{k}) + \phi^{(3)}(\mathbf{k}) + \dots \\ &= \phi(\mathbf{k}) + f_{\text{NL}} \left[\int \frac{d^3 q_1}{(2\pi)^3} \int d^3 q_2 \phi(\mathbf{q}_1) \phi(\mathbf{q}_2) \delta^D(\mathbf{k} - \mathbf{q}_{12}) - \langle \phi^2 \rangle \delta^D(\mathbf{k}) \right] \\ &\quad + g_{\text{NL}} \int \frac{d^3 q_1}{(2\pi)^3} \int \frac{d^3 q_2}{(2\pi)^3} \int d^3 q_3 \phi(\mathbf{q}_1) \phi(\mathbf{q}_2) \phi(\mathbf{q}_3) \delta^D(\mathbf{k} - \mathbf{q}_{123}) \\ &\quad + \dots \end{aligned} \quad (\text{C.2})$$

The bispectrum of $\Phi(\mathbf{k})$ is defined as a three point correlator in Fourier space:

$$\langle \Phi(\mathbf{k}_1) \Phi(\mathbf{k}_2) \Phi(\mathbf{k}_3) \rangle = (2\pi)^3 B(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) \delta^D(\mathbf{k}_{123}). \quad (\text{C.3})$$

We calculate the primordial bispectrum by plugging equation (C.2) into left hand side of equation (C.3). As ϕ follows Gaussian distribution, the leading order bispectrum comes from $\langle \phi^{(1)} \phi^{(1)} \phi^{(2)} \rangle$. Let us pick up one of three such terms.

$$\begin{aligned} &\langle \phi^{(1)}(\mathbf{k}_1) \phi^{(1)}(\mathbf{k}_2) \phi^{(2)}(\mathbf{k}_3) \rangle \\ &= f_{\text{NL}} \int \frac{d^3 q_1}{(2\pi)^3} \int d^3 q_2 \delta^D(\mathbf{k}_3 - \mathbf{q}_{12}) \langle \phi(\mathbf{k}_1) \phi(\mathbf{k}_2) \phi(\mathbf{q}_1) \phi(\mathbf{q}_2) \rangle \\ &\quad - f_{\text{NL}} \langle \phi^2 \rangle \delta^D(\mathbf{k}_3) \langle \phi(\mathbf{k}_1) \phi(\mathbf{k}_2) \rangle \\ &= 2f_{\text{NL}} (2\pi)^3 \int d^3 q_1 \int d^3 q_2 \delta^D(\mathbf{k}_3 - \mathbf{q}_{12}) P_\phi(\mathbf{k}_1) P_\phi(\mathbf{k}_2) \delta^D(\mathbf{k}_1 + \mathbf{q}_1) \delta^D(\mathbf{k}_2 + \mathbf{q}_2) \\ &= (2\pi)^3 [2f_{\text{NL}} P_\phi(\mathbf{k}_1) P_\phi(\mathbf{k}_2)] \delta^D(\mathbf{k}_{123}) \end{aligned} \quad (\text{C.4})$$

In order to evaluate the four-point correlator, we use the Wick's theorem [Eq. (B.1)]. By adding up cyclic combinations of equation (C.4), we calculate the primordial bispectrum of curvature perturbation as

$$B_\Phi(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = 2f_{\text{NL}} [P_\phi(\mathbf{k}_1)P_\phi(\mathbf{k}_2) + P_\phi(\mathbf{k}_2)P_\phi(\mathbf{k}_3) + P_\phi(\mathbf{k}_3)P_\phi(\mathbf{k}_1)]. \quad (\text{C.5})$$

The trispectrum of $\Phi(\mathbf{k})$ is the four-point correlator in Fourier space:

$$\langle \Phi(\mathbf{k}_1)\Phi(\mathbf{k}_2)\Phi(\mathbf{k}_3)\Phi(\mathbf{k}_4) \rangle_c = (2\pi)^3 T(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) \delta^D(\mathbf{k}_{1234}). \quad (\text{C.6})$$

Explicitly expanding the left hand side by using equation (C.2) yields $(\phi_i^{(n)})$ is the short-hand notation of $\phi^{(n)}(\mathbf{k}_i)$

$$\begin{aligned} & \langle \Phi(\mathbf{k}_1)\Phi(\mathbf{k}_2)\Phi(\mathbf{k}_3)\Phi(\mathbf{k}_4) \rangle \\ &= \langle \phi_1^{(1)}\phi_2^{(1)}\phi_3^{(1)}\phi_4^{(1)} \rangle + \left[\langle \phi_1^{(1)}\phi_2^{(1)}\phi_3^{(1)}\phi_4^{(3)} \rangle + (3 \text{ cyclic}) \right] \\ & \quad + \left[\langle \phi_1^{(1)}\phi_2^{(1)}\phi_3^{(2)}\phi_4^{(2)} \rangle + (5 \text{ cyclic}) \right], \end{aligned}$$

and among them, trispectrum picks up only *connected* correlator:

$$\begin{aligned} & \langle \Phi(\mathbf{k}_1)\Phi(\mathbf{k}_2)\Phi(\mathbf{k}_3)\Phi(\mathbf{k}_4) \rangle_c \\ &= \left[\langle \phi_1^{(1)}\phi_2^{(1)}\phi_3^{(1)}\phi_4^{(3)} \rangle_c + (3 \text{ cyclic}) \right] + \left[\langle \phi_1^{(1)}\phi_2^{(1)}\phi_3^{(2)}\phi_4^{(2)} \rangle_c + (5 \text{ cyclic}) \right]. \end{aligned} \quad (\text{C.7})$$

Each piece of connected trispectrum is a product of three power spectra of Gaussian field ϕ . First, let us calculate one of $\langle \phi^{(1)}\phi^{(1)}\phi^{(1)}\phi^{(3)} \rangle_c$ terms.

$$\begin{aligned} \langle \phi_1^{(1)}\phi_2^{(1)}\phi_3^{(1)}\phi_4^{(3)} \rangle_c &= g_{\text{NL}} \int \frac{d^3 q_1}{(2\pi)^3} \int \frac{d^3 q_2}{(2\pi)^3} \int d^3 q_3 \delta^D(\mathbf{k}_4 - \mathbf{q}_{123}) \\ & \quad \times \langle \phi(\mathbf{k}_1)\phi(\mathbf{k}_2)\phi(\mathbf{k}_3)\phi(\mathbf{q}_1)\phi(\mathbf{q}_2)\phi(\mathbf{q}_3) \rangle_c \end{aligned} \quad (\text{C.8})$$

As the *connected* correlator does not count the case where any wave vectors \mathbf{k}_i are factored out as a delta function (e.g. $\delta^D(\mathbf{k}_i)$ or $\delta^D(\mathbf{k}_i + \mathbf{k}_j)$) except for $\delta^D(\mathbf{k}_{1234})$, only six (3!) of 15 (5!!) Wick contraction terms are qualified as a part of trispectrum. For all of six Wick contractions, we contract one $\phi(\mathbf{k}_i)$ with one $\phi(\mathbf{q}_i)$, and equation (C.8) becomes

$$\begin{aligned} & \langle \phi_1^{(1)}\phi_2^{(1)}\phi_3^{(1)}\phi_4^{(3)} \rangle_c \\ &= (2\pi)^3 \sum_{\text{perm.}} g_{\text{NL}} \int d^3 q_1 \int d^3 q_2 \int d^3 q_3 \delta^D(\mathbf{k}_4 - \mathbf{q}_{123}) \\ & \quad \times P_\phi(k_1)P_\phi(k_2)P_\phi(k_3)\delta^D(\mathbf{k}_1 + \mathbf{q}_i)\delta^D(\mathbf{k}_2 + \mathbf{q}_j)\delta^D(\mathbf{k}_3 + \mathbf{q}_k), \end{aligned} \quad (\text{C.9})$$

where sum over permutation means that (i, j, k) can be any permutation of the set $\{1, 2, 3\}$. After integrating all three delta functions, we find

$$\langle \phi_1^{(1)} \phi_2^{(1)} \phi_3^{(1)} \phi_4^{(3)} \rangle_c = (2\pi)^3 \left[6g_{\text{NL}} P_\phi(k_1) P_\phi(k_2) P_\phi(k_3) \right] \delta^D(\mathbf{k}_{1234}). \quad (\text{C.10})$$

We also calculate $\langle \phi^{(1)} \phi^{(1)} \phi^{(2)} \phi^{(2)} \rangle_c$ terms with similar method. Let us focus on one of 1122-terms

$$\begin{aligned} & \langle \phi_1^{(1)} \phi_2^{(1)} \phi_3^{(2)} \phi_4^{(2)} \rangle_c \\ &= f_{\text{NL}}^2 \int \frac{d^3 q_1}{(2\pi)^3} \int \frac{d^3 q_2}{(2\pi)^3} \langle \phi(\mathbf{k}_1) \phi(\mathbf{k}_2) \phi(\mathbf{q}_1) \phi(\mathbf{k}_3 - \mathbf{q}_1) \phi(\mathbf{q}_2) \phi(\mathbf{k}_4 - \mathbf{q}_2) \rangle_c. \end{aligned}$$

In order to calculate the connected correlator, we should not contract $\phi(\mathbf{k}_1)$ with $\phi(\mathbf{k}_2)$, $\phi(\mathbf{q}_1)$ with $\phi(\mathbf{k}_3 - \mathbf{q}_1)$, and $\phi(\mathbf{q}_2)$ with $\phi(\mathbf{k}_4 - \mathbf{q}_2)$. As a result, there are eight different way of pairing six ϕ s: first contracting one of $\{\phi(\mathbf{q}_1), \phi(\mathbf{k}_3 - \mathbf{q}_1)\}$ with one of $\{\phi(\mathbf{q}_2), \phi(\mathbf{k}_4 - \mathbf{q}_2)\}$ (4 ways), there are two ways of contracting $\phi(\mathbf{k}_1)$ and $\phi(\mathbf{k}_2)$ with two left-overs of the previous contraction. After contracting all three pairs, the correlator becomes

$$\begin{aligned} & \langle \phi(\mathbf{k}_1) \phi(\mathbf{k}_2) \phi(\mathbf{q}_1) \phi(\mathbf{k}_3 - \mathbf{q}_1) \phi(\mathbf{q}_2) \phi(\mathbf{k}_4 - \mathbf{q}_2) \rangle_c \\ &= (2\pi)^9 P_\phi(k_1) P_\phi(k_2) \\ & \times \left[\delta^D(\mathbf{k}_1 + \mathbf{q}_1) \delta^D(\mathbf{k}_2 + \mathbf{q}_2) P_\phi(\mathbf{k}_3 - \mathbf{q}_1) \delta^D(\mathbf{k}_4 - \mathbf{q}_2) \right. \\ & \quad + \delta^D(\mathbf{k}_1 + \mathbf{q}_1) \delta^D(\mathbf{k}_2 + \mathbf{k}_4 - \mathbf{q}_2) P_\phi(\mathbf{k}_3 - \mathbf{q}_1) \delta^D(\mathbf{k}_3 - \mathbf{q}_1 + \mathbf{q}_2) \\ & \quad + \delta^D(\mathbf{k}_1 + \mathbf{k}_3 - \mathbf{q}_1) \delta^D(\mathbf{k}_2 + \mathbf{q}_2) P_\phi(\mathbf{q}_1) \delta^D(\mathbf{k}_4 - \mathbf{q}_2 + \mathbf{q}_1) \\ & \quad \left. + \delta^D(\mathbf{k}_1 + \mathbf{k}_3 - \mathbf{q}_1) \delta^D(\mathbf{k}_2 + \mathbf{k}_4 - \mathbf{q}_2) P_\phi(\mathbf{q}_1) \delta^D(\mathbf{q}_{12}) + (\mathbf{k}_1 \leftrightarrow \mathbf{k}_2) \right]. \end{aligned}$$

After integrating the delta functions, we find

$$\begin{aligned} & \langle \phi_1^{(1)} \phi_2^{(1)} \phi_3^{(2)} \phi_4^{(2)} \rangle_c \\ &= (2\pi)^3 \left[f_{\text{NL}}^2 P_\phi(k_1) P_\phi(k_2) \{4P_\phi(k_{13}) + 4P_\phi(k_{23})\} \right] \delta^D(\mathbf{k}_{1234}). \end{aligned}$$

By using $\mathbf{k}_{1234} = 0$, we can rewrite the equation in the form that the symmetry is most apparent:

$$\begin{aligned} & \langle \phi_1^{(1)} \phi_2^{(1)} \phi_3^{(2)} \phi_4^{(2)} \rangle_c \\ &= (2\pi)^3 \left[2f_{\text{NL}}^2 P_\phi(k_1) P_\phi(k_2) \right. \\ & \quad \left. \times \{P_\phi(k_{13}) + P_\phi(k_{14}) + P_\phi(k_{23}) + P_\phi(k_{24})\} \right] \delta^D(\mathbf{k}_{1234}). \quad (\text{C.11}) \end{aligned}$$

By combining equation (C.10) and equation (C.11), we calculate the primordial trispectrum induced by local type of non-Gaussianity [Eq. (C.7)]:

$$\begin{aligned}
& T_{\Phi}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) \\
& = 6g_{\text{NL}}P_{\phi}(k_1)P_{\phi}(k_2)P_{\phi}(k_3) + (3 \text{ cyclic}) \\
& \quad + 2f_{\text{NL}}^2P_{\phi}(k_1)P_{\phi}(k_2)\left\{P_{\phi}(k_{13}) + P_{\phi}(k_{14}) + P_{\phi}(k_{23}) + P_{\phi}(k_{24})\right\} + (5 \text{ cyclic}) \\
& = 6g_{\text{NL}}P_{\phi}(k_1)P_{\phi}(k_2)P_{\phi}(k_3) + (3 \text{ cyclic}) \\
& \quad + 2f_{\text{NL}}^2P_{\phi}(k_1)P_{\phi}(k_2)\left\{P_{\phi}(k_{13}) + P_{\phi}(k_{14})\right\} + (11 \text{ cyclic}). \tag{C.12}
\end{aligned}$$

Appendix D

$$P_{gs,22}^{(rest)} \text{ and } P_{gs,13}^{(rest)}$$

The rest term for $P_{gs,22}^{(rest)}(k, \mu; c_1, c_2)$ and $P_{gs,13}^{(rest)}(k, \mu; c_1)$ are defined as terms which cannot be absorbed as either re-definition of bias or previously known terms such as P_m , P_{b2} , P_{b22} , $P_{\delta\theta}$, and $P_{\theta\theta}$. In this appendix, we present the explicit functional form of them. They are

$$\begin{aligned}
& P_{gs,22}^{(rest)}(k, \mu; c_1, c_2) \\
&= 2 \int \frac{d^3 q}{(2\pi)^3} P_L(q) P_L(|\mathbf{k} - \mathbf{q}|) \\
& \times \left[c_1^2 \frac{(fk\mu)^2}{4} \left(\frac{q_z}{q^2} \mathcal{W}_R(|\mathbf{k} - \mathbf{q}|) + \frac{k_z - q_z}{|\mathbf{k} - \mathbf{q}|^2} \mathcal{W}_R(q) \right)^2 + \frac{(fk\mu)^4}{4} \left(\frac{q_z(k_z - q_z)}{q^2 |\mathbf{k} - \mathbf{q}|^2} \right)^2 \right. \\
& + c_2 f \mu^2 \mathcal{W}_R(q) \mathcal{W}_R(|\mathbf{k} - \mathbf{q}|) G_2^{(s)}(\mathbf{q}, \mathbf{k} - \mathbf{q}) \\
& + c_1 c_2 \frac{fk\mu}{2} \mathcal{W}_R(q) \mathcal{W}_R(|\mathbf{k} - \mathbf{q}|) \left(\frac{q_z}{q^2} \mathcal{W}_R(|\mathbf{k} - \mathbf{q}|) + \frac{k_z - q_z}{|\mathbf{k} - \mathbf{q}|^2} \mathcal{W}_R(q) \right) \\
& + c_2 \frac{(fk\mu)^2}{2} \mathcal{W}_R(q) \mathcal{W}_R(|\mathbf{k} - \mathbf{q}|) \left(\frac{q_z(k_z - q_z)}{q^2 |\mathbf{k} - \mathbf{q}|^2} \right) \\
& + c_1^2 f k \mu \mathcal{W}_R(k) F_2^{(s)}(\mathbf{q}, \mathbf{k} - \mathbf{q}) \left(\frac{q_z}{q^2} \mathcal{W}_R(|\mathbf{k} - \mathbf{q}|) + \frac{k_z - q_z}{|\mathbf{k} - \mathbf{q}|^2} \mathcal{W}_R(q) \right) \\
& + c_1 (fk\mu)^2 \mathcal{W}_R(k) F_2^{(s)}(\mathbf{q}, \mathbf{k} - \mathbf{q}) \left(\frac{q_z(k_z - q_z)}{q^2 |\mathbf{k} - \mathbf{q}|^2} \right) \\
& + c_1 f^2 k \mu^3 G_2^{(s)}(\mathbf{q}, \mathbf{k} - \mathbf{q}) \left(\frac{q_z}{q^2} \mathcal{W}_R(|\mathbf{k} - \mathbf{q}|) + \frac{k_z - q_z}{|\mathbf{k} - \mathbf{q}|^2} \mathcal{W}_R(q) \right) \\
& + (f^3 k^2 \mu^4) G_2^{(s)}(\mathbf{q}, \mathbf{k} - \mathbf{q}) \left(\frac{q_z(k_z - q_z)}{q^2 |\mathbf{k} - \mathbf{q}|^2} \right) \\
& \left. + c_1 \frac{(fk\mu)^3}{2} \left(\frac{q_z}{q^2} \mathcal{W}_R(|\mathbf{k} - \mathbf{q}|) + \frac{k_z - q_z}{|\mathbf{k} - \mathbf{q}|^2} \mathcal{W}_R(q) \right) \left(\frac{q_z(k_z - q_z)}{q^2 |\mathbf{k} - \mathbf{q}|^2} \right) \right], \tag{D.1}
\end{aligned}$$

and

$$\begin{aligned}
& P_{gs,13}^{(rest)}(k, \mu; c_1) \\
&= 3(c_1 \mathcal{W}_R(k) + f\mu^2)P_L(k) \int \frac{d^3q}{(2\pi)^3} P_L(q) \\
&\quad \times \left[c_1 \frac{fk\mu}{3} \left\{ \mathcal{W}_R(|\mathbf{k} - \mathbf{q}|) F_2^{(s)}(\mathbf{k}, -\mathbf{q}) \frac{q_z}{q^2} - \mathcal{W}_R(|\mathbf{k} + \mathbf{q}|) F_2^{(s)}(\mathbf{k}, \mathbf{q}) \frac{q_z}{q^2} \right. \right. \\
&\quad \left. \left. + \mathcal{W}_R(q) \left(G_2^{(s)}(\mathbf{k}, -\mathbf{q}) \frac{k_z - q_z}{|\mathbf{k} - \mathbf{q}|^2} + G_2^{(s)}(\mathbf{k}, \mathbf{q}) \frac{k_z + q_z}{|\mathbf{k} + \mathbf{q}|^2} \right) \right\} \right. \\
&\quad \left. + \frac{(fk\mu)^2}{3} \left\{ \left(G_2^{(s)}(\mathbf{k}, -\mathbf{q}) \frac{q_z(k_z - q_z)}{q^2 |\mathbf{k} - \mathbf{q}|^2} - G_2^{(s)}(\mathbf{k}, \mathbf{q}) \frac{q_z(k_z + q_z)}{q^2 |\mathbf{k} + \mathbf{q}|^2} \right) \right. \right. \\
&\quad \left. \left. - c_1 \mathcal{W}_R(k) \frac{(fk\mu)^2}{6} \frac{q_z^2}{q^4} \right\} \right. \\
&\quad \left. - \frac{(fk\mu)^3}{6} \frac{q_z^2 k_z}{q^4 k^2} \right]. \tag{D.2}
\end{aligned}$$

When we implement that, we set $\mathcal{W}_R = 1$ (see, the discussion in Section 2.8.1), then $P_{gs,22}^{(rest)}(k, \mu; c_1, c_2)$ and $P_{gs,13}^{(rest)}(k, \mu; c_1)$ are simplified as

$$\begin{aligned}
P_{gs,22}^{(rest)}(k, \mu, c_1, c_2) &= \sum_{m,n,i,j} \mu^{2m} f^n c_1^i c_2^j \frac{k^3}{(2\pi)^3} \int dr P_L(kr) \int_{-1}^1 dx \\
&\quad \times P_L \left[k(1 + r^2 - 2rx)^{1/2} \right] \frac{A_{mnij}(r, x)}{(1 + r^2 - 2rx)^2} \tag{D.3}
\end{aligned}$$

$$P_{gs,13}^{(rest)}(k, \mu, c_1) = (c_1 + f\mu^2)P_L(k) \sum_{m,n,i} \mu^{2m} f^n c_1^i \frac{k^3}{(2\pi)^3} \int dr P_L(kr) B_{mni}(r), \tag{D.4}$$

where the non-zero $A_{mnij}(r, x)$ and $B_{mni}(r)$ are

$$A_{1120}(r, x) = \frac{1}{7} ((20x^4 - 16x^2 + 3)r^2 - (24x^3 - 10x)r + 7x^2) \quad (D.5)$$

$$A_{1111}(r, x) = -r ((2x^2 - 1)r^3 - (4x^3 - x)r^2 + (4x^2 - 1)r - x) \quad (D.6)$$

$$A_{1101}(r, x) = -\frac{1}{7} ((6x^2 + 1)r^3 - 3x(4x^2 + 3)r^2 + (20x^2 + 1)r - 7x) \quad (D.7)$$

$$A_{1220}(r, x) = \frac{1}{4} (1 - x^2)(1 - 2rx)^2 \quad (D.8)$$

$$A_{1210}(r, x) = \frac{1}{14} (1 - x^2)((10x^2 - 3)r - 7x)r \quad (D.9)$$

$$A_{1201}(r, x) = -\frac{1}{2} (1 - x^2)(r^2 - 2rx + 1)r^2 \quad (D.10)$$

$$A_{2220}(r, x) = \frac{1}{4} (2(6x^4 - 6x^2 + 1)r^2 - (12x^3 - 8x)r + 3x^2 - 1) \quad (D.11)$$

$$A_{2210}(r, x) = \frac{1}{14} ((54x^4 - 27x^2 + 1)r^2 - (81x^3 - 25x)r + 28x^2) \quad (D.12)$$

$$A_{2201}(r, x) = \frac{1}{2} r(-(3x^2 + 1)r^3 + 6x^3r^2 - (7x^2 - 1)r + 2x) \quad (D.13)$$

$$A_{2310}(r, x) = \frac{1}{2} (1 - x^2)((6x^2 - 1)r^2 - 5xr + 1) \quad (D.14)$$

$$A_{2300}(r, x) = \frac{1}{14} (1 - x^2)((6x^2 + 1)r - 7x)r \quad (D.15)$$

$$A_{2400}(r, x) = \frac{3}{16} r^2(x^2 - 1)^2 \quad (D.16)$$

$$A_{3310}(r, x) = \frac{1}{2} ((10x^4 - 9x^2 + 1)r^2 - (11x^3 - 7x)r + 3x^2 - 1) \quad (D.17)$$

$$A_{3300}(r, x) = \frac{1}{14} ((18x^4 - 3x^2 - 1)r^2 - (33x^3 - 5x)r + 14x^2) \quad (D.18)$$

$$A_{3400}(r, x) = \frac{1}{8} (1 - x^2)(3(5x^2 - 1)r^2 - 12xr + 2) \quad (D.19)$$

$$A_{4400}(r, x) = \frac{1}{16} ((35x^4 - 30x^2 + 3)r^2 - 8x(5x^2 - 3)r + 12x^2 - 4) \quad (D.20)$$

and

$$B_{111}(r) = \frac{1}{84} \left[-2(9r^4 - 24r^2 + 19) + \frac{9}{r}(r^2 - 1)^3 \ln \left(\frac{1+r}{|1-r|} \right) \right] \quad (\text{D.21})$$

$$B_{121}(r) = -\frac{1}{3} \quad (\text{D.22})$$

$$B_{120}(r) = -\frac{1}{336r^3} \left[2(-9r^7 + 33r^5 + 33r^3 - 9r) + 9(r^2 - 1)^4 \ln \left(\frac{1+r}{|1-r|} \right) \right] \quad (\text{D.23})$$

$$B_{220}(r) = \frac{1}{336r^3} \left[2r(-27r^6 + 63r^4 - 109r^2 + 9) + 9(3r^2 + 1)(r^2 - 1)^3 \ln \left(\frac{1+r}{|1-r|} \right) \right] \quad (\text{D.24})$$

$$B_{230}(r) = -\frac{1}{3} \quad (\text{D.25})$$

Appendix E

Lagrangian Perturbation Theory and Initial condition for Cosmological N-body Simulation

We generate the initial condition for cosmological N-body simulation by using Lagrangian perturbation theory. In this section, we review the linear and second order Lagrangian perturbation theory and compare the cosmological initial condition generated from two theories.

E.1 Lagrangian perturbation theory formalism

Let us summarize the result of the Lagrangian perturbation theory. The reader can find a review on the subject in Bouchet et al. (1995); Bernardeau et al. (2002).

While Eulerian perturbation theory (Chapter 2) describes the density and velocity fields of matter at a fixed (‘comoving’ in cosmology) coordinate system, Lagrangian perturbation theory concentrates on the trajectory of individual particle. We denote the Eulerian (comoving physical) coordinate \mathbf{x} , and the Lagrangian (comoving initial) coordinate \mathbf{q} . As we define the both coordinate in comoving sense, the expansion of Universe does not change them. In Lagrangian perturbation theory, the dynamical variable is the Lagrangian displacement field $\Psi(\mathbf{q}, \tau)$, which is defined by

$$\mathbf{x}(\tau) = \mathbf{q} + \Psi(\mathbf{q}, \tau). \quad (\text{E.1})$$

Note that $\Psi = 0$ initially so that \mathbf{q} is the same as the usual comoving coordinate at initial time, $\tau = 0$.

The particle trajectory in the expanding universe is governed by the equation of motion:

$$\frac{d^2 \mathbf{x}}{d\tau^2} + \mathcal{H}(\tau) \frac{d\mathbf{x}}{d\tau} = -\nabla_{\mathbf{x}} \Phi, \quad (\text{E.2})$$

where Φ is the peculiar gravitational potential, and

$$\mathcal{H}(\tau) \equiv \frac{1}{a} \frac{da}{d\tau} = a(t)H(t) \quad (\text{E.3})$$

is the modified Hubble parameter. Note that we are using a conformal time τ which is related to the Robertson-Walker coordinate time by $dt = a d\tau$. Taking a divergence of this equation, we get

$$\nabla_{\mathbf{x}} \cdot \left[\frac{d^2 \mathbf{x}}{d\tau^2} + \mathcal{H}(\tau) \frac{d\mathbf{x}}{d\tau} \right] = -\nabla_{\mathbf{x}}^2 \Phi = -\frac{3}{2} \mathcal{H}^2 \Omega_m \delta(\mathbf{x}), \quad (\text{E.4})$$

where $\delta(\mathbf{x})$ is the density contrast, $\delta(\mathbf{x}) \equiv \rho(\mathbf{x})/\bar{\rho} - 1$.

The particle density in the Lagrangian coordinate is the same as the average density of the universe. Therefore, by using the mass conservation, we have

$$\bar{\rho}(\tau) d^3 q = \rho(\mathbf{x}, \tau) d^3 x = \bar{\rho}(\tau) [1 + \delta(\mathbf{x}, \tau)] d^3 x. \quad (\text{E.5})$$

By using the equation above, we can relate the Eulerian density contrast $\delta(\mathbf{x}, \tau)$ to the Lagrangian displacement vector $\Psi(\mathbf{q}, \tau)$ as

$$1 + \delta(\mathbf{x}, \tau) = \left| \frac{d^3 q}{d^3 x} \right| = \frac{1}{J(\mathbf{q}, \tau)}, \quad (\text{E.6})$$

where

$$J(\mathbf{q}, \tau) = \det(\delta_{ij} + \Psi_{i,j}(\mathbf{q}, \tau)), \quad (\text{E.7})$$

is a Jacobian of the Lagrangian to Eulerian coordinate transform. Here, we abbreviate the partial derivative with respect q_j coordinate as $\Psi_{i,j} \equiv \partial \Psi_i / \partial q_j$.

By using equation (E.6), the equation of motion becomes

$$J(\mathbf{q}, \tau) \nabla_{\mathbf{x}} \cdot \left[\frac{d^2 \mathbf{x}}{d\tau^2} + \mathcal{H}(\tau) \frac{d\mathbf{x}}{d\tau} \right] = \frac{3}{2} \mathcal{H}^2(\tau) \Omega_m(\tau) (J - 1). \quad (\text{E.8})$$

Using the chain rule

$$\frac{\partial}{\partial x_i} = \left[\frac{d^3 q}{d^3 x} \right]_{ij} \frac{\partial}{\partial q_j} = [\delta_{ij} + \Psi_{i,j}]^{-1} \frac{\partial}{\partial q_j},$$

the equation for displacement vector Ψ becomes

$$\begin{aligned} J(\mathbf{q}, \tau) [\delta_{ij} + \Psi_{i,j}(\mathbf{q}, \tau)]^{-1} \left[\frac{d^2 \Psi_{i,j}(\mathbf{q}, \tau)}{d\tau^2} + \mathcal{H}(\tau) \frac{d\Psi_{i,j}(\mathbf{q}, \tau)}{d\tau} \right] \\ = \frac{3}{2} \mathcal{H}^2(\tau) \Omega_m(\tau) [J(\mathbf{q}, \tau) - 1]. \end{aligned} \quad (\text{E.9})$$

Equation (E.9) is the master equation of the Lagrangian perturbation theory. In order to get the perturbative solution, we solve the equation perturbatively in $\Psi(\mathbf{q}, \tau)$:

$$\Psi(\mathbf{q}, \tau) = \Psi^{(1)}(\mathbf{q}, \tau) + \Psi^{(2)}(\mathbf{q}, \tau) + \dots \quad (\text{E.10})$$

E.2 Linear Lagrangian perturbation theory

In order to get the linear solution, let's first approximate the Jacobian as

$$J(\mathbf{q}, \tau) = \det[\delta_{ij} + \Psi_{i,j}(\mathbf{q}, \tau)] \simeq 1 + \Psi_{i,i}(\mathbf{q}, \tau), \quad (\text{E.11})$$

and the inverse of the Jacobian matrix as

$$[\delta_{ij} + \Psi_{i,j}]^{-1} \simeq \delta_{ij} - \Psi_{i,j}. \quad (\text{E.12})$$

This approximation is justified by using the matrix approximation of $\det(I + A) = 1 + \text{tr}(A) + \mathcal{O}(A^2)^1$. Using these linear approximations, equation (E.9) becomes

$$(1 + \Psi_{k,k})[\delta_{ij} - \Psi_{i,j}] \left(\frac{d^2 \Psi_{i,j}}{d\tau^2} + \mathcal{H}(\tau) \frac{d\Psi_{i,j}}{d\tau} \right) = \frac{3}{2} \mathcal{H}^2(\tau) \Omega_m(\tau) \Psi_{k,k}, \quad (\text{E.15})$$

and we can find the linear equation for $\Psi^{(1)}$:

$$\frac{d^2 \Psi_{i,i}^{(1)}}{d\tau^2} + \mathcal{H}(\tau) \frac{d\Psi_{i,i}^{(1)}}{d\tau} = \frac{3}{2} \mathcal{H}^2(\tau) \Omega_m(\tau) \Psi_{i,i}^{(1)}. \quad (\text{E.16})$$

Since the dynamical variable of the equation of motion is only time, we can separate the time dependent part as $D_1(\tau)$. Also, in linear approximation,

$$1 + \delta_1(\mathbf{x}, \tau) = J^{-1} \simeq 1 - \Psi_{i,i},$$

therefore, the linear solution becomes

$$\nabla_q \cdot \Psi^{(1)} = -\delta_1(\mathbf{x}, \tau), \quad (\text{E.17})$$

where the time evolution of $\delta_1(\mathbf{x}, \tau)$ is governed by following equation.

$$D_1''(\tau) + \mathcal{H}(\tau) D_1'(\tau) = \frac{3}{2} \mathcal{H}^2(\tau) \Omega_m(\tau) D_1(\tau). \quad (\text{E.18})$$

¹This identity can be proved as follows. Suppose we have a matrix A, whose eigenvalues are small. From the matrix identity

$$\det(C) = \exp[\text{tr}(\ln C)], \quad (\text{E.13})$$

and when $C = I + A$, where I is the identity matrix, we get

$$\begin{aligned} \det(I + A) = \exp[\text{tr}(\ln(I + A))] &= \exp \left[\text{tr} \left(\sum_{n=1}^{\infty} \frac{(-1)^{n-1} A^n}{n} \right) \right] \\ &= 1 + \text{tr} \left(\sum_{n=1}^{\infty} \frac{(-1)^{n-1} A^n}{n} \right) + \text{tr} \left(\sum_{n=1}^{\infty} \frac{(-1)^{n-1} A^n}{n} \right)^2 + \dots \\ &= 1 + \text{tr}(A) + \frac{1}{2} [\text{tr}^2(A) - \text{tr}(A^2)] + \mathcal{O}(A^3). \end{aligned} \quad (\text{E.14})$$

Note that the linear growth function $D_1(\tau)$ here is the same as the linear growth factor in Eulerian perturbation theory [Eq. (2.11)].

Therefore, the particle position in the linear Lagrangian perturbation theory is given by

$$\mathbf{x} = \mathbf{q} - \nabla_q^{-1} \delta_1(\mathbf{x}, \tau), \quad (\text{E.19})$$

and corresponding peculiar velocity is

$$\mathbf{v} \equiv \frac{d\mathbf{x}}{d\tau} = -\mathcal{H} f_1 \nabla_q^{-1} \delta_1(\mathbf{x}, \tau). \quad (\text{E.20})$$

Here, $f_1 \equiv d \ln D_1 / d \ln a$ is a logarithmic derivative of the linear growth factor.

We can further simplify the result by using the curl-free condition of the Lagrangian displacement. The ir-rotational condition $\nabla \times \Psi^{(1)} = 0$ implies the existence of a scalar potential satisfying $\Psi^{(1)}(\mathbf{q}, \tau) = -\nabla_q \phi^{(1)}(\mathbf{q}, \tau)$. Therefore, in terms of the linear Lagrangian potential $\phi^{(1)}$, which is related to the linear density field as

$$\nabla_q \cdot \Psi^{(1)}(\mathbf{q}, \tau) = -\nabla_q^2 \phi^{(1)}(\mathbf{q}, \tau) = -\delta_1(\mathbf{x}, \tau), \quad (\text{E.21})$$

the linear solution can be also written as

$$\mathbf{x}(\mathbf{q}, \tau) = \mathbf{q} - \nabla_q \phi^{(1)}(\mathbf{q}, \tau) \quad (\text{E.22})$$

$$\mathbf{v}(\mathbf{q}, \tau) = -\mathcal{H} f_1 \nabla_q \phi^{(1)}(\mathbf{q}, \tau). \quad (\text{E.23})$$

E.2.1 Zel'dovich approximation

The Zel'dovich approximation (Zel'dovich, 1970) extrapolates the linear solution [Eq. (E.22) and Eq. (E.23)] into the non-linear regime. As a consequence of such an approximation, the theory predict what's called *Zel'dovich pancake*.

In this context, the Jacobian matrix is often called the ‘tensor of deformation’:

$$\mathfrak{D}_{ij} \equiv \delta_{ij} + \Psi_{i,j}(\mathbf{q}, \tau). \quad (\text{E.24})$$

We can find the three eigenvectors of the deformation tensor \mathfrak{D}_{ij} and using these eigensystem as a basis (with eigenvalues $-\alpha$, $-\beta$, and $-\gamma$ ²), we can diagonalize³ it as

$$J(\mathbf{q}, \tau) = \left\| \begin{array}{ccc} 1 - \alpha D_1(\tau) & 0 & 0 \\ 0 & 1 - \beta D_1(\tau) & 0 \\ 0 & 0 & 1 - \gamma D_1(\tau) \end{array} \right\|, \quad (\text{E.25})$$

²Negative eigenvalues correspond to the growing mode (positive $\delta(\mathbf{x}, \tau)$) See, for example, equation (E.17).

³This diagonalization is justified because, we assume the ir-rotational perturbation; using a scalar potential $\phi^{(1)}$, \mathfrak{D}_{ij} is symmetric (i.e. diagonalizable).

where $D_1(\tau)$ is the linear growth function. By using the three eigenvalues, we can write the density contrast as [Eq. (E.6)]

$$\delta(\mathbf{x}, \tau) = [(1 - D_1(\tau)\alpha)(1 - D_1(\tau)\beta)(1 - D_1(\tau)\gamma)]^{-1} - 1. \quad (\text{E.26})$$

What is the consequence of it? In order to see that, without loss of generality, let α be the largest eigenvalues of deformation tensor. For example, for the ellipsoidal shape initial perturbation, α corresponds to the shortest axis. Initially, $D_1(\tau)$ is small, but it gets bigger as linear perturbation grows, and eventually, reaches $1/\alpha$. At that moment, the density of α -direction become infinity, and this stage is called *Zeldovich pancake*. Physically, it mean that every sheet element sliced perpendicular to α -direction finally merging at one point. In reality, however, Zeldovich approximation breaks down before the volume element reaches its infinite density stage. It is because for the particles within such a small volume element, we have to take into account the pressure effect as well as shell crossing.

E.3 Second order Lagrangian perturbation theory (2LPT)

Up to second order, the Jacobian is approximated as [Eq. (E.14)]

$$J \simeq 1 + \Psi_{k,k}^{(1)} + \Psi_{k,k}^{(2)} + \frac{1}{2} \left[\left(\Psi_{k,k}^{(1)} \right)^2 - \Psi_{i,j}^{(1)} \Psi_{j,i}^{(1)} \right], \quad (\text{E.27})$$

therefore equation of motion,

$$J [\delta_{ij} - \Psi_{i,j}] \left(\frac{d^2 \Psi_{i,j}}{d\tau^2} + \mathcal{H} \frac{d\Psi_{i,j}}{d\tau} \right) = \frac{3}{2} \mathcal{H}^2 \Omega_m (J - 1), \quad (\text{E.28})$$

becomes

$$\begin{aligned} & \left(\frac{d^2 \Psi_{i,i}^{(2)}}{d\tau^2} + \mathcal{H} \frac{d\Psi_{i,i}^{(2)}}{d\tau} \right) + \Psi_{k,k}^{(1)} \left(\frac{d^2 \Psi_{i,i}^{(1)}}{d\tau^2} + \mathcal{H} \frac{d\Psi_{i,i}^{(1)}}{d\tau} \right) - \Psi_{i,j}^{(1)} \left(\frac{d^2 \Psi_{i,j}^{(1)}}{d\tau^2} + \mathcal{H} \frac{d\Psi_{i,j}^{(1)}}{d\tau} \right) \\ &= \frac{3}{2} \mathcal{H}^2 \Omega_m \left[\Psi_{k,k}^{(2)} + \frac{1}{2} \left(\Psi_{k,k}^{(1)} \right)^2 - \frac{1}{2} \Psi_{i,j}^{(1)} \Psi_{j,i}^{(1)} \right]. \end{aligned} \quad (\text{E.29})$$

Using the solution for the linear displacement field $\Psi^{(1)}(\mathbf{x}, \tau)$, we simplify the equation of motion as

$$\left(\frac{d^2 \Psi_{i,i}^{(2)}}{d\tau^2} + \mathcal{H} \frac{d\Psi_{i,i}^{(2)}}{d\tau} - \frac{3}{2} \mathcal{H}^2 \Omega_m \Psi_{i,i}^{(2)} \right) = -\frac{3}{2} \mathcal{H}^2 \Omega_m \left[\frac{1}{2} \left(\Psi_{k,k}^{(1)} \right)^2 - \frac{1}{2} \Psi_{i,j}^{(1)} \Psi_{j,i}^{(1)} \right]. \quad (\text{E.30})$$

Here, we use the following symmetry: $\Psi_{i,j}^{(1)} = \Psi_{j,i}^{(1)} = -\phi_{,ij}^{(1)}$. The equation of motion above is, again, separable, because the spatial derivatives only appears as divergence. Let's denote

the time dependent part of $\Psi^{(2)}$ as $D_2(\tau)$. The second order time evolution of $D_2(\tau)$ is governed by

$$D_2''(\tau) + \mathcal{H}D_2'(\tau) - \frac{3}{2}\mathcal{H}^2\Omega_m D_2(\tau) = -\frac{3}{2}\mathcal{H}^2\Omega_m [D_1(\tau)]^2. \quad (\text{E.31})$$

Note that, in flat Λ CDM universe, $D_2(\tau) \simeq -3D_1^2(\tau)\Omega_m^{-1/143}/7$ approximates $D_2(\tau)$ better than 0.6 per cent (Bouchet et al. 1995).

The space part of the second order solution describes the effect of gravitational tide,

$$\frac{\Psi_{k,k}^{(2)}(\mathbf{q}, \tau)}{D_2(\tau)} = \frac{1}{D_1^2(\tau)} \frac{1}{2} \sum_{i \neq j} \left\{ \Psi_{i,i}^{(1)}(\mathbf{q}, \tau) \Psi_{j,j}^{(1)}(\mathbf{q}, \tau) - \Psi_{i,j}^{(1)}(\mathbf{q}, \tau) \Psi_{j,i}^{(1)}(\mathbf{q}, \tau) \right\}. \quad (\text{E.32})$$

By using a second order scalar potential $\Psi^{(2)}(\mathbf{q}, \tau) = \nabla_{\mathbf{q}} \phi^{(2)}(\mathbf{q}, \tau)$, the equation becomes

$$\nabla_{\mathbf{q}}^2 \phi^{(2)}(\mathbf{q}, \tau) = \frac{D_2(\tau)}{D_1^2(\tau)} \frac{1}{2} \sum_{i \neq j} \left\{ \phi_{,ii}^{(1)}(\mathbf{q}, \tau) \phi_{,jj}^{(1)}(\mathbf{q}, \tau) - \phi_{,ij}^{(1)}(\mathbf{q}, \tau) \phi_{,ji}^{(1)}(\mathbf{q}, \tau) \right\} \quad (\text{E.33})$$

$$\simeq -\frac{3}{7}\Omega_m^{-1/143} \sum_{i > j} \left\{ \phi_{,ii}^{(1)}(\mathbf{q}, \tau) \phi_{,jj}^{(1)}(\mathbf{q}, \tau) - \left[\phi_{,ij}^{(1)}(\mathbf{q}, \tau) \right]^2 \right\}. \quad (\text{E.34})$$

Using the scalar potential, the solution for the position and velocity up to the second order are

$$\mathbf{x}(\mathbf{q}, \tau) = \mathbf{q} - \nabla_{\mathbf{q}} \phi^{(1)}(\mathbf{q}, \tau) + \nabla_{\mathbf{q}} \phi^{(2)}(\mathbf{q}, \tau) \quad (\text{E.35})$$

$$\mathbf{v}(\mathbf{q}, \tau) = \frac{d\mathbf{x}}{d\tau} = -\mathcal{H}f_1 \nabla_{\mathbf{q}} \phi^{(1)}(\mathbf{q}, \tau) + \mathcal{H}f_2 \nabla_{\mathbf{q}} \phi^{(2)}(\mathbf{q}, \tau). \quad (\text{E.36})$$

The logarithmic derivatives of the growth factor $f_i \equiv d \ln D_i / d \ln a$ is well approximated as

$$f_1 \simeq [\Omega_m(z)]^{5/9}, \quad f_2 \simeq 2 [\Omega_m(z)]^{6/11}$$

to better than 10 and 12 percent, respectively, for flat- Λ CDM universe with $0.01 < \Omega_m < 1$. The accuracy of these two fits improves significantly for $\Omega_m \geq 0.1$ (Bouchet et al. 1995).

E.4 Generating initial condition using Linear solution

In this section, we present the way how to generate the initial condition by using the linear, and second order Lagrangian perturbation theory solutions. For the detailed normalization, and the way to impose the Hermitian condition in Fourier space, see, Appendix A.

The solutions of Lagrangian perturbation theory is summarized as below: The second order particle position and velocity are given by

$$\mathbf{x}(\mathbf{q}, \tau) = \mathbf{q} - \nabla_{\mathbf{q}}\phi^{(1)}(\mathbf{q}, \tau) + \nabla_{\mathbf{q}}\phi^{(2)}(\mathbf{q}, \tau) \quad (\text{E.37})$$

$$\mathbf{v}(\mathbf{q}, \tau) = \frac{d\mathbf{x}}{d\tau} = -\mathcal{H}f_1\nabla_{\mathbf{q}}\phi^{(1)}(\mathbf{q}, \tau) + \mathcal{H}f_2\nabla_{\mathbf{q}}\phi^{(2)}(\mathbf{q}, \tau), \quad (\text{E.38})$$

where the linear and the second order Lagrangian potential is

$$\nabla_{\mathbf{q}}^2\phi^{(1)}(\mathbf{q}, \tau) = \delta_1(\mathbf{x}, \tau), \quad (\text{E.39})$$

and

$$\nabla_{\mathbf{q}}^2\phi^{(2)}(\mathbf{q}, \tau) \simeq -\frac{3}{7}[\Omega_m(\tau)]^{-1/143} \sum_{i>j} \left\{ \phi_{,ii}^{(1)}(\mathbf{q}, \tau)\phi_{,jj}^{(1)}(\mathbf{q}, \tau) - \left[\phi_{,ij}^{(1)}(\mathbf{q}, \tau) \right]^2 \right\}, \quad (\text{E.40})$$

respectively. Here, $\mathcal{H}(\tau)$ is a reduced Hubble parameter

$$\mathcal{H}(z) = \frac{H_0}{1+z} \sqrt{\Omega_\Lambda + \Omega_m(1+z)^3}, \quad (\text{E.41})$$

and $f_i \equiv d \ln D_i / d \ln a$ is the logarithmic derivative of growth factor, which can be approximated as

$$f_1 \simeq [\Omega_m(z)]^{5/9}, \quad f_2 \simeq 2[\Omega_m(z)]^{6/11}, \quad (\text{E.42})$$

for flat- Λ CDM universe. For different cosmology, one has to solve the linear and non-linear growth equations: [Eq. (E.18)] and [Eq. (E.31)].

Let us suppose that we want to generate the initial condition for N^3 matter particles inside of cubic box of volume L^3 at redshift z . The initial density field follows the Gaussian statistics, and the linear power spectrum is given $P_L(k, z)$. Then, the procedure is as follow.

- (1) Imagine we divide the cubic box into N^3 regular grid points. Those are the Lagrangian coordinate \mathbf{q} .
- (2) As Lagrangian displacement vector is determined by the density contrast, we need to generate the Gaussian random density contrast. From the definition of power spectrum

$$\langle \delta_1(\mathbf{k}, z) \delta_1(\mathbf{k}', z) \rangle = (2\pi)^3 P_L(k, z) \delta^D(\mathbf{k} + \mathbf{k}'),$$

we relate the density contrast to the power spectrum as

$$\langle |\delta(\mathbf{k}, z)|^2 \rangle = \frac{(2\pi)^3}{k_F^3} P(k, z) = V P(k, z), \quad (\text{E.43})$$

where we use $\delta^D(\mathbf{k} + \mathbf{k}') = \delta_{\mathbf{n}_k, \mathbf{n}_{k'}}^k / k_F^3$ with the fundamental frequency $k_F = (2\pi)/L$. Here, δ_{ij}^k is the Kronecker delta.

Generate density contrast in the Fourier space $\delta(\mathbf{k}) \equiv \delta_r(\mathbf{k}) + i\delta_i(\mathbf{k})$ as random variables obeying a Gaussian statistics with mean and variance given by:

$$\langle \delta_r(\mathbf{k}) \rangle = 0, \quad \langle \delta_r^2(\mathbf{k}) \rangle = \sqrt{\frac{VP(k)}{2}} \quad (\text{E.44})$$

$$\langle \delta_i(\mathbf{k}) \rangle = 0, \quad \langle \delta_i^2(\mathbf{k}) \rangle = \sqrt{\frac{VP(k)}{2}}. \quad (\text{E.45})$$

When generating the random variable, we have to explicitly impose the Hermitian condition of $\delta(\mathbf{k})$:

$$\delta(-\mathbf{k}) = \delta^*(\mathbf{k}). \quad (\text{E.46})$$

If one want to calculate the real space density field, do the inverse Fourier transform by

$$\delta(\mathbf{q}) = \frac{1}{V} \sum \delta(\mathbf{k}) e^{i\mathbf{k} \cdot \mathbf{q}} = \frac{1}{V} \delta^{\text{FFTW}}(\mathbf{q}). \quad (\text{E.47})$$

(3) The linear Lagrangian potential $\phi^{(1)}$ is given by

$$-k^2 \phi^{(1)}(\mathbf{k}, z) = \delta_1(\mathbf{k}, z). \quad (\text{E.48})$$

Therefore, calculate the linear Lagrangian displacement $\Psi^{(1)}$ in Fourier space by

$$\Psi^{(1)}(\mathbf{k}, z) = -i\mathbf{k} \phi^{(1)}(\mathbf{k}, z) = i\mathbf{k} \frac{\delta_1(\mathbf{k}, z)}{k^2}. \quad (\text{E.49})$$

(4) Inverse Fourier transform to get the displacement field in real space.

$$\Psi^{(1)}(\mathbf{q}, z) = \frac{1}{V} \sum i\mathbf{k} \frac{\delta_1(\mathbf{k}, z)}{k^2} e^{i\mathbf{k} \cdot \mathbf{q}}. \quad (\text{E.50})$$

Then, move particles at each grid point \mathbf{q} by the displacement vector at that point.

$$\mathbf{x} = \mathbf{q} + \Psi^{(1)}(\mathbf{q}, z) \quad (\text{E.51})$$

Also, assign the velocity as

$$\mathbf{v} = f_1(z) \mathcal{H}(z) \Psi^{(1)}(\mathbf{q}, z). \quad (\text{E.52})$$

Now, we have generated the Zeldovich initial condition.

- (5) In order to calculate the initial condition by using second order Lagrangian perturbation theory (2LPT), we have to calculate the second order Lagrangian potential $\phi^{(2)}$ from equation (E.40). We shall solve this equation in Fourier space with following order.

- (5)-1 Calculate the Fourier transform of $\phi_{,ij}^{(1)}(\mathbf{q}, \tau)$ in Fourier space:

$$-k_i k_j \phi^{(1)}(\mathbf{k}, \tau) = \frac{k_i k_j \delta_1(\mathbf{k}, \tau)}{k^2}. \quad (\text{E.53})$$

- (5)-2 Inverse Fourier transform to the real space, now we have six $\phi_{,ij}^{(1)}(\mathbf{q}, z)$ s. Calculate the right hand side of equation (E.40) as

$$F(\mathbf{q}) = \left[\phi_{,22}^{(1)}(\mathbf{q}, z) \phi_{,11}^{(1)}(\mathbf{q}, z) + \phi_{,33}^{(1)}(\mathbf{q}, z) \phi_{,22}^{(1)}(\mathbf{q}, z) + \phi_{,33}^{(1)}(\mathbf{q}, z) \phi_{,11}^{(1)}(\mathbf{q}, z) \right. \\ \left. - \left[\phi_{,32}^{(1)}(\mathbf{q}, z) \right]^2 - \left[\phi_{,31}^{(1)}(\mathbf{q}, z) \right]^2 - \left[\phi_{,21}^{(1)}(\mathbf{q}, z) \right]^2 \right]. \quad (\text{E.54})$$

- (5)-3 Fourier transform $F(\mathbf{q})$. Then, $\phi^{(2)}(\mathbf{k}, z)$ is

$$\phi^{(2)}(\mathbf{k}, z) = \frac{3}{7} [\Omega_m(z)]^{-1/143} \frac{F(\mathbf{k}, z)}{k^2}, \quad (\text{E.55})$$

and we calculate the second order Lagrangian displacement $\Psi^{(2)}$ by

$$\Psi^{(2)}(\mathbf{k}, z) = i\mathbf{k}\phi^{(2)}(\mathbf{k}, z) = \frac{3}{7} [\Omega_m(z)]^{-1/143} \frac{i\mathbf{k}F(\mathbf{k}, z)}{k^2}. \quad (\text{E.56})$$

- (5)-4 Inverse Fourier transform to get the second order displacement field in real space, and move particle further by $\Psi^{(2)}(\mathbf{q}, z)$

$$\mathbf{x} = \mathbf{q} + \Psi^{(1)}(\mathbf{q}, z) + \Psi^{(2)}(\mathbf{q}, z). \quad (\text{E.57})$$

Finally, assign the velocity

$$\mathbf{v} = f_1(z)\mathcal{H}(z)\Psi^{(1)}(\mathbf{q}, z) + f_2(z)\mathcal{H}(z)\Psi^{(2)}(\mathbf{q}, z). \quad (\text{E.58})$$

Now, we have generated the second order Lagrangian initial condition.

E.5 Starting redshifts, initial condition generators, and convergence tests

How early in redshift should one start N -body simulations? Is the usual first-order Lagrangian perturbation theory, which is traditionally known as the Zel'dovich approximation (ZA) (e.g., Efstathiou et al., 1985), accurate enough for generating initial conditions

for N -body simulations? Will ZA converge as one increases the starting redshift? In this Appendix we show that ZA converges only very slowly: even $z_{start} = 400$ leaves an artificial suppression of power at the level of $\sim 1\%$ relative to a more accurate initial condition generated by the second-order Lagrangian perturbation theory (2LPT). This suppression is persistent at all the redshifts we have studied, $z = 1 - 6$. On the other hand, simulations starting from 2LPT initial conditions at $z_{start} = 300$ show convergence at $z < 6$, i.e., simulations starting at $z_{start} = 400$ give very similar results at $z < 6$.

Crocce et al. (2006) have shown that ZA-generated initial conditions yield an artificial suppression of power spectrum measured from N -body simulations. They call this effect the “transient effect”. The lower the starting redshift, z_{start} , is, the larger the artificial suppression of power becomes. This is a very important systematic error and must be taken into account when one is interested in making precision predictions for the power spectrum (Jeong & Komatsu, 2006) as well as for the mass function of dark matter halos (Lukić et al., 2007).

The transient effects occur when decaying modes are excited artificially by inaccurate initial condition generators. Although these decaying modes decay, they decay only slowly as a^{-1} when ZA is used to generate initial conditions (Crocce et al., 2006). For example, the transient effect reported in Crocce et al. (2006) is about 4% at $k = 1 \ h \text{ Mpc}^{-1}$ at $z = 3$ for a simulation starting at $z_{start} = 49$, and decays only slowly toward lower redshifts. The error at this level is unacceptable for testing precision calculations of the power spectrum (Jeong & Komatsu, 2006; McDonald, 2007; Matarrese & Pietroni, 2007).

Crocce et al. (2006) also show that simulations starting from 2LPT initial conditions, which are more accurate than ZA, still yield transient effects with the opposite sign: there is an artificial amplification of power. However, an advantage of 2LPT is that the transient modes decay much more quickly than those from ZA.

A natural question then arises: would ZA perform better as z_{start} is raised, and if so, how large should z_{start} be? The same question would apply to 2LPT as well. To answer these questions, we ran 10 simulations with 5 different starting redshifts, $z_{start} = 50, 100, 150, 300, \text{ and } 400$, whose initial conditions were generated from either ZA or 2LPT. We have used a publicly-available 2LPT initial condition generator developed by Roman Scoccimarro⁴ to generate initial conditions for N -body simulations. We have then used the **Gadget-2** code to evolve density fields. The cosmological parameters are exactly the same as those used in

⁴<http://cosmo.nyu.edu/roman/2LPT>

the main body of this paper. We chose to run simulations with a $(200 h^{-1} \text{ Mpc})^3$ box and 512^3 particles for these runs. Finally, we have used the same initial random seed for all of these runs to facilitate head-to-head comparison.

Figure E.1 shows $P(k)$ measured from 5 ZA runs with $z_{start} = 50, 100, 150, 300$, and 400 (bottom to top), divided by $P(k)$ from a 2LPT run with $z_{start} = 300$. In all cases the power is suppressed relative to 2LPT, and the suppression is persistent at all redshifts from $z = 6$ to 1. The amount of suppression decreases only slowly as we raise z_{start} .

Figure E.2 shows $P(k)$ measured from 5 2LPT runs with $z_{start} = 50, 100, 150, 300$, and 400 (top to bottom), divided by $P(k)$ from a 2LPT run with $z_{start} = 300$. The situation is reversed: the transient effects amplify the power, but the amount of amplification decays very quickly with z and z_{start} . We conclude that the transient effect is unimportant ($< 1\%$) at $z < 6$, if initial conditions are generated at $z_{start} = 300$ using 2LPT.

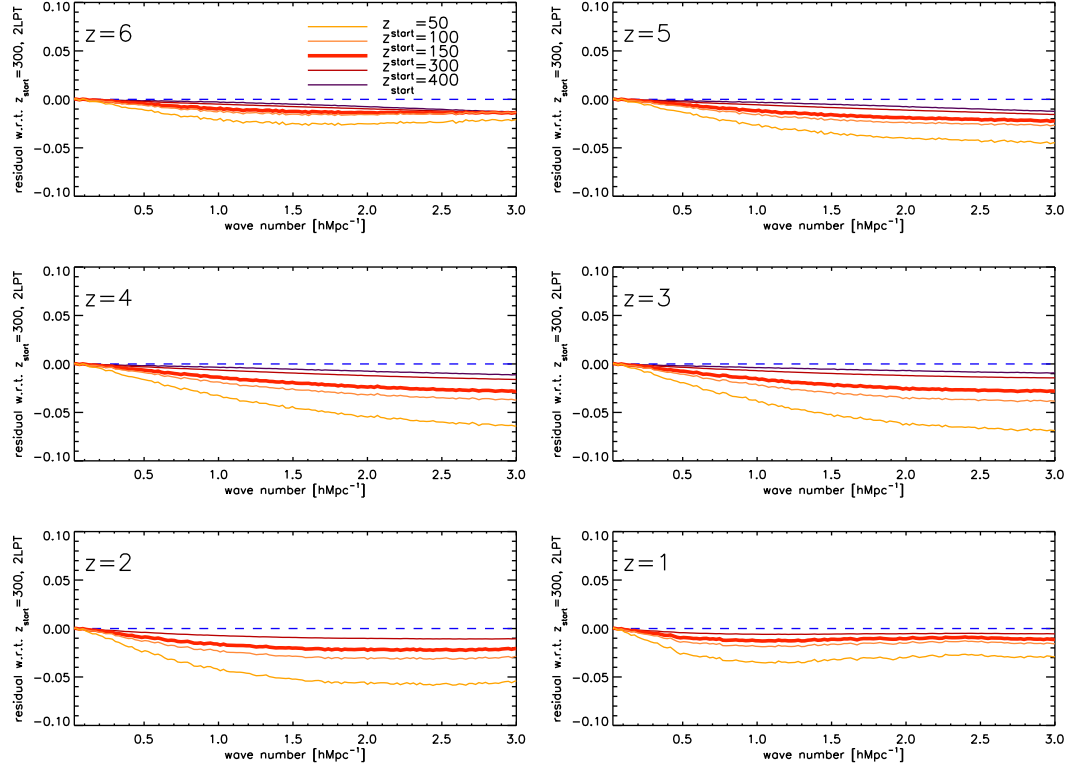


Figure E.1: Comparison between the power spectra calculated from 5 different ZA runs using different starting redshifts, $z_{start} = 50, 100, 150, 300$, and 400 (from bottom to top). The power spectra are divided by the one from 2LPT with $z_{start} = 300$, to facilitate comparison.

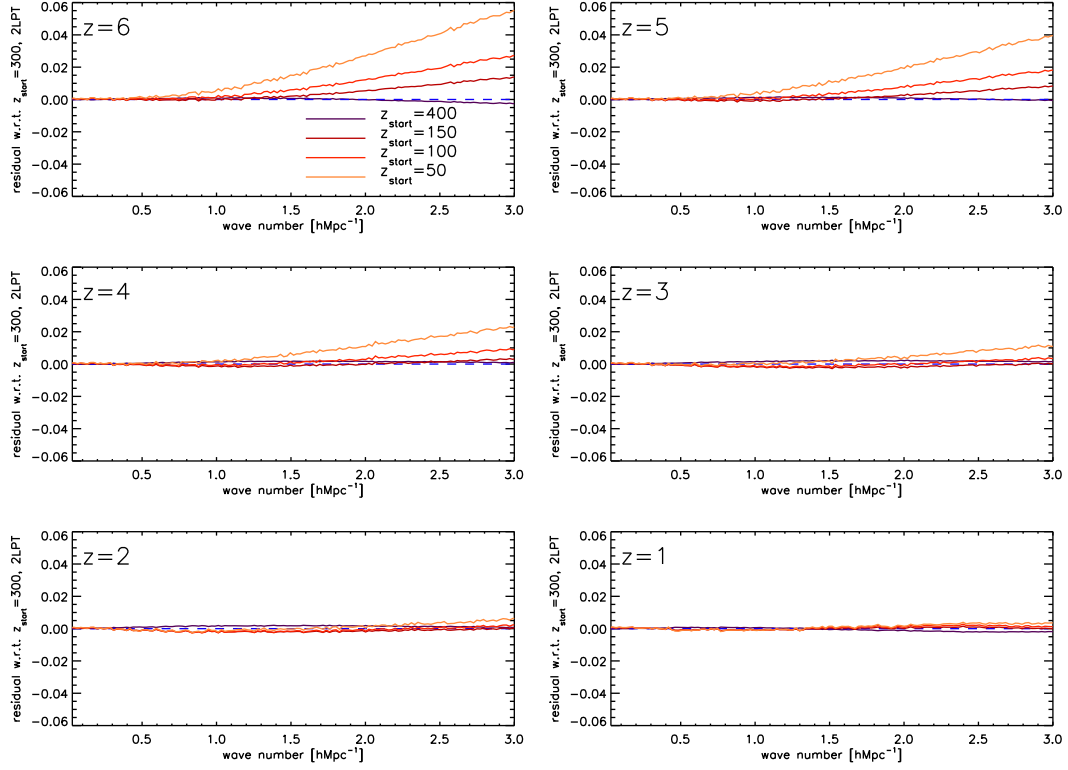


Figure E.2: Comparison between the power spectra calculated from 5 different 2LPT runs using different starting redshifts, $z_{start} = 50, 100, 150, 300$, and 400 (from bottom to top). The power spectra are divided by the one from 2LPT with $z_{start} = 300$, to facilitate comparison. Note that the power spectra with $z_{start} = 300$ and 400 agree very well, which suggests convergence at $z_{start} = 300$.

Appendix F

Error on power spectrum and bispectrum

F.1 Error on power spectrum

Besides the normalization, an estimator for the power spectrum may be written as

$$P_{obs}(k) = \frac{1}{N_k} \sum_{i=1}^{N_k} |\delta(\mathbf{k}_i)|^2 \Big|_{|\mathbf{k}_i - k| \leq \Delta k} \quad (\text{F.1})$$

where $\delta(\mathbf{k}_i)$ is a Fourier transform of the density field in position space, Δk is the fundamental wavenumber of either survey volume or simulation box, and N_k is the number of independent k -modes available per bin. This estimator is unbiased because

$$\langle P_{obs}(k) \rangle = \frac{1}{N_k} \sum_{i=1}^{N_k} \langle |\delta(\mathbf{k}_i)|^2 \rangle \Big|_{|\mathbf{k}_i - k| \leq \Delta k} = \langle |\delta(k)|^2 \rangle = P(k), \quad (\text{F.2})$$

where $P(k)$ is the underlying power spectrum. The variance of this estimator is given by

$$\left\langle \left(\frac{P_{obs}(k) - P(k)}{P(k)} \right)^2 \right\rangle = 1 - 2 \frac{\langle P_{obs} \rangle}{P(k)} + \frac{1}{N_k^2 P(k)^2} \sum_{i=1}^{N_k} \sum_{j=1}^{N_k} \langle \delta^*(\mathbf{k}_i) \delta(\mathbf{k}_i) \delta^*(\mathbf{k}_j) \delta(\mathbf{k}_j) \rangle. \quad (\text{F.3})$$

Assuming that the density field is a Gaussian random variable with its variance given by $P(k)$, i.e.,

$$\langle \delta_i^* \delta_j \rangle = P(k) \delta_{ij}, \quad (\text{F.4})$$

we use Wick's theorem for evaluating the last double summation:

$$\begin{aligned} \sum_{i=1}^{N_k} \sum_{j=1}^{N_k} \langle \delta_i^* \delta_i \delta_j^* \delta_j \rangle &= \sum_{i=1}^{N_k} \sum_{j=1}^{N_k} [\langle \delta_i^* \delta_i \rangle \langle \delta_j^* \delta_j \rangle + \langle \delta_i^* \delta_j \rangle \langle \delta_j^* \delta_i \rangle + \langle \delta_i^* \delta_j^* \rangle \langle \delta_i \delta_j \rangle] \\ &= N_k^2 [P(k)]^2 + N_k [P(k)]^2. \end{aligned} \quad (\text{F.5})$$

Therefore, the variance is given by

$$\left\langle [P_{obs}(k) - P(k)]^2 \right\rangle = \frac{[P(k)]^2}{N_k}, \quad (\text{F.6})$$

and the standard deviation is given by

$$\sigma_{P(k)} \equiv \left\langle [P_{obs}(k) - P(k)]^2 \right\rangle^{1/2} = \sqrt{\frac{1}{N_k}} P(k). \quad (\text{F.7})$$

Note that this formula is valid only when δ is a Gaussian random field. When δ is non-Gaussian due to, e.g., non-linear evolution, primordial non-Gaussianity, non-linear bias, etc., we must add the connected four-point function to Eq. (F.5). See also Takahashi et al. (2008) for the study of finite box size effects on the four-point function.

How do we calculate N_k ? As the Fourier transformation of a real-valued field has symmetry given by $\delta^*(\mathbf{k}) = \delta(-\mathbf{k})$, the number of independent k -modes is exactly a half of the number of modes available in a spherical shell at a given k . We find

$$N_k = \frac{1}{2} \frac{4\pi k^2 \delta k}{(\delta k)^3} = 2\pi \left(\frac{k}{\delta k} \right)^2, \quad (\text{F.8})$$

where δk is the fundamental wavenumber given by $\delta k = 2\pi/L$, where L is the survey size or simulation box size.

In the literature one may often find a different formula such as

$$\sigma_{P(k)}^{literature} = \sqrt{\frac{2}{N_k^{literature}}} P(k). \quad (\text{F.9})$$

Here, there is an extra factor of $\sqrt{2}$, as $N_k^{literature}$ is the number of modes available in a spherical shell at a given k , *without symmetry*, $\delta^*(\mathbf{k}) = \delta(-\mathbf{k})$, *taken into account*, i.e., $N_k^{literature} = 2N_k$. Both formulas give the same results, provided that we understand what we mean by N_k in these formulas.

We have tested the formula Eq. (F.7) by comparing it to the standard deviation estimated the ensemble of dark matter simulations used in Chapter 3. Figure F.1 and F.2 show the result of this comparison. The formula Eq. (F.7) and the simulation data agree well.

F.2 Error on bispectrum

Similarly, we can calculate the expected Gaussian error on the measurement of the bispectrum. Let us suppose that we estimate the bispectrum from a cubic box of size L . As the fundamental frequency is $k_F \equiv 2\pi/L$, we can estimate the bispectrum for a triangle with

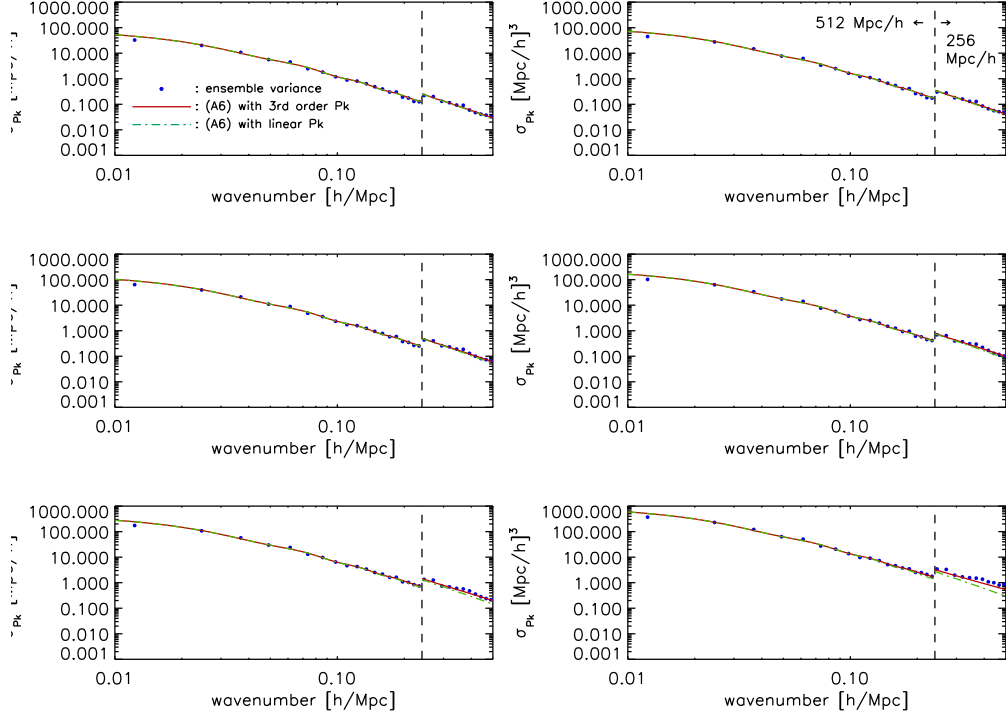


Figure F.1: Standard deviation of the matter power spectrum: analytical versus simulations. The symbols show the standard deviations directly measured from 120 independent N -body simulations whose box sizes are $L = 512$ Mpc/h (60 realizations for $k < 0.24h/\text{Mpc}$) and $L = 256$ Mpc/h (60 realizations for $0.24 < k < 0.5h/\text{Mpc}$). Each simulation contains 256^3 particles. The solid and dot-dashed lines show the analytical formula (Eq. (F.7)) with the 3rd-order PT non-linear $P(k)$ and the linear $P(k)$, respectively. Note that the graph is discontinuous at $k = 0.24h/\text{Mpc}$ because the number of k modes, N_k , for a given wavenumber k is different for different box sizes.

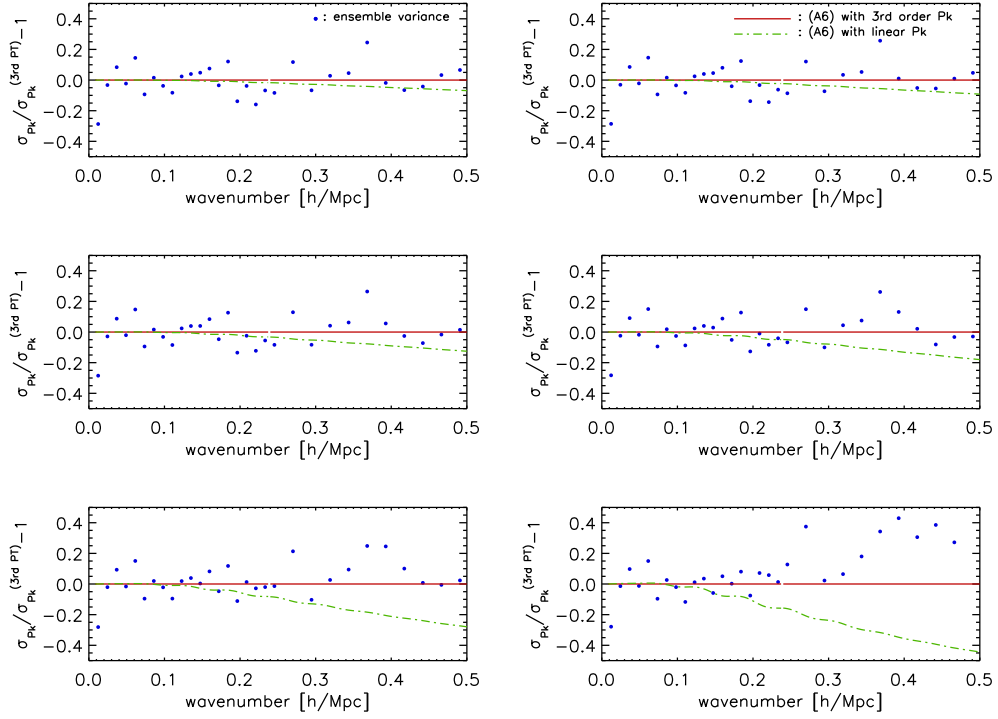


Figure F.2: Residuals. We divide both analytical estimation and simulation results by the analytical formula (Eq. (F.7)) with the 3rd-order PT nonlinear $P(k)$.

side of $k_F(n_1, n_2, n_3)$, where n_i are integers. From the unbiased estimator of the bispectrum (See, Chapter 7 for this estimator)

$$B_{obs}(n_1 k_F, n_2 k_F, n_3 k_F) = \frac{V^2}{N^9} \left(\frac{1}{N_t} \sum_{\mathbf{m} \in \text{Tri}_{123}} \delta^{FTW}(\mathbf{m}_1) \delta^{FTW}(\mathbf{m}_2) \delta^{FTW}(\mathbf{m}_3) \right), \quad (\text{F.10})$$

one can estimate the errors on calculating the bispectrum as

$$\begin{aligned} & \left\langle \left(\frac{\Delta B}{B}(n_1, n_2, n_3) \right)^2 \right\rangle \\ &= 1 - 2 \left\langle \frac{B_{obs}}{B} \right\rangle + \frac{V^4}{N^{18}} \frac{1}{H^{18}} \frac{1}{N_t^2 B^2} \sum_{\mathbf{q} \in \text{Tri}_{123}} \sum_{\mathbf{q}' \in \text{Tri}_{123}} \langle \delta(\mathbf{q}_1) \delta(\mathbf{q}_2) \delta(\mathbf{q}_3) \delta(\mathbf{q}'_1) \delta(\mathbf{q}'_2) \delta(\mathbf{q}'_3) \rangle, \end{aligned} \quad (\text{F.11})$$

where N_t is the total number of triangle with side of $(k_1, k_2, k_3) = k_F(n_1, n_2, n_3)$, and N is the size of one dimensional Fourier grid (therefore, we perform the Fourier transform on grid of size N^3 .)

Let's calculate the six point correlator in equation (F.11). At the lowest order, the six point correlator is given by the sum of the product of three power spectra and two bispectra, which are related to $\delta(\mathbf{q}_i)$ as (See, Chapter 7 for the coefficients)

$$\langle \delta(\mathbf{q}_i) \delta(\mathbf{q}_j) \rangle = \frac{(2\pi)^3}{k_F^3} P(k_i) \delta_{ij}^k$$

and

$$\langle \delta(\mathbf{q}_i) \delta(\mathbf{q}_j) \delta(\mathbf{q}_k) \rangle = \frac{(2\pi)^3}{k_F^3} B(k_{ijk}) \delta_{ijk}^k.$$

Here, δ_{ijk}^k denotes the Kronecker delta. That is, the six point correlator is given by

$$\begin{aligned} & \sum_{\mathbf{q} \in \text{Tri}_{123}} \sum_{\mathbf{q}' \in \text{Tri}_{123}} \langle \delta(\mathbf{q}_1) \delta(\mathbf{q}_2) \delta(\mathbf{q}_3) \delta(\mathbf{q}'_1) \delta(\mathbf{q}'_2) \delta(\mathbf{q}'_3) \rangle \\ &= N_t^2 \frac{(2\pi)^6}{k_F^6} B^2(k_{123}) + s_B N_t \frac{(2\pi)^9}{k_F^9} P(k_1) P(k_2) P(k_3) \\ &+ \mathcal{O}(P(k_i) T(k_l, k_m, k_n, k_o)) + \mathcal{O}(H(k_1, k_2, k_3, k_1, k_2, k_3)) \\ &\simeq N_t^2 V^2 B^2(k_{123}) + s_B N_t V^3 P(k_1) P(k_2) P(k_3), \end{aligned} \quad (\text{F.12})$$

where we approximate the equation by only taking the 'Gaussian' part. Here, T and H denote the trispectrum, and hexaspectrum generated by non-linear gravitational interaction. The symmetric factor s_B is 6, 2, 1 for equilateral, isosceles and general triangles, respectively.

Using this approximation, we calculate the variance of bispectrum

$$\langle [B_{obs}(k_{123}) - B(k_{123})]^2 \rangle = s_B \frac{V}{N_t} P(k_1)P(k_2)P(k_3) = s_B \frac{(2\pi)^3}{N_t k_F^3} P(k_1)P(k_2)P(k_3) \quad (\text{F.13})$$

Finally, the standard deviation is given by

$$\sigma_B(k_{123}) = \sqrt{\frac{s_B V}{N_t}} [P(k_1)P(k_2)P(k_3)]^{1/2} = \sqrt{\frac{s_B (2\pi)^3}{N_t k_F^3}} [P(k_1)P(k_2)P(k_3)]^{1/2}. \quad (\text{F.14})$$

Note that this result is consistent with Equation (28) of Sefusatti & Komatsu (2007). In that paper, variance on bispectrum is given by

$$(\Delta B)^2 \simeq k_F^3 \frac{s_{123}}{V_B} P(k_1)P(k_2)P(k_3) \quad (\text{F.15})$$

where

$$V_B \equiv \int_{k_1} d^3 q_1 \int_{k_2} d^3 q_2 \int_{k_3} d^3 q_3 \delta_D(\mathbf{q}_{123}) \simeq 8\pi^2 k_1 k_2 k_3 (\Delta k)^3 \quad (\text{F.16})$$

measures 9-dimensional volume within which $\mathbf{q}_1 + \mathbf{q}_2 + \mathbf{q}_3 = 0$ is satisfied. Note that their Fourier convention is different from the convention we are using in this dissertation (See Appendix A); there is $(2\pi)^3$ factor difference in the integral measure in k -space. In this convention, the number of triangles (N_t) is given by dividing the volume V_B by the volume of fundamental cell:

$$N_t = \frac{V_B}{k_F^6}, \quad (\text{F.17})$$

and we rewrite the variance in Sefusatti & Komatsu (2007) as

$$(\Delta B)^2 \simeq \frac{s_{123}}{N_t k_F^3} P(k_1)P(k_2)P(k_3). \quad (\text{F.18})$$

One can easily show that the factor of $(2\pi)^3$ difference comes from the difference in the definition of power spectrum and bispectrum. They define them as

$$\langle \delta(\mathbf{k}_1) \delta(\mathbf{k}_2) \rangle = P(\mathbf{k}_1) \delta^D(\mathbf{k}_{12}) \quad (\text{F.19})$$

$$\langle \delta(\mathbf{k}_1) \delta(\mathbf{k}_2) \delta(\mathbf{k}_3) \rangle = B(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) \delta^D(\mathbf{k}_{123}). \quad (\text{F.20})$$

Appendix G

Analytical marginalization of the likelihood function over \tilde{b}_1^2 and P_0

In this appendix we derive the analytical formulas for the likelihood function marginalized over \tilde{b}_1^2 and P_0 .

The likelihood function, Eq. (4.6), is given by

$$\mathcal{L}(\tilde{b}_1, \tilde{b}_2, P_0, \theta_n) = \left(\prod_i \frac{1}{\sqrt{2\pi\sigma_{P_i}^2}} \right) \exp \left[- \sum_i \frac{\left(P_{obs,i} - \tilde{b}_1^2(P_{\delta\delta,i} + \tilde{b}_2 P_{b2,i} + \tilde{b}_2^2 P_{b22,i}) - P_0 \right)^2}{2\sigma_{P_i}^2} \right], \quad (\text{G.1})$$

where θ_n are the cosmological parameters that do not depend on any of the bias parameters. The subscript i denotes bins, k_i .

Integrating the likelihood function over P_0 , we find

$$\begin{aligned} \mathcal{L}(\tilde{b}_1, \tilde{b}_2, \theta_n) &= \int_{-\infty}^{\infty} dP_0 \mathcal{L}(\tilde{b}_1, \tilde{b}_2, P_0, \theta_n) \\ &= \mathcal{N} \sqrt{\frac{2\pi}{\sum_i w_i}} \exp \left[- \frac{1}{2} \frac{\sum_{i>j} w_i w_j (a_j - a_i)^2}{\sum_i w_i} \right], \end{aligned} \quad (\text{G.2})$$

where we have defined new variables

$$\mathcal{N} \equiv \prod_i \frac{1}{\sqrt{2\pi\sigma_{P_i}^2}} \quad (\text{G.3})$$

$$a_i \equiv P_{obs,i} - \tilde{b}_1^2(P_{\delta\delta,i} + \tilde{b}_2 P_{b2,i} + \tilde{b}_2^2 P_{b22,i}) \quad (\text{G.4})$$

$$w_i \equiv \frac{1}{\sigma_{P_i}^2}. \quad (\text{G.5})$$

We then integrate this function over \tilde{b}_1^2 . Introducing new variables given by

$$\bar{\mathcal{N}} \equiv \mathcal{N} \sqrt{\frac{2\pi}{\sum_i w_i}}, \quad (\text{G.6})$$

$$P_{th,i} \equiv P_{\delta\delta,i} + \tilde{b}_2 P_{b2,i} + \tilde{b}_2^2 P_{b22,i}, \quad (\text{G.7})$$

and $a_i = P_{obs,i} - \tilde{b}_1^2 P_{th,i}$, we rewrite Eq. (G.2) as

$$\begin{aligned}\mathcal{L}(\tilde{b}_1, \tilde{b}_2, \theta_n) &= \bar{N} \exp \left[-\frac{1}{2} \frac{\sum_{i>j} w_i w_j \left\{ (P_{th,i} - P_{th,j}) \tilde{b}_1^2 - (P_{obs,i} - P_{obs,j}) \right\}^2}{\sum_i w_i} \right] \\ &= \bar{N} \exp \left[-\frac{1}{2} (A \tilde{b}_1^4 - 2B \tilde{b}_1^2 + C) \right],\end{aligned}\tag{G.8}$$

where

$$A \equiv \frac{\sum_{i>j} w_i w_j (P_{th,i} - P_{th,j})^2}{\sum_i w_i}\tag{G.9}$$

$$B \equiv \frac{\sum_{i>j} w_i w_j (P_{th,i} - P_{th,j})(P_{obs,i} - P_{obs,j})}{\sum_i w_i}\tag{G.10}$$

$$C \equiv \frac{\sum_{i>j} w_i w_j (P_{obs,i} - P_{obs,j})^2}{\sum_i w_i}.\tag{G.11}$$

Assuming a flat prior on \tilde{b}_1^2 , we integrate the likelihood function to find the desired result:

$$\begin{aligned}\mathcal{L}(\tilde{b}_2, \theta_n) &= \bar{N} \int_0^\infty \exp \left[-\frac{1}{2} (A \tilde{b}_1^4 - 2B \tilde{b}_1^2 + C) \right] d(\tilde{b}_1^2) \\ &= \bar{N} \exp \left[\frac{B^2 - AC}{2A} \right] \sqrt{\frac{\pi}{2A}} \left\{ 1 + \operatorname{erf} \left(\frac{B}{\sqrt{2A}} \right) \right\}.\end{aligned}\tag{G.12}$$

Note that the convergence of the likelihood function is ensured by Cauchy's inequality, $B^2 - AC < 0$.

Appendix H

distribution of errors on the distance scale

We find that the error on D extracted from the halo power spectrum of Millennium Simulation is about 2.17% for $k_{max} = 1.5 \ h/\text{Mpc}$ at $z = 6$. (See Figure 4.16.) On the other hand, the error on D calculated from the Fisher information matrix is 1.57%. Are they consistent?

In order to test whether it is possible to get the error on D far from the value derived from the Fisher matrix, we generate 1000 realizations of mock power spectra with the best-fitting bias parameters for halo with $k_{max} = 1.5 \ h/\text{Mpc}$ at $z = 6$. Then, we calculate the best-fitting value of D as well as the $1-\sigma$ (68.27% CL) range from the one-dimensional marginalized likelihood function of D for each realization.

We find that the mean $1-\sigma$ error on D calculated from these realizations is 1.66%, and their standard deviation is 0.43%. Figure H.1 shows the distribution of the fractional $1-\sigma$ error on D compared with D_{true} . While the error derived from the Fisher matrix is close to the mean, the error calculated from the Millennium Simulation is on the tail of the distribution. The probability of having an error on D greater than that from the Millennium Simulation is about 6%, which is acceptable.

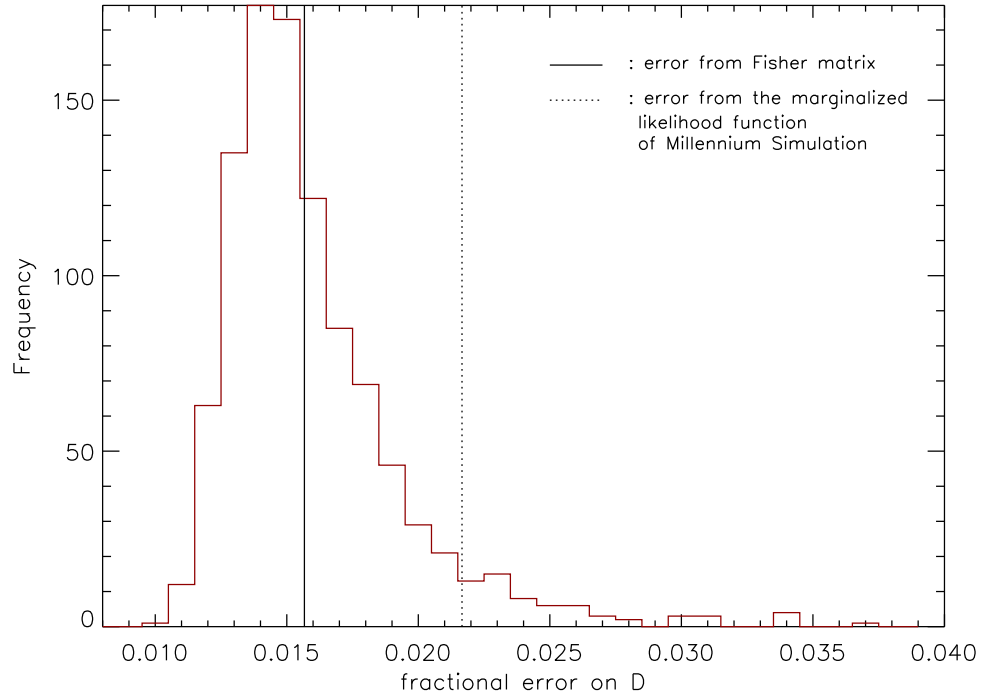


Figure H.1: Histogram for the $1\text{-}\sigma$ errors on D calculated from 1000 Monte Carlo realizations generated with the best-fitting bias parameters of halo power spectrum of Millennium Simulation with $k_{max} = 1.5 \text{ } h/\text{Mpc}$ at $z = 6$. The error derived from the Fisher matrix is close to the mean, while the error from the marginalized one-dimensional likelihood function of Millennium Simulation is on the tail of the distribution. The probability of having an error on D greater than that from the Millennium Simulation is about 6%.

Appendix I

Theory on statistics of density peaks: Mass function and halo bias

I.1 Mass function of halos

I.1.1 Variables

First, we summarize the variables we use to describe the mass function of halos.

σ_R^2 : The variance of the smoothed linear matter density contrast is defined as

$$\sigma_R^2(z) = \int \frac{d^3k}{(2\pi)^3} P_L(k, z) W^2(k; R) = \int \frac{dk}{k} \Delta_L^2(k, \tau) W^2(k; R), \quad (\text{I.1})$$

where $P_L(k, z) = P_L(k) D^2(z)$ is the linear power spectrum, and $\Delta_L^2(k, z) = k^3 P_L(k, z) / (2\pi^2)$ is the dimensionless linear power spectrum. The smoothing is usually done by the following two k-space Filters;

– **Gaussian filter**

$$W_G(k; R) = \exp \left[-\frac{R^2 k^2}{2} \right] \quad (\text{I.2})$$

– **Top-Hat filter**

$$W_{TH}(k; R) = 3 \frac{\sin x - x \cos x}{x^3}, \quad x \equiv kR \quad (\text{I.3})$$

We can explicitly separate the time dependence of $\sigma_R^2(z)$ in terms of the linear growth function, $D(z)$.

$$\sigma_R^2(z) = D^2(z) \sigma_R^2, \quad (\text{I.4})$$

where σ_R^2 is the value at present.

$\delta_c(\Omega_M)$: In the spherical collapse model, a spherical clump becomes a virialized object when linear density contrast of the clump extrapolated to the moment reaches the critical over-density, δ_c . In Einstein de-Sitter (flat, matter-dominated) Universe, $\delta_c \equiv$

$0.15(12\pi)^{2/3} \simeq 1.686$ regardless of the time when halo forms. In flat- Λ CDM universe, the critical over-density changes as Ω_m changes, and it is approximately given in Kitayama & Suto (1996):

$$\delta_c(z) \simeq \frac{3(12\pi)^{2/3}}{20} [1 + 0.0123 \log_{10} \Omega_m(z)], \quad (\text{I.5})$$

In order to avoid confusion between time dependence due to Λ and time dependence due to growth factor (in random walk view), we refer to the former as $\delta_c(\Omega_m)$.

R, M, ν : In this chapter, we shall use these three quantities interchangeably. They are related by

$$\frac{m}{1M_\odot} = \frac{4\pi}{3} (2.7754 \times 10^{11}) \Omega_m h^2 \left(\frac{r}{1\text{Mpc}} \right)^3 \quad (\text{I.6})$$

$$\nu = \frac{\delta_c}{\sigma_R} = \frac{\delta_c}{\sigma_m}, \quad (\text{I.7})$$

where σ_R and δ_c are given above, and σ_m is defined as $\sigma_{R(m)}$.

I.1.2 The random walk view

A region of radius R becomes a halo of mass M , when the linear density contrast δ_1 of that region grows and exceeds the critical value δ_c . Therefore, the statistics of halos are determined by comparing the growth of linear density contrast $\delta_1(z)$ to a fixed critical value δ_c .

The same process can also be described by the alternative view. In this view, instead of following the growth of the linear density field (we call it *physical view*), we freeze the density field at the fixed time (usually at present time) and change the critical density as $\delta_c(\Omega_m)/D(z)$, so that the relative ratio between the density field and the critical density remains same (we call it *random walk view*). The excursion set approach is based on this view. For the details of this view, see Bond et al. (1991) and Lacey & Cole (1993).

Formally, these two views can be written as

$$\delta_{h,R}(\mathbf{x}, z_f) = \Theta[\delta_R(\mathbf{x}, z_f) - \delta_c(\Omega_{m,f})] = \Theta[\delta_R(\mathbf{x}) - \delta_c(\Omega_{m,f})/D(z_f)] \quad (\text{I.8})$$

with a step function, Θ , and a formation redshift of halo, z_f . Here, $\delta_R(\mathbf{x}, z)$ is the smoothed linear density field (value at present without z dependence). Note that, for high mass halos, we usually assume that observed redshift z_o is very close to the formation redshift: $z_f \simeq z_o$.

Although these two views give essentially the same answer for the halo mass function, random walk view is more convenient to calculate the conditional mass function; thus, quantities which is based on it, like a halo merger rate, halo bias, etc. Therefore, we shall use the random walk view in this appendix whenever excursion set approach is used.

I.1.3 Press-Schechter mass function

As smoothed linear density field follows Gaussian statistics,

$$P_R(\delta) = \frac{1}{\sqrt{2\pi\sigma_R^2}} \exp\left[-\frac{\delta^2}{2\sigma_R^2}\right], \quad (\text{I.9})$$

we calculate the fraction of collapsed objects with mass larger than $m \equiv 4/3\bar{\rho}\pi R^3$ by integrating the PDF:

$$F(> m) = \int_{\delta_c}^{\infty} P_R(\delta) d\delta = \frac{1}{2} \text{erfc}\left[\frac{\delta_c}{\sqrt{2}\sigma_R}\right]. \quad (\text{I.10})$$

Using this, the fraction of collapsed objects within mass range between $m \sim m + \delta m$ is

$$f(m)dm \equiv F(> m) - F(> m + \delta m) = -\frac{\partial F}{\partial m} dm. \quad (\text{I.11})$$

We calculate the comoving number density of the collapsed object within that mass range as

$$\begin{aligned} n(m)dm &= \frac{\text{number of smoothed patch} \times f(m)}{V} = \frac{V/(M/\bar{\rho})f(m)dm}{V} \\ &= -\bar{\rho} \frac{\partial F}{\partial m} \frac{dm}{m}. \end{aligned} \quad (\text{I.12})$$

Multiplying the Cloud-In-Cloud *magic factor* 2, (Bond et al., 1991) we find

$$n_{PS}(m)dm = -2\bar{\rho} \frac{\partial F}{\partial m} \frac{dm}{m}. \quad (\text{I.13})$$

Changing the parameter from δ to $x \equiv \delta/\sigma_R$ and with new parameter $\nu \equiv \delta_c/\sigma_R$,

$$F = \int_{\nu}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{x^2}{2}\right] dx. \quad (\text{I.14})$$

Therefore,

$$\frac{\partial F}{\partial m} = \frac{\partial F}{\partial \nu} \frac{\partial \nu}{\partial m} = \frac{1}{\sqrt{2\pi}} \nu e^{-\nu^2/2} \frac{d \ln \sigma_R}{dm}, \quad (\text{I.15})$$

and the PS mass function is now

$$n_{PS}(m)mdm = -2\bar{\rho}\nu \frac{e^{-\nu^2/2}}{\sqrt{2\pi}} \frac{d\sigma_R}{\sigma_R}, \quad (\text{I.16})$$

or

$$n_{PS}(m) = -2 \frac{\bar{\rho}}{m^2} \nu \frac{e^{-\nu^2/2}}{\sqrt{2\pi}} \frac{d \ln \sigma_R}{d \ln m}. \quad (\text{I.17})$$

Sometimes, people define $f(\nu)$ as

$$n(m) m dm = -2 \bar{\rho} f(\nu) d \ln \sigma_R = 2 \bar{\rho} f(\nu) d \ln \nu. \quad (\text{I.18})$$

Then,

$$f_{PS}(\nu) = \sqrt{\frac{1}{2\pi}} \nu e^{-\nu^2/2}. \quad (\text{I.19})$$

I.1.3.1 ST convention

Sheth and Tormen use different definition of ν . They set $\nu = (\delta_c/\sigma_R)^2$. In this case,

$$F = \int_{\sqrt{\nu}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{x^2}{2} \right] dx, \quad (\text{I.20})$$

and with the following partial differentiation,

$$\frac{\partial F}{\partial m} = \frac{\partial F}{\partial \nu} \frac{\partial \nu}{\partial m} = -\frac{1}{\sqrt{2\pi}} e^{-\nu/2} \frac{d\sqrt{\nu}}{dm} = -\frac{1}{\sqrt{2\pi}} e^{-\nu/2} \frac{1}{2\sqrt{\nu}} \frac{d\nu}{dm}, \quad (\text{I.21})$$

we get

$$n_{PS}(m) dm = \frac{\bar{\rho}}{m} \frac{1}{\sqrt{2\pi\nu}} e^{\nu/2} d\nu. \quad (\text{I.22})$$

Defining $f(\nu)$ as

$$\nu f(\nu) \equiv m^2 \frac{n(m)}{\bar{\rho}} \frac{d \ln m}{d \ln \nu}, \quad (\text{I.23})$$

leads

$$\nu f(\nu) = \left(\frac{\nu}{2} \right)^{1/2} \frac{e^{-\nu/2}}{\sqrt{\pi}}. \quad (\text{I.24})$$

I.1.4 Mass function from excursion set approach

Some halos reside in the bigger halo, and others are isolated. In the excursion set approach (Bond et al., 1991), we can formulate isolated halo of mass m which is just formed at redshift z (mass function) as “trajectories that are above the threshold $\delta_c(z)$ at some mass scale $m \sim m + \delta m$ but are below this threshold for all larger values of m ”. Let’s calculate the probability of having those halos.

First, we consider a region which has overdensity $\delta(< \delta_c(z))$ after smoothing with a filter of mass m . Then, probability that this region does NOT exceed the critical density $\delta_c(z)$ for any larger smoothing scale is given by

$$q(\delta, \delta_c; m, z)d\delta = \frac{1}{\sqrt{2\pi\sigma_m^2}} \left[\exp\left(-\frac{\delta^2}{2\sigma_m^2}\right) - \exp\left(-\frac{(2\delta_c(z) - \delta)^2}{2\sigma_m^2}\right) \right] d\delta.$$

It is easy to prove this equation because for every random walk path ends up δ which passed the *barrier* of $\delta_c(z)$ in the past, there exists a unique path ends up $2\delta_c(z) - \delta$ which has the same probability as the former path by reflective symmetry. Therefore, the chance that the patch does not exceed $\delta_c(z)$ for larger smoothing scale is simply

$$\begin{aligned} & [\text{Prob. of random walk path ends up } \delta] \\ & - [\text{Prob. of random walk path ends up } 2\delta_c(z) - \delta], \end{aligned}$$

which is basically the equation above.

The probability that a trajectory hit the barrier at redshift z between filtering scale m and $m + \delta m$ has to equal the reduction of trajectories which does NOT hit the barrier. That is,

$$\begin{aligned} f(m, z)d\sigma_m^2 &= -\frac{\partial}{\partial\sigma_m^2} \left[\int_{-\infty}^{\delta_c(z)} q(\delta, \delta_c; m, z)d\delta \right] d\sigma_m^2 \\ &= -\frac{\partial}{\partial\sigma_m^2} \left[\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\nu} dx \left(e^{-x^2/2} - e^{-(x-2\nu)^2/2} \right) \right] d\sigma_m^2 \\ &= -\frac{\partial\nu}{\partial\sigma_m^2} \frac{\partial}{\partial\nu} \left[\frac{2}{\sqrt{2\pi}} \int_0^{\nu} dx e^{-x^2/2} \right] d\sigma_m^2 \\ &= \frac{1}{\sqrt{2\pi}} \frac{\delta_c}{\sigma_m^3} \exp\left[-\frac{\delta_c^2}{2\sigma_m^2}\right] d\sigma_m^2. \end{aligned} \tag{I.25}$$

$f(m, z)$ represents the fraction of mass associated with halos in range $m \sim m + \delta m$. Therefore the comoving number density of halos of mass m at redshift z (mass function) is

$$n_{es}(m)dm = \frac{\bar{\rho}}{m} f(m, z) \left| \frac{d\sigma_m^2}{dm} \right| dm = -2 \frac{\bar{\rho}}{m^2} \nu \frac{e^{-\nu^2/2}}{\sqrt{2\pi}} \frac{d \ln \sigma_R}{d \ln m} dm. \tag{I.26}$$

where subscript ‘es’ stands for ‘excursion set approach’. Note that the resulting mass function is the same as the Press-Schechter mass function, but we do not need a *magic factor* of 2!

I.1.5 Halo merger rate

I.1.5.1 Conditional mass-function

One can also calculate the conditional mass function in the onset of the Press-Schechter formalism. Conditional mass function $f(m_1, z_1 | m_0, z_0)$ is the conditional probability that the halo of mass m_0 at later time $z = z_0$ was found as the halo with mass m_1 at earlier time $z = z_1$. Note that $z_0 < z_1$ and $m_0 > m_1$, which leads $\delta_c(z_0) < \delta_c(z_1)$ and $\sigma_0^2 < \sigma_1^2$.

In the previous section, we calculate that the probability of forming a halo of mass m at redshift z by using an excursion set approach. Note that this probability is the result from analyzing random walk trajectories from $m = 0$, $\delta = 0$.

$$f(m, z)dm = \frac{1}{\sqrt{2\pi}} \frac{\delta_c(z)}{\sigma^3(m)} \exp \left[-\frac{\delta_c^2(z)}{2\sigma^2(m)} \right] \frac{d\sigma^2(m)}{dm} dm. \quad (\text{I.27})$$

Because the random walk trajectories are translationally invariant, we can set up the similar equation for a trajectory starting from $m = m_0$ and $\delta = \delta_c(z_0)$. Especially, the probability that this trajectory reaches the bar of $\delta_c(z_1)$ for the first time at $m = m_1$ is the conditional probability we want to calculate.

$$\begin{aligned} & f(m_1, z_1 | m_0, z_0) d\sigma^2(m_1) \\ &= \frac{1}{\sqrt{2\pi}} \frac{\delta_c(z_1) - \delta_c(z_0)}{[\sigma^2(m_1) - \sigma^2(m_0)]^{3/2}} \exp \left\{ -\frac{[\delta_c(z_1) - \delta_c(z_0)]^2}{2[\sigma^2(m_1) - \sigma^2(m_0)]} \right\} d\sigma^2(m_1). \end{aligned} \quad (\text{I.28})$$

Remember that we use σ^2 as a time step of random walk.

To simplify the notation, we use put subscript 0 for a quantity related to the ‘later time’ and subscript 1 for a quantity related to the ‘earlier time’. Plus, density contrast with subscript has to be understood as a ‘critical density’ at that epoch, e.g. $\delta_1 = \delta_c(z_1)$.

I.1.5.2 Halo merger rate

To get the merger rate, we had to invert the conditional mass function using Bayes theorem:

$$f(m_0, z_0 | m_1, z_1) d\sigma_0^2 = \frac{f(m_1, z_1 | m_0, z_0) d\sigma_1^2 f(m_0, z_0) d\sigma_0^2}{f(m_1, z_1) d\sigma_1^2}. \quad (\text{I.29})$$

Therefore,

$$\begin{aligned}
& f(m_0, z_0 | m_1, z_1) \\
&= \frac{1}{\sqrt{2\pi}} \frac{\delta_1 - \delta_0}{(\sigma_1^2 - \sigma_0^2)^{3/2}} \exp \left\{ -\frac{(\delta_1 - \delta_0)^2}{2(\sigma_1^2 - \sigma_0^2)} \right\} \frac{\delta_0}{\sigma_0^3} \exp \left(-\frac{\delta_0^2}{2\sigma_0^2} \right) \frac{\sigma_1^3}{\delta_1} \exp \left(\frac{\delta_1^2}{2\sigma_1^2} \right) \\
&= \frac{1}{\sqrt{2\pi}} \left[\frac{\sigma_1^2}{\sigma_0^2(\sigma_0^2 - \sigma_1^2)} \right]^{3/2} \frac{\delta_0(\delta_1 - \delta_0)}{\delta_1} \exp \left[-\frac{(\delta_0\sigma_1^2 - \delta_1\sigma_0^2)^2}{2\sigma_0^2\sigma_1^2(\sigma_1^2 - \sigma_0^2)} \right].
\end{aligned} \tag{I.30}$$

This is the probability that a halo of mass m_1 at redshift z_1 will merge later and by the time z_0 have mass between m_0 and $m_0 + \delta m_0$. In the language of excursion set approach, merger event corresponds to the sudden decrease of $\sigma^2(m)$, equivalently sudden increase of m , in time. Since the parameter controls the time is $\delta_c(z)$, this can be achieved by setting $\delta_0 \rightarrow \delta_1$. Then, the using the equation (I.30), we can define the probability of the merger (*i.e.* sudden increase in mass) from m_1 to m_0 , equivalently from σ_1^2 to σ_0^2 , within time interval $d\delta(z)$, which is also called ‘mean transition rate’, as

$$\frac{d^2 p}{d\sigma_0^2 d\delta} d\delta d\sigma_0^2 = \frac{1}{\sqrt{2\pi}} \left[\frac{\sigma_1^2}{\sigma_0^2(\sigma_0^2 - \sigma_1^2)} \right]^{3/2} \exp \left[-\frac{\delta_0^2(\sigma_1^2 - \sigma_0^2)}{2\sigma_0^2\sigma_1^2} \right] d\delta_0 d\sigma_0^2. \tag{I.31}$$

Finally, we can get the merger rate by changing the variables from (δ, σ^2) to (t, m) .

$$\begin{aligned}
\frac{d^2 p}{d(\ln \Delta m) dt} &= \left| \frac{d\sigma_0^2}{dm} \right| \Delta m \left| \frac{d\delta}{dt} \right| \frac{d^2 p}{d\sigma_0^2 d\delta} \\
&= \left(\frac{2}{\pi} \right)^{1/2} \frac{1}{t} \left| \frac{d \ln \delta_c(z)}{d \ln t} \right| \left(\frac{\Delta m}{m_0} \right) \left| \frac{d \ln \sigma_0}{d \ln m_0} \right| \frac{\delta_c(z)}{\sigma_0} \frac{1}{(1 - \sigma_0^2/\sigma_1^2)^{3/2}} \\
&\quad \times \exp \left[-\frac{\delta_c(z)^2}{2} \left(\frac{1}{\sigma_0^2} - \frac{1}{\sigma_1^2} \right) \right]
\end{aligned} \tag{I.32}$$

I.1.6 Sheth-Tormen mass function

Sheth & Tormen (1999) refines the PS mass function in order to fit the mass function from GIF/Virgo simulation. Their modification of the PS mass function is

$$f_{ST}(\nu) = A \sqrt{\frac{\alpha \nu^2}{2\pi}} \left[1 + \frac{1}{(\alpha \nu^2)^p} \right] \exp \left(-\frac{\alpha \nu^2}{2} \right), \tag{I.33}$$

where $\alpha = 0.707$, $p = 0.3$, and A is determined by requiring that

$$\int n_{ST}(m) m dm = \bar{\rho} = 2 \int_0^\infty \bar{\rho} f_{ST}(\nu) d \ln \nu. \tag{I.34}$$

Therefore,

$$A = \frac{1}{1 + \frac{\Gamma(0.5-p)}{2^p \sqrt{\pi}}} \simeq 0.322184. \tag{I.35}$$

Later, Sheth et al. (2001) show that this mass function may be related to a ellipsoidal collapse. Note that, in the special case of $A = 1/2$, $p = 0$, and $\alpha = 1$, ST mass function is reduced to PS mass function.

I.1.6.1 ST convention

Using Sheth-Tormen definition of $\nu \equiv (\delta_c/\sigma_R)^2$

$$\nu f(\nu) = A \left(1 + \frac{1}{(\alpha\nu)^p} \right) \left(\frac{\alpha\nu}{2} \right)^{1/2} \frac{e^{-\alpha\nu/2}}{\sqrt{\pi}} \quad (\text{I.36})$$

I.2 Halo bias with Gaussian linear density field

I.2.1 Kaiser (1984)

The first paper which calculate the linear bias factor is Kaiser (1984). We summarize his analysis here. Let's assume that halos form where “the primordial density enhancement, when averaged over a rich cluster mass scale, lay above some moderately high threshold”. Then, “at large separations, the cluster correlation function is amplified by a factor, $A \equiv \xi_c/\xi_\rho$ which depends only on the statistical distribution of the smoothed density fluctuations.”

Let's think about a density field $\delta_R(\mathbf{x})$ which is smoothed with a window radius R_W . We denote the threshold as $\nu\sigma$, where $\sigma^2 = \xi(0)$ is the rms density fluctuation. We calculate $\xi_{>\nu}(r)$ which is “the correlation function of *regions* lying above the threshold”, which is closely related to what we want to calculate, $\xi_{\max>\nu}(r)$. “ $\xi_{>\nu}(r)$ is defined to be the fractional excess probability that $\delta_2 \equiv \delta(\mathbf{x}_2) > \nu\sigma$ given $\delta_1 \equiv \delta(\mathbf{x}_1) > \nu\sigma$ and $|\mathbf{x}_1 - \mathbf{x}_2| > r$.”

I.2.1.1 Gaussian density field

Let's do the analysis for Gaussian case first. Then, for any point, the probability of having a density field of size y is

$$P(y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{y^2}{2\sigma^2}},$$

and the probability of having a density field of size y_1 at \mathbf{x}_1 and y_2 at \mathbf{x}_2 , whose separation is r , becomes

$$P(y_1, y_2) = \frac{1}{2\pi\sqrt{|C|}} \exp \left[-\frac{1}{2} y^T C^{-1} y \right],$$

where

$$y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

and the covariant matrix can be calculated by using there auto-correlation and correlation:

$$C = \begin{pmatrix} \xi(0) & \xi(r) \\ \xi(r) & \xi(0) \end{pmatrix}.$$

By using the p.d.f. above, probability that a randomly chosen point lies above the threshold becomes

$$P_1 = \int_{\nu\sigma}^{\infty} P(y)dy = \frac{1}{\sqrt{2\pi}} \int_{\nu}^{\infty} e^{-x^2/2} dx = \frac{1}{2} \operatorname{erfc} \left[\frac{\nu}{\sqrt{2}} \right]$$

and the probability that δ_1 and δ_2 both lie above the threshold becomes

$$\begin{aligned} P_2 &= \int_{\nu\sigma}^{\infty} dy_1 \int_{\nu\sigma}^{\infty} dy_2 P(y_1, y_2) \\ &= \int_{\nu\sigma}^{\infty} dy_1 \int_{\nu\sigma}^{\infty} dy_2 \frac{1}{2\pi\sqrt{\xi^2(0) - \xi^2(r)}} \exp \left[-\frac{\xi(0)y_1^2 + \xi(0)y_2^2 - 2\xi(r)y_1y_2}{2[\xi^2(0) - \xi^2(r)]} \right] \\ &= \frac{1}{2\pi\sqrt{1 - \alpha^2}} \int_{\nu}^{\infty} dx_1 \int_{\nu}^{\infty} dx_2 \exp \left[-\frac{x_1^2 + x_2^2 - 2\alpha x_1x_2}{2(1 - \alpha^2)} \right] \\ &= \sqrt{\frac{1}{8\pi}} \int_{\nu}^{\infty} dx e^{-x^2/2} \operatorname{erfc} \left[\frac{\nu - \alpha x}{\sqrt{2(1 - \alpha^2)}} \right], \end{aligned}$$

where $\alpha \equiv \xi(r)/\xi(0) < 1$. Therefore, $\xi_{>\nu}(r)$ is

$$1 + \xi_{>\nu}(r) = \frac{P_2}{P_1^2} = \sqrt{\frac{2}{\pi}} \left[\operatorname{erfc} \left(\frac{\nu}{\sqrt{2}} \right) \right]^{-2} \int_{\nu}^{\infty} dx e^{-x^2/2} \operatorname{erfc} \left[\frac{\nu - \alpha x}{\sqrt{2(1 - \alpha^2)}} \right]. \quad (\text{I.37})$$

Let's think about the case when $\alpha = \xi(r)/\xi(0) \ll 1$. This assumption is valid for large scale, because correlation function is very small for large scale. In that limit, we can approximate erfc of the right hand side of equation (I.37) as

$$\operatorname{erfc} \left[\frac{\nu - \alpha x}{\sqrt{2(1 - \alpha^2)}} \right] = \operatorname{erfc} \left(\frac{\nu}{\sqrt{2}} \right) + \sqrt{\frac{2}{\pi}} e^{-\nu^2/2} x\alpha + \mathcal{O}(\alpha^2).$$

Therefore, from equation (I.37), we find

$$\xi_{>\nu}(r) = \frac{2}{\pi} e^{-\nu^2/2} \left[\operatorname{erfc} \left(\frac{\nu}{\sqrt{2}} \right) \right]^{-2} \left[\int_{\nu}^{\infty} dx x e^{-x^2/2} \right] \frac{\xi(r)}{\xi(0)} \quad (\text{I.38})$$

$$\simeq \left(e^{\nu^2/2} \int_{\nu}^{\infty} dy e^{-y^2/2} \right)^{-2} \frac{\xi(r)}{\xi(0)}. \quad (\text{I.39})$$

We can further simplify the equation by assuming $\nu \gg 1$ by using

$$\lim_{x \rightarrow \infty} \sqrt{\frac{\pi}{2}} e^{x^2/2} \operatorname{erfc}\left(\frac{x}{\sqrt{2}}\right) = \frac{1}{x}.$$

That is, in the limit of $\xi(r) \ll \xi(0)$ and $\nu \gg 1$, the relation between the correlation function of regions above the density threshold and the correlation function of underlying density field becomes¹

$$\xi_{>\nu}(r) \simeq \frac{\nu^2}{\sigma^2} \xi(r). \quad (\text{I.42})$$

Note that *correlation of over-dense region* $\xi_{>\nu}(r)$ *become stronger and stronger when we increase threshold* ν . It is because $(\operatorname{erfc}[\nu/\sqrt{2}])^2$ (\propto probability of having two disjoint region have more than threshold density) decays slightly more rapidly than $\exp[-x^2]$ (\propto conditional probability of finding an object at a distance r) as ν becomes large.

Finally, We write the time dependence explicitly:

$$\xi_{>\nu}(r, z) = \left[\frac{\nu(z)}{\sigma(z)} \right]^2 \xi(r, z) = \left[\frac{\delta_c(\Omega_m)}{\sigma^2(z)} \right]^2 \xi(r, z) = \left[\frac{\nu^2(z)}{\delta_c(\Omega_m)} \right]^2 \xi(r, z) \quad (\text{I.43})$$

In the last equality, we use $\nu(z) = \delta_c/\sigma(z)$.

¹While this result is consistent with the Peak-background split method in its leading order, the next-to-leading order correction term is different. We can show that from the general approximation of

$$\int_{x_0}^{\infty} dx f(x) e^{-x^2/2} = e^{-x_0^2/2} \left\{ \frac{f(x_0)}{x_0} \left(1 - \frac{1}{x^2} + \dots \right) + \frac{f'(x_0)}{x_0^2} + \dots \right\}.$$

Using the formula, $\operatorname{erfc}(x/\sqrt{2})$ and $\int_{\nu}^{\infty} dx x \exp[-x^2/2]$ can be approximated as

$$\operatorname{erfc}\left(\frac{x}{\sqrt{2}}\right) \equiv \sqrt{\frac{2}{\pi}} \int_{\nu}^{\infty} dx e^{-x^2/2} = \sqrt{\frac{2}{\pi}} e^{-\nu^2/2} \left\{ \frac{1}{\nu} \left(1 - \frac{1}{\nu^2} + \dots \right) \right\}$$

and

$$\int_{\nu}^{\infty} dx x e^{-x^2/2} = \sqrt{\frac{2}{\pi}} e^{-\nu^2/2} \left\{ \left(1 - \frac{1}{\nu^2} + \dots \right) + \frac{1}{\nu^2} + \dots \right\}.$$

Therefore, from equation (I.39), we find

$$\xi_{>\nu}(r) = \frac{2}{\pi} e^{-\nu^2/2} \left[\operatorname{erfc}\left(\frac{\nu}{\sqrt{2}}\right) \right]^{-2} \left[\int_{\nu}^{\infty} dx x e^{-x^2/2} \right] \frac{\xi(r)}{\xi(0)} = \left[\frac{1}{\nu} \left(1 - \frac{1}{\nu^2} + \dots \right) \right]^{-2} \frac{\xi(r)}{\sigma^2}.$$

Therefore, the next-to-leading order linear bias parameter is

$$b(\nu) \simeq \nu \left(1 + \frac{1}{\nu^2} \right) \frac{1}{\sigma} = \frac{1}{\sigma} \left(\nu + \frac{1}{\nu} \right) \quad (\text{I.40})$$

This result has to be compared to equation (I.47) from the Peak-background split method, where the linear bias is given by

$$b(\nu) = \frac{1}{\sigma} \left(\nu - \frac{1}{\nu} \right). \quad (\text{I.41})$$

I.2.1.2 General pdf : peak/background split method

As an example, let's assume that we can divide the density field by two parts

$$\delta(\mathbf{x}) = \delta_S(\mathbf{x}) + \delta_N(\mathbf{x}),$$

where $\delta_N(\mathbf{x})$ (“peak”) has a much smaller correlation length l_C and $\delta_S(\mathbf{x})$ (“background”, whose correlation length is $l_S \gg l_C$) has relatively smaller amplitude than $\delta_N(\mathbf{x})$. That is, the probability distribution of $\delta(\mathbf{x})$ is effectively the same as that of $\delta_N(\mathbf{x})$ whereas the correlation function of $\delta(\mathbf{x})$ is just equal to that of $\delta_S(\mathbf{x})$ for the length scale much larger than l_C .

Consider a region of size l ($l_C \ll l \ll l_S$). Since its length scale is smaller than l_S , $\delta_S(\mathbf{x})$ is almost stationary in this region, and $\delta_N(\mathbf{x})$ fluctuates many times. Therefore, placing the object when $\delta = \delta_S + \delta_N > \nu\sigma$, is equivalent to place the object when $\delta_N > \nu\sigma - \delta_s(\mathbf{x}) \equiv \nu_{eff}\sigma$. The probability of finding an object is then spatially modulated as:

$$\begin{aligned} P(> \nu, \mathbf{x}) &= P_N(> \nu - \delta_s(\mathbf{x})/\sigma) \\ &= P(> \nu) \left[1 - \frac{d \ln P(> \nu)}{d\nu} \frac{\delta_s(\mathbf{x})}{\sigma} \right]. \end{aligned} \quad (\text{I.44})$$

Here, $P(> \nu)$ denotes the *homogeneous*² probability of exceeding the threshold. This equation says that the probability of finding object is higher where δ_s is higher, and lower when δ_s is lower. How much it becomes higher? It's not quite δ_s , but a multiplet of it. Thus the objects can be considered to be a random sample from a field with density contrast:

$$\delta_c(\mathbf{x}) = -\frac{d \ln P(> \nu)}{d\nu} \frac{\delta_s(\mathbf{x})}{\sigma} = b\delta_s(\mathbf{x}) \quad (\text{I.45})$$

where

$$b \equiv -\frac{1}{\sigma} \frac{d \ln P(> \nu)}{d\nu}$$

is the linear bias parameter. One can find the more elaborated version of peak/background split method for objects at density peaks of 3D Gaussian density field in Bardeen et al. (1986).

I.2.2 Linear bias with Press-Schechter mass function

The linear bias calculated in Kaiser (1984) becomes more explicit when aided by Press-Schechter formalism (Cole & Kaiser, 1989). The background number density of halo

²We assume that the statistical properties of δ_N is homogeneous on this scale as $l > l_C$.

of mass m is simply a PS mass function,

$$\bar{n}(m) = -2 \frac{\bar{\rho} \nu}{m^2} \frac{e^{-\nu^2/2}}{\sqrt{2\pi}} \frac{d \ln \sigma_R}{d \ln m}.$$

With background/peak split argument, the number density of halo in the denser region with linear density contrast δ_s is a PS mass function, with $\nu_{\text{eff}} = \nu - \delta_s(\mathbf{x})/\sigma_R$:

$$n(m, \mathbf{x}) \simeq \bar{n}(m) - \frac{\partial \bar{n}}{\partial \nu} \frac{\delta_s(\mathbf{x})}{\sigma_R}$$

Therefore, the halo density contrast is

$$\delta_h(\mathbf{x}) \equiv \frac{n(m, \mathbf{x}) - \bar{n}(m)}{\bar{n}(m)} = - \frac{\partial \ln \bar{n}}{\partial \nu} \frac{\delta_s(\mathbf{x})}{\sigma_R}, \quad (\text{I.46})$$

and linear bias parameter $\delta_h(\mathbf{x}, z) = b(m, z) \delta_s(\mathbf{x}, z)$ is

$$b(m, z) \equiv \frac{\delta_h(\mathbf{x})}{\delta_s(\mathbf{x})} = - \frac{1}{\sigma_R} \frac{\partial \ln \bar{n}}{\partial \nu} = - \frac{1}{\sigma_R} \left(\frac{1}{\nu} - \nu \right) = \frac{\nu^2(z) - 1}{\nu(z) \sigma_R(z)} = \frac{\nu^2(z) - 1}{\delta_c(\Omega_m)}. \quad (\text{I.47})$$

If we use $\nu(z) = \delta_c(\Omega_m)/\sigma_R(z)$ in last equality. Note that it is consistent with Mo & White (1996):

$$\delta_h(\mathbf{x}, z) = \frac{\nu^2(z) - 1}{\delta_c(\Omega_m)} \delta_s(\mathbf{x}, z) = \frac{\nu^2(z) - 1}{\delta_c(\Omega_m)} D(z) \delta_s(\mathbf{x}, z=0) = \frac{\nu^2(z) - 1}{\delta_c(z)} \delta_s(\mathbf{x}, z=0).$$

I.2.3 Linear bias parameters from excursion set approach

Mo & White (1996) calculated the bias parameters by using Press-Schechter formalism. Let us start from the conditional mass function, which can be also interpreted as “the fraction of the mass in a region of initial radius R_0 and linear overdensity δ_0 ³ which at redshift z_1 is contained in halos of mass m_1 ” ($m_1 < m_0 \equiv 4\pi\bar{\rho}(1+\delta)R_0^3/3$):

$$f(\sigma_1, \delta_1 | \sigma_0, \delta_0) \frac{d\sigma_1^2}{dm_1} dm_1 = \frac{1}{(2\pi)^{1/2}} \frac{\delta_1 - \delta_0}{(\sigma_1^2 - \sigma_0^2)^{3/2}} \exp \left[-\frac{(\delta_1 - \delta_0)^2}{2(\sigma_1^2 - \sigma_0^2)} \right] \frac{d\sigma_1^2}{dm_1} dm_1$$

Using this, the average number of m_1 halos identified at redshift z_1 in a spherical region with comoving radius R_0 and overdensity δ_0 is

$$\mathcal{N}(1|0) dm_1 \equiv \frac{m_0}{m_1} f(1|0) \frac{d\sigma_1^2}{dm_1} dm_1, \quad (\text{I.48})$$

³Remember that this density contrast has the value extrapolated to $z = 0$.

where $(1|0)$ is the shot-hand notation for $(\sigma_1, \delta_1 | \sigma_0, \delta_0)$. Note that $\delta_1 > \delta_0$ because m_1 is a collapsed halo at z_1 and m_0 is un-collapsed even at $z = 0$. Using this definition, the average over-abundance of halos in the spherical region becomes

$$\delta_h^L(1|0) = \frac{\mathcal{N}(1|0)}{n(m_1, z_1)V_0} - 1, \quad (\text{I.49})$$

where $V_0 = 4\pi R_0^3/3$. Especially, when $m_0 \gg m_1$, (or $\sigma_0 \ll \sigma_1$) and $\delta_0 \ll \delta_1$, it becomes

$$\delta_h^L(1|0) = \frac{\nu_1^2 - 1}{\delta_1} \delta_0 = \frac{\nu_1^2 - 1}{\delta_c} \delta_0(z_1), \quad (\text{I.50})$$

which coincide with the result of previous section. That is, “the halo overdensity in these *Lagrangian* sphere is directly proportional to the linear mass overdensity.” So, the proportionality constant in equation (I.50) is called a “Lagrangian bias”, in the sense that we calculate the halo overdensity in the initial sphere.

How are halo over-densities related to the current over-dense region? In order to answer the question, we have to relate the initial density contrast, δ_0 , and radius, R_0 , of the spherical region to those in the recent values, δ and R , respectively. First, from the mass conservation, we get

$$R_0^3 = R^3(1 + \delta).$$

The initial spherical region evolves as described by ‘spherical collapse mode’, whose radius evolves by (for positive δ_0)

$$\frac{R(R_0, \delta_0, z)}{R_0} = \frac{3}{10} \frac{1 - \cos \theta}{\delta_0},$$

and

$$\frac{1}{1+z} = \frac{3 \times 6^{2/3}}{20} \frac{(\theta - \sin \theta)^{1/2}}{\delta_0}.$$

By equating those two equation, we can express $\delta \equiv (R_0/R)^3 - 1$ as a function of δ_0 :

$$\delta_0 \simeq -1.35(1 + \delta)^{-2/3} + 0.78785(1 + \delta)^{-0.58661} - 1.12431(1 + \delta)^{-1/2} + 1.68647$$

Finally, combining all of these, we find the average overdensity of halos in sphere with *current* radius R and current mass density contrast δ as

$$\delta_h(1|0) = \frac{\mathcal{N}(1|0)}{n(m_1, z_1)V} - 1, \quad (\text{I.51})$$

where $V = 4\pi R^3/3 = V_0/(1 + \delta)$. Again, for $m_0 \gg m_1$ and $\delta_0 \ll \delta_1$, we have

$$\delta_h(1|0) = \left(1 + \frac{\nu_1^2 - 1}{\delta_1}\right) \delta. \quad (\text{I.52})$$

We can also derive the higher order expression adding more terms from the spherical collapse model as Mo et al. (1997). They use the following expansion :

$$\delta_0(z) = \sum_{n=0}^{\infty} a_n \delta^n, \quad (\text{I.53})$$

where first five coefficients calculated using spherical collapse model are

$$a_0 = 0, a_1 = 1, a_2 = -\frac{17}{21}, a_3 = \frac{341}{567}, a_4 = -\frac{55805}{130977}. \quad (\text{I.54})$$

I.2.4 Linear bias with Shech-Tormen mass function

In the linear region, relation between Eulerian and Lagrangian bias are easy to show by using the relation $R_0^3 = R^3(1 + \delta)$. It is first shown in this paper :

$$\begin{aligned} \delta_h^E(1|0) &= \frac{\mathcal{N}(1|0)}{n(m_1, z_1)V} - 1 \\ &\simeq (1 + \delta) [1 + \delta_h^L(1|0)] - 1 \simeq [1 + b_{Lag}(m_1, z_1)] \delta \end{aligned} \quad (\text{I.55})$$

Motivated by this equation, Sheth & Tormen (1999) define the Eulerian bias as

$$b_{\text{Eul}}(m_1, z_1) \equiv 1 + b_{\text{Lag}}(m_1, z_1). \quad (\text{I.56})$$

Physically, this additional factor 1 is added because, the Eulerian overdensity region of density $\bar{\rho}(1 + \delta)$ spans larger volume in the initial Lagrangian space. Therefore, more masses are enclosed within that sphere, which leads more number of halos by factor of $1 + \delta$.

In this paper, they apply the “peak background split method” to newly found mass function (so called ‘Sheth-Tormen’ mass function) which is, with our notation,

$$n_{ST}(m) = -2 \frac{\bar{\rho}}{m^2} f(\nu) \frac{d \ln \sigma_R}{d \ln m}, \quad (\text{I.57})$$

where

$$f_{ST}(\nu) = A \sqrt{\frac{\alpha \nu^2}{2\pi}} \left[1 + \frac{1}{(\alpha \nu^2)^p} \right] \exp \left(-\frac{\alpha \nu^2}{2} \right). \quad (\text{I.58})$$

From peak background split method in Cole & Kaiser (1989), we calculate Lagrangian bias parameter as

$$\begin{aligned} b_{ST}^L(m, z) &= -\frac{1}{\sigma_R} \frac{\partial \ln \bar{n}_{ST}}{\partial \nu} = -\frac{1}{\sigma_R} \frac{\partial}{\partial \nu} \left[\ln \nu + \ln \left(1 + \frac{1}{(\alpha \nu^2)^p} \right) - \frac{\alpha \nu^2}{2} \right] \\ &= \frac{1}{\delta_c(\Omega_m)} \left[\alpha \nu^2(z) - 1 + \frac{2p}{1 + (\alpha \nu^2(z))^p} \right]. \end{aligned} \quad (\text{I.59})$$

Therefore, the Eulerian bias is

$$b_{ST} = 1 + \epsilon_1 + E_1 = 1 + \frac{\alpha \nu^2(z) - 1}{\delta_c(\Omega_m)} + \frac{2p/\delta_c(\Omega_m)}{1 + (\alpha \nu^2(z))^p}.$$

I.2.5 Nonlinear halo bias and its mass dependence

Halo bias for the more general mass function is given in Scoccimarro et al. (2001b).

$$\delta_h(m_1, z_1 | m_0, z_0) = \sum_k b_k(m, z_1) \delta^k, \quad (\text{I.60})$$

will lead the bias parameters as following:

$$b_1(m) = 1 + \epsilon_1 + E_1 \quad (\text{I.61})$$

$$b_2(m) = 2(1 + a_2)(\epsilon_1 + E_1) + \epsilon_2 + E_2 \quad (\text{I.62})$$

$$b_3(m) = 6(a_2 + a_3)(\epsilon_1 + E_1) + 3(1 + 2a_2)(\epsilon_2 + E_2) + \epsilon_3 + E_3 \quad (\text{I.63})$$

$$b_4(m) = 24(a_3 + a_4)(\epsilon_1 + E_1) + 12[a_2^2 + 2(a_2 + a_3)](\epsilon_2 + E_2) \quad (\text{I.64})$$

$$+ 4(1 + 3a_2)(\epsilon_3 + E_3) + \epsilon_4 + E_4 \quad (\text{I.65})$$

where, the coefficients are

$$\epsilon_1 = \frac{\alpha\nu^2 - 1}{\delta_f}$$

$$\epsilon_2 = \frac{\alpha\nu^2}{\delta_f^2}(\alpha\nu^2 - 3)$$

$$\epsilon_3 = \frac{\alpha\nu^2}{\delta_f^3}(\alpha^2\nu^4 - 6\alpha\nu^2 + 3)$$

$$\epsilon_4 = \left(\frac{\alpha\nu^2}{\delta_f^2} \right)^2 (\alpha^2\nu^4 - 10\alpha\nu^2 + 15)$$

$$E_1 = \frac{2p/\delta_f}{1 + (\alpha\nu^2)^p}$$

$$\frac{E_2}{E_1} = \left(\frac{1 + 2p}{\delta_f} + 2\epsilon_1 \right)$$

$$\frac{E_3}{E_1} = \left[\frac{4(p^2 - 1) + 6p\alpha\nu^2}{\delta_f^2} + 3\epsilon_1^2 \right]$$

$$\frac{E_4}{E_1} = \frac{2\alpha\nu^2}{\delta_f^2} \left(2\frac{\alpha^2\nu^4}{\delta_f} - 15\epsilon_1 \right) + 2\frac{(1+p)}{\delta_f^2} \left[\frac{4(p^2 - 1) + 8(p-1)\alpha\nu^2 + 3}{\delta_f} + 6\alpha\nu^2\epsilon_1 \right].$$

Here, $\nu = \delta_c/\sigma_m(z_1)$ with $\delta_c \simeq 1.686$, and $\delta_f \equiv \delta_c D(z_0)/D(z_1)$ which is equal to δ_c with an assumption that the formation redshift (z_1) is the same as the observing redshift (z_0).

For the case of PS formula, we have all the E_n 's to be zero, and $\alpha = 1$. And the formulae reduce to those in Mo et al. (1997). Note that, by construction, bias parameters

have to satisfy the consistency relation:

$$\int dm \frac{nm(m, z)}{\bar{\rho}} b_k(m, z) = \delta_{k1}. \quad (\text{I.66})$$

I.3 Halo bias with local type primordial non-Gaussianity

Dalal et al. (2008) has derived the bias parameter in the presence of the *local type* non-Gaussianity by using a method similar to Kaiser (1984). Local type non-Gaussianity is defined by constructing a non-Gaussian Bardeen's potential outside of horizon as

$$\Phi(\mathbf{x}) = \phi(\mathbf{x}) + f_{NL}(\phi^2(\mathbf{x}) + \langle \phi^2 \rangle), \quad (\text{I.67})$$

where ϕ is a Gaussian random field whose power spectrum is similar to Φ . Then, the Laplacian of Φ becomes

$$\nabla^2 \Phi = \nabla^2 \phi + 2f_{NL} [\phi \nabla^2 \phi + |\nabla \phi|^2].$$

We set $\nabla \phi = 0$, as we are interested in the density peak region where ϕ is also maximum:

$$\nabla^2 \Phi = \nabla^2 \phi + 2f_{NL} \phi \nabla^2 \phi \quad (\text{I.68})$$

We can relate the Laplacian of ϕ with the density field by Poisson equation. The *exact* relation inside the horizon is

$$\delta_{NG}(k, z) = M(k, z) \Phi(k) = \frac{3}{2} \frac{D(z)}{H_0^2 \Omega_{m0}} k^2 T(k) \Phi(k). \quad (\text{I.69})$$

I.3.1 Basic idea

To make analysis simple, let's focus on the large scale where $T(k) \simeq 1$. Then, we found the simple relation between the non-Gaussian density field, δ_{NG} and Gaussian 'density' field $\delta \propto -\nabla^2 \phi$ as

$$\delta_{NG} \simeq \delta(1 + 2f_{NL} \phi_p) = \delta \left[1 + 2f_{NL} \frac{\phi}{g(z)} \right]. \quad (\text{I.70})$$

Note that we explicitly write down the time dependence in the last equation using time evolution function of gravitational potential $\phi(z) = g(z) \phi_p \propto D(z)(1+z)\phi_p$.

What does equation (I.70) mean? It says that the density field is modified by $\Delta \delta = 2\delta f_{NL} \phi_p$. Since $\phi_p \delta > 0$, that is the Bardeen potential (or spatial curvature) is

positive (negative) in the over-dense (under-dense) region, the non-Gaussian correction to the density field is positive (negative) for positive (negative) f_{NL} . Therefore, we generally expect that the number of the regions whose overdensity exceed δ_c , thus halos, increases (decrease) in the presence of positive (negative) f_{NL} . As a result, the linear bias parameter increases (decreases) for positive (negative) f_{NL} .

Let's calculate the correction to the bias parameter due to non-Gaussianity. We consider the long wavelength mode with density perturbation δ and corresponding curvature perturbation of ϕ . As Kaiser (1984) showed, if density field is Gaussian, the presence of this 'background' density field boosted the 'peak' over-densities, so that the number of halos increases, and becomes $\delta_h = b_L \delta$. In addition to that, due to the non-Gaussianity, 'peak' height δ_{pk} is enhanced by the long-wavelength curvature perturbation by $2f_{NL}\phi_p\delta_{pk}$. If we focus on the peaks near threshold, $\delta_{pk} \sim \delta_c$, the amount of enhancement becomes $2f_{NL}\phi_p\delta_c$, then halo density is now

$$\delta_h = b_L(\delta + 2f_{NL}\phi_p\delta_c) = b_L(1 + 2f_{NL}\delta_c\frac{\phi_p}{\delta})\delta. \quad (\text{I.71})$$

Therefore, correction to the bias is

$$\Delta b_L(m, z) = \frac{2b_L f_{NL} \delta_c}{M(k, z)} = 2b_L(m, z) f_{NL} \delta_c(\Omega_m) \frac{2H_0^2 \Omega_{m0}}{3D(z)k^2} \quad (\text{I.72})$$

$$= 2[b(m, z) - 1] f_{NL} \delta_c(\Omega_m) \frac{2H_0^2 \Omega_{m0}}{3D(z)k^2}. \quad (\text{I.73})$$

For last line, I use $b = b_L + 1$. Note that this argument is valid only for the small k where $T(k) \simeq 1$. Matarrese & Verde (2008) provide more general derivation based on MLB formula (Appendix K).

I.3.2 Calculation

In this section, we will calculate the Lagrangian bias of halos in the non-Gaussian density field. We utilize that both δ and ϕ are Gaussian field, and $f_{NL}|\phi|$ is very small. In order to use Gaussianity of δ and ϕ , we invert the equation $\delta_{NG} \simeq \delta(1 + 2f_{NL}\phi)$ as

$$\delta \simeq (1 - 2f_{NL}\phi_p)\delta_{NG}.$$

Then, the condition for forming halos ($\delta_{NG} > \delta_c$) becomes

$$\delta > (1 - 2f_{NL}\phi_p)\delta_c.$$

Calculating one point probability P_1

We first calculate the one-point probability that a spatial region has a overdensity more than δ_c , i.e. that region contains a halo:

$$P_1 = \int_{-\infty}^{\infty} d\phi \int_{\delta_c - 2f_{NL}\delta_c\phi_p}^{\infty} d\delta \frac{1}{\sqrt{(2\pi)^2|\Sigma|}} \exp \left[-\frac{1}{2}(\phi \ \delta) \cdot \Sigma^{-1} \cdot (\phi \ \delta)^T \right] \quad (\text{I.74})$$

Here, Σ is a covariant matrix of $(\delta \ \phi)$, which is

$$\Sigma = \begin{pmatrix} \sigma_\phi^2 & \langle \phi \delta \rangle \\ \langle \phi \delta \rangle & \sigma_\delta^2 \end{pmatrix}. \quad (\text{I.75})$$

After redefining variables: $\mu = \phi/\sigma_\phi$, $\nu = \delta/\sigma_\delta$, $\nu_c = \delta_c/\sigma_\delta$ and $\eta = 2f_{NL}\sigma_\phi\nu_c$, and covariant matrix

$$S = \begin{pmatrix} 1 & r \\ r & 1 \end{pmatrix} \quad (\text{I.76})$$

with $r = \langle \mu \nu \rangle$, integration becomes

$$P_1 = \int_{-\infty}^{\infty} d\mu \int_{\nu_c - \mu\eta}^{\infty} d\nu \frac{1}{2\pi\sqrt{1-r^2}} \exp \left[-\frac{\mu^2 + \nu^2 - 2r\mu\nu}{2(1-r^2)} \right]. \quad (\text{I.77})$$

Change of variable $(\mu, \nu) \rightarrow (\mu, v \equiv \nu + \eta\mu)$ does not changes integration measures, but disentangle the integrations as

$$\begin{aligned} P_1 &= \int_{-\infty}^{\infty} d\mu \int_{\nu_c}^{\infty} dv \frac{1}{2\pi\sqrt{1-r^2}} \exp \left[-\frac{\sigma_v^2\mu^2 - 2v(r+\eta)\mu + v^2}{2(1-r^2)} \right] \\ &= \int_{\nu_c}^{\infty} dv \frac{1}{2\pi\sqrt{1-r^2}} \sqrt{\frac{2\pi(1-r^2)}{\sigma_v^2}} \exp \left(-\frac{[\sigma_v^2 - (r+\eta)^2]v^2}{2(1-r^2)\sigma_v^2} \right) \\ &= \int_{\nu_c}^{\infty} dv \frac{1}{\sqrt{2\pi\sigma_v^2}} \exp \left[-\frac{v^2}{2\sigma_v^2} \right] = \frac{1}{2} \text{erfc} \left[\frac{x_c}{\sqrt{2}} \right], \end{aligned} \quad (\text{I.78})$$

where we use $\sigma_v^2 \equiv \langle v^2 \rangle = \langle (\nu + \eta\mu)^2 \rangle = 1 + 2\eta r + \eta^2$ and

$$x_c \equiv \frac{\nu_c}{\sqrt{1 + 2\eta r + \eta^2}}.$$

Therefore, the effect of non-Gaussianity on the abundance of the peak is simply to *rescale* the threshold density ν_c by an inverse of f_{NL} and cross correlation of $\langle \delta \phi \rangle$. Note that the threshold decreases, that is, we have more halo for given density field, as f_{NL} increases as we expected from the earlier discussion.

Calculating two point probability P_2

We calculate two point probability by integrating four Gaussian parameters $\mathbf{u} = (\mu_1, \mu_2, \nu_1, \nu_2)$:

$$P_2 = \int d^4 u \frac{\exp \left[-\frac{1}{2} \mathbf{u} \cdot \Sigma^{-1} \cdot \mathbf{u} \right]}{(2\pi)^2 |\Sigma|^{1/2}} \Theta(\nu_1 + \eta\mu_1 - \nu_c) \Theta(\nu_2 + \eta\mu_2 - \nu_c), \quad (\text{I.79})$$

where we denote the covariant matrix Σ as

$$\Sigma = \begin{pmatrix} 1 & \gamma & r & \beta \\ \gamma & 1 & \beta & r \\ r & \beta & 1 & \psi \\ \beta & r & \psi & 1 \end{pmatrix}, \quad (\text{I.80})$$

by using $\psi = \langle \nu_1 \nu_2 \rangle$, $\gamma = \langle \mu_1 \mu_2 \rangle$, $r = \langle \mu_1 \nu_1 \rangle = \langle \mu_2 \nu_2 \rangle$, $\beta = \langle \mu_1 \nu_2 \rangle = \langle \mu_2 \nu_1 \rangle$.

Again, after change the variable as, $v_i \equiv \nu_i - \eta\mu_i$, we can integrate out the two potential variables μ_1 and μ_2 :

$$P_2 = \int_{\nu_c}^{\infty} dv_1 \int_{\nu_c}^{\infty} dv_2 \frac{1}{2\pi |S|^{1/2}} \exp \left[-\frac{1}{2} \mathbf{v} \cdot S^{-1} \cdot \mathbf{v} \right], \quad (\text{I.81})$$

where $\mathbf{v} = (v_1, v_2)$ and there covariant matrix is

$$S = \begin{pmatrix} \sigma_v^2 & \psi + 2\eta\beta + \gamma\eta^2 \\ \psi + 2\eta\beta + \gamma\eta^2 & \sigma_v^2 \end{pmatrix}.$$

Let's define a matrix $C \equiv S/\sigma_v^2$, and two variables $x_i = v_i/\sigma_v$ and $\chi = (\psi + 2\eta\beta + \gamma\eta^2)/\sigma_v$ in order to simplify the calculation. Then, $\mathbf{v} \cdot S^{-1} \cdot \mathbf{v} = \mathbf{x} \cdot C^{-1} \cdot \mathbf{x}$ and $dv_1 dv_2 / |S|^{1/2} = dx_1 dx_2 / |C|^{1/2}$ makes the expression

$$P_2 = \int_{x_c}^{\infty} dx_1 \int_{x_c}^{\infty} dx_2 \frac{1}{2\pi |C|^{1/2}} \exp \left[-\frac{1}{2} \mathbf{x} \cdot C^{-1} \cdot \mathbf{x} \right], \quad (\text{I.82})$$

where x_c is defined when we calculate P_1 , and the covariant matrix is simply

$$C = \begin{pmatrix} 1 & \chi \\ \chi & 1 \end{pmatrix}.$$

Integration at this point has already done when we follow Kaiser (1984) in the previous section, and we can write the answer immediately.

$$P_2 = \sqrt{\frac{1}{8\pi}} \int_{x_c}^{\infty} dx e^{-x^2/2} \text{erfc} \left[\frac{x_c - \chi x}{\sqrt{2(1 - \chi^2)}} \right] \quad (\text{I.83})$$

Therefore, when $\nu_c \gg 1$ and $\chi \ll 1$, we get

$$\xi_{pk} = \frac{P_2}{P_1^2} - 1 \simeq \left(e^{x_c^2/2} \int_{x_c}^{\infty} dy e^{-y^2/2} \right)^{-2} \chi \simeq x_c^2 \chi, \quad (\text{I.84})$$

which to lowest order in η becomes

$$\begin{aligned}\xi_{pk} &\simeq \nu_c^2(\psi + 2\eta\beta + \gamma\eta^2)(1 + 2r\eta + \eta^2)^{-2} \\ &\simeq \nu_c^2(\psi + 2\beta\eta)(1 - 4r\eta) \\ &= \nu_c^2(\psi + 2(\beta - 2r\psi)\eta).\end{aligned}$$

Now we further simplify the equation by neglecting $r\psi \equiv \sigma_{\delta\phi}\xi_{\delta\phi}(r_{12})/(\sigma_\delta^3\sigma_\phi) < \xi_{\delta\delta}(r_{12})/\sigma_\delta^2$ compare to $\beta \equiv \xi_{\delta\phi}(r_{12})/(\sigma_\delta\sigma_\phi)$, because $\delta/\phi \sim k^2 \sim 1/r^2$ leads

$$\frac{\xi_{\delta\phi}}{\sigma_{\delta\phi}} \sim \frac{\xi_{\delta\delta}r_{12}^2}{\sigma_\delta^2 R^2} = \left(\frac{r_{12}}{R}\right)^2 \frac{\xi_{\delta\delta}}{\sigma_\delta^2} \gg \frac{\xi_{\delta\delta}}{\sigma_\delta^2}.$$

Then,

$$\xi_{pk}(r) \simeq \nu_c^2(\psi + 2\beta\eta) = b_L^2 \left[\xi_{\delta\delta}(r) + 4f_{NL}\delta_c \xi_{\delta\phi}(r) \right].$$

Finally, Fourier transformation of it gives the peak (halo) power spectrum

$$P_h(k) = b_L^2 \left[P_{\delta\delta}(k) + 4f_{NL}\delta_c P_{\delta\phi}(k) \right], \quad (\text{I.85})$$

and scale dependence bias as

$$\Delta b(k) = 2b_L f_{NL} \delta_c \frac{P_{\delta\phi}(k)}{P_{\delta\delta}(k)} = \frac{2b_L f_{NL} \delta_c}{M(k, z)} = \frac{2(b-1)f_{NL}\delta_c}{M(k, z)}. \quad (\text{I.86})$$

In the last equation, we write Δb in terms of the Eulerian bias b . Again, because it stems from the approximated relation between δ_{NG} and δ , this result only works for the very large scale when we make approximation of $T(k) \simeq 1$.

Appendix J

Integration of T_R^{1112}

In the standard perturbation theory, the four-point correlator contained in the definition of T_R^{1112} (see Eq. (5.28)) is given by¹

$$\begin{aligned} & \langle \delta^{(1)}(\mathbf{k}_1) \delta^{(1)}(\mathbf{k}_2) \delta^{(1)}(\mathbf{k}_3) \delta^{(2)}(\mathbf{k}_4) \rangle \\ &= \int \frac{d^3 q}{(2\pi)^3} F_2^{(s)}(\mathbf{q}, \mathbf{k}_4 - \mathbf{q}) \langle \delta^{(1)}(\mathbf{k}_1) \delta^{(1)}(\mathbf{k}_2) \delta^{(1)}(\mathbf{k}_3) \delta^{(1)}(\mathbf{k}_4 - \mathbf{q}) \delta^{(1)}(\mathbf{q}) \rangle. \end{aligned} \quad (\text{J.1})$$

For non-Gaussian density fields, the leading order of Eq. (J.1) contains the ensemble average of products of six Gaussian variables, ϕ , which gives products of three power spectra, P_ϕ . We find

$$\begin{aligned} & \int \frac{d^3 q}{(2\pi)^3} F_2^{(s)}(\mathbf{q}, \mathbf{k}_4 - \mathbf{q}) \langle \delta^{(1)}(\mathbf{k}_1) \delta^{(1)}(\mathbf{k}_2) \delta^{(1)}(\mathbf{k}_3) \delta^{(1)}(\mathbf{k}_4 - \mathbf{q}) \delta^{(1)}(\mathbf{q}) \rangle \\ &= \int \frac{d^3 q}{(2\pi)^3} F_2^{(s)}(\mathbf{q}, \mathbf{k}_4 - \mathbf{q}) \langle \delta^{(1)}(\mathbf{k}_1) \delta^{(1)}(\mathbf{k}_2) \rangle \langle \delta^{(1)}(\mathbf{k}_3) \delta^{(1)}(\mathbf{k}_4 - \mathbf{q}) \delta^{(1)}(\mathbf{q}) \rangle \\ & \quad + 2 \int \frac{d^3 q}{(2\pi)^3} F_2^{(s)}(\mathbf{q}, \mathbf{k}_4 - \mathbf{q}) \langle \delta^{(1)}(\mathbf{k}_1) \delta^{(1)}(\mathbf{q}) \rangle \langle \delta^{(1)}(\mathbf{k}_2) \delta^{(1)}(\mathbf{k}_3) \delta^{(1)}(\mathbf{k}_4 - \mathbf{q}) \rangle \\ & \quad + (\text{cyclic } 123) \\ &= (2\pi)^3 \left[2f_{\text{NL}} P_m(k_1) \mathcal{M}(k_3) \int d^3 q \mathcal{M}(\mathbf{q}) \mathcal{M}(|\mathbf{k}_4 - \mathbf{q}|) P_\phi(q) \right. \\ & \quad \times \{P_\phi(|\mathbf{k}_4 - \mathbf{q}|) + 2P_\phi(k_3)\} F_2^{(s)}(\mathbf{q}, \mathbf{k}_4 - \mathbf{q}) \delta^D(\mathbf{k}_{12}) \\ & \quad + 4f_{\text{NL}} \mathcal{M}(k_2) \mathcal{M}(k_3) \mathcal{M}(k_{14}) P_m(k_1) F_2^{(s)}(-\mathbf{k}_1, \mathbf{k}_{14}) \\ & \quad \times \{P_\phi(k_2) P_\phi(k_3) + P_\phi(k_2) P_\phi(k_{14}) + P_\phi(k_3) P_\phi(k_{14})\} \\ & \quad \left. + (\text{cyclic } 123) \right] \delta^D(\mathbf{k}_{1234}). \end{aligned} \quad (\text{J.2})$$

¹Sefusatti (2009) also derived and studied this term independently.

Therefore, T_R^{1112} is given by

$$\begin{aligned}
& T_R^{1112}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) \\
= & \mathcal{W}_R(k_1)\mathcal{W}_R(k_2)\mathcal{W}_R(k_3)\mathcal{W}_R(k_4) \left[2f_{\text{NL}}P_m(k_1)\mathcal{M}(k_3) \right. \\
& \times \int d^3q \mathcal{M}(q)\mathcal{M}(|\mathbf{k}_4 - \mathbf{q}|)P_\phi(q) \{P_\phi(|\mathbf{k}_4 - \mathbf{q}|) + 2P_\phi(k_3)\} \\
& \times F_2^{(s)}(\mathbf{q}, \mathbf{k}_4 - \mathbf{q})\delta^D(\mathbf{k}_{12}) + 4f_{\text{NL}}\mathcal{M}(k_2)\mathcal{M}(k_3)\mathcal{M}(k_{14})P_m(k_1)F_2^{(s)}(-\mathbf{k}_1, \mathbf{k}_{14}) \\
& \left. \times \{P_\phi(k_2)P_\phi(k_3) + P_\phi(k_2)P_\phi(k_{14}) + P_\phi(k_3)P_\phi(k_{14})\} + (\text{cyclic } 123) \right], \quad (\text{J.3})
\end{aligned}$$

where $\mathcal{M}(k) \equiv \mathcal{M}_R(k)/\mathcal{W}_R(k)$, $\mathbf{k}_{ij} = \mathbf{k}_i + \mathbf{k}_j$, and (cyclic 123) denotes that the cyclic changes among $(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)$. We calculate the sum of $\{1112\}$ terms in Eq. (5.29) by integrating

Eq. (J.3):

$$\begin{aligned}
& \int \frac{d^3 q}{(2\pi)^3} T_R^{(2)}(\mathbf{q}, \mathbf{k}_1 - \mathbf{q}, \mathbf{k}_2, \mathbf{k}_3) \\
= & 8f_{\text{NL}} \mathcal{W}_R(k_2) \mathcal{W}_R(k_3) \mathcal{M}(k_1) \mathcal{M}(k_2) \mathcal{M}(k_3) \\
& \times \{P_\phi(k_2) P_\phi(k_3) + P_\phi(k_2) P_\phi(k_1) + P_\phi(k_3) P_\phi(k_1)\} \\
& \times \int \frac{d^3 q}{(2\pi)^3} \mathcal{W}_R(|\mathbf{k}_1 - \mathbf{q}|) \mathcal{W}_R(q) P_m(q) F_2^{(s)}(-\mathbf{q}, \mathbf{k}_1) \\
+ & 4f_{\text{NL}} \mathcal{W}_R(k_2) \mathcal{W}_R(k_3) \mathcal{M}(k_1) \\
& \times \left[P_m(k_2) F_2^{(s)}(\mathbf{k}_2, \mathbf{k}_1) + P_m(k_3) F_2^{(s)}(\mathbf{k}_3, \mathbf{k}_1) \right] \\
& \times \int \frac{d^3 q}{(2\pi)^3} \mathcal{W}_R(|\mathbf{k}_1 - \mathbf{q}|) \mathcal{W}_R(q) \mathcal{M}(q) \mathcal{M}(|\mathbf{k}_1 - \mathbf{q}|) \\
& \times \{P_\phi(q) P_\phi(|\mathbf{k}_1 - \mathbf{q}|) + 2P_\phi(q) P_\phi(k_1)\} \\
+ & 8f_{\text{NL}} \mathcal{W}_R(k_2) \mathcal{W}_R(k_3) \mathcal{M}(k_3) P_m(k_2) \int \frac{d^3 q}{(2\pi)^3} \mathcal{W}_R(|\mathbf{k}_1 - \mathbf{q}|) \mathcal{W}_R(q) \\
& \times \mathcal{M}(|\mathbf{k}_1 - \mathbf{q}|) \mathcal{M}(|\mathbf{k}_2 + \mathbf{q}|) F_2^{(s)}(-\mathbf{k}_2, \mathbf{k}_2 + \mathbf{q}) \\
& \times \{P_\phi(k_3) P_\phi(|\mathbf{k}_1 - \mathbf{q}|) + P_\phi(k_3) P_\phi(|\mathbf{k}_2 + \mathbf{q}|) + P_\phi(|\mathbf{k}_1 - \mathbf{q}|) P_\phi(|\mathbf{k}_2 + \mathbf{q}|)\} \\
& + (k_2 \leftrightarrow k_3) \\
+ & 8f_{\text{NL}} \mathcal{W}_R(k_2) \mathcal{W}_R(k_3) \mathcal{M}(k_3) \int \frac{d^3 q}{(2\pi)^3} \mathcal{W}_R(|\mathbf{k}_1 - \mathbf{q}|) \mathcal{W}_R(q) \\
& \times \mathcal{M}(|\mathbf{k}_1 - \mathbf{q}|) \mathcal{M}(|\mathbf{k}_2 + \mathbf{q}|) P_m(q) F_2^{(s)}(-\mathbf{q}, \mathbf{k}_2 + \mathbf{q}) \\
& \times \{P_\phi(|\mathbf{k}_1 - \mathbf{q}|) P_\phi(k_3) + P_\phi(|\mathbf{k}_1 - \mathbf{q}|) P_\phi(|\mathbf{k}_2 + \mathbf{q}|) + P_\phi(k_3) P_\phi(|\mathbf{k}_2 + \mathbf{q}|)\} \\
& + (k_2 \leftrightarrow k_3) \\
+ & 8f_{\text{NL}} (\mathcal{W}_R(k_2) \mathcal{W}_R(k_3))^2 P_m(k_3) \mathcal{M}(k_2) \int \frac{d^3 q}{(2\pi)^3} \mathcal{M}(q) \mathcal{M}(|\mathbf{k}_2 - \mathbf{q}|) \\
& \times P_\phi(q) \{P_\phi(|\mathbf{k}_2 - \mathbf{q}|) + 2P_\phi(k_2)\} F_2^{(s)}(\mathbf{q}, \mathbf{k}_2 - \mathbf{q}) \\
& + (k_2 \leftrightarrow k_3). \tag{J.4}
\end{aligned}$$

Appendix K

The three-point correlation function (bispectrum) of density peaks

Consider the smoothed density field, $\rho_R(\mathbf{x}) = \bar{\rho}[1 + \delta_R(\mathbf{x})]$, with a general smoothing kernel $W(\mathbf{x})$

$$\delta_R(\mathbf{x}) = \int d^3y W_R(\mathbf{x} - \mathbf{y}) \delta(\mathbf{y}). \quad (\text{K.1})$$

We define the *peaks* as regions in the space where the smoothed density contrast exceed a certain threshold value, δ_c :

$$n_p(\mathbf{x}) = \frac{1}{V_R} \theta[\delta_R(\mathbf{x}) - \delta_c], \quad (\text{K.2})$$

where $\theta(x)$ is a step-function, and V_R is the volume of the smoothed region¹. In this chapter, we shall calculate the three-point correlation function of peaks when the density field follows general probability distribution. We shall use the functional integration method adopted by Matarrese et al. (1986), but, instead of presenting the general formula, we explicitly calculate the two point correlation function and the three point correlation function.

Let us denote the density contrast of the peaks as $\delta_p(\mathbf{x}) \equiv n_p(\mathbf{x})/\langle n_p \rangle - 1$. The probability $P_2(\mathbf{x}_1, \mathbf{x}_2)$ of finding two peaks at two different locations \mathbf{x}_1 and \mathbf{x}_2 , and the probability $P_3(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ of finding three peaks at three different locations \mathbf{x}_1 , \mathbf{x}_2 and \mathbf{x}_3 are related to the two-point correlation function $\xi_p(\mathbf{x}_1, \mathbf{x}_2)$, and the three-point correlation

¹For a spherical top-hat filter

$$W_R(\mathbf{x}) = \frac{1}{V_R} \begin{cases} 1, & |\mathbf{x}| < R \\ 0, & \text{otherwise} \end{cases},$$

the volume of the smoothed region is $V_R = 4\pi R^3/3$. However, V_R may not be well-defined for different filters, e.g. Gaussian filter, where smoothing function is extended to infinity. Fortunately, it is not important to calculate the correlation function from the method we use here, as mean number density cancels out in equation (K.3) and (K.4).

function $\zeta_p(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ of density contrast of peaks as (Peebles, 1980):

$$\begin{aligned} \frac{P_2(\mathbf{x}_1, \mathbf{x}_2)}{P_1^2} &= \frac{\langle n_p(\mathbf{x}_1) n_p(\mathbf{x}_2) \rangle}{\langle n_p \rangle^2} \\ &= \langle (1 + \delta_p(\mathbf{x}_1))(1 + \delta_p(\mathbf{x}_2)) \rangle = 1 + \xi_p(\mathbf{x}_1, \mathbf{x}_2) \end{aligned} \quad (\text{K.3})$$

$$\begin{aligned} \frac{P_3(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)}{P_1^3} &= \frac{\langle n_p(\mathbf{x}_1) n_p(\mathbf{x}_2) n_p(\mathbf{x}_3) \rangle}{\langle n_p \rangle^3} \\ &= \langle (1 + \delta_p(\mathbf{x}_1))(1 + \delta_p(\mathbf{x}_2))(1 + \delta_p(\mathbf{x}_3)) \rangle \\ &= 1 + \xi_p(\mathbf{x}_1, \mathbf{x}_2) + \xi_p(\mathbf{x}_2, \mathbf{x}_3) + \xi_p(\mathbf{x}_3, \mathbf{x}_1) + \zeta_p(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) \end{aligned} \quad (\text{K.4})$$

Therefore, in order to calculate the two-point correlation function and the three-point correlation function, we have to calculate P_N up to $N = 3$.

In Appendix K.1 we calculate the probability P_N : starting from the general functional integration method for calculating P_N , we explicitly calculate P_1 (Appendix K.1.1), P_2 (Appendix K.1.2) and P_3 (Appendix K.1.3). Then, in Appendix K.2, we calculate the two-point correlation function (power spectrum) of peaks by using equation (K.3). Finally, We calculate the three-point correlation function (bispectrum) of peaks in Appendix K.3 by using equation (K.4).

K.1 Probability of finding N distinct peaks

The *probability functional* $\mathcal{P}[\delta(\mathbf{x})]$ is the probability distribution function of $\delta(\mathbf{x})$ at all point \mathbf{x} in space. It is normalized to be

$$\int [\mathcal{D}\delta] \mathcal{P}[\delta(\mathbf{x})] = 1,$$

with a suitable measure $[\mathcal{D}\delta]$. The probability distribution function of the density field δ_0 as a specific position \mathbf{x}_0 can be calculated as

$$P(\delta_0) = \int [\mathcal{D}\delta] \mathcal{P}[\delta(\mathbf{x})] \delta^D(\delta(\mathbf{x}_0) - \delta_0), \quad (\text{K.5})$$

which means that fixing the density at \mathbf{x}_0 to be δ_0 and marginalize over all other points.

As defined in equation (K.2), the *peaks* are the regions where the smoothed density field [Eq. (K.1)] exceeds δ_c . By using a $\mathcal{P}[\delta(\mathbf{x})]$ we can formulate the probability

$P_N(\mathbf{x}_1, \dots, \mathbf{x}_N)$, which is the probability of finding N peaks at $\mathbf{x}_1, \dots, \mathbf{x}_N$, as

$$\begin{aligned}
& P_N(\mathbf{x}_1, \dots, \mathbf{x}_N) \\
&= \int_{\delta_c}^{\infty} d\alpha_1 \cdots \int_{\delta_c}^{\infty} d\alpha_N \int [\mathcal{D}\delta] \mathcal{P}[\delta(\mathbf{x})] \prod_{r=1}^N \delta^D(\delta_R(\mathbf{x}_r) - \alpha_r) \\
&= \int [\mathcal{D}\delta] \mathcal{P}[\delta(\mathbf{x})] \prod_{r=1}^N \int_{-\infty}^{\infty} \frac{d\phi_r}{2\pi} \int_{\nu\sigma_R}^{\infty} d\alpha_r e^{i\phi_r [\int d^3y W_R(\mathbf{x}_r - \mathbf{y}) \delta(\mathbf{y}) - \alpha_r]}. \tag{K.6}
\end{aligned}$$

Here, in the second equality we use $\int \delta^D(x) \equiv 1/(2\pi) \int d\phi e^{i\phi x}$ representation of the Dirac delta function and the definition of the smoothed density field. In order to simplify the notation later, we define a variable $\nu \equiv \delta_c/\sigma_R$ which quantify the density threshold in a unit of the root-mean-squared (r.m.s.) value of the smoothed density contrast.

Let us define the *partition (generating) functional* $Z[J]$ as

$$\begin{aligned}
Z[J] &\equiv \int [\mathcal{D}\delta] \mathcal{P}[\delta(\mathbf{x})] e^{i \int d^3y \delta(\mathbf{y}) \sum_{r=1}^N \phi_r W_R(|\mathbf{x}_r - \mathbf{y}|)} \\
&= \left\langle \exp \left[i \int d^3y \delta(\mathbf{y}) J(\mathbf{y}) \right] \right\rangle, \tag{K.7}
\end{aligned}$$

with following source function:

$$J(\mathbf{y}) \equiv \sum_{r=1}^N \phi_r W_R(\mathbf{x}_r - \mathbf{y}). \tag{K.8}$$

Then, from equation (K.7), we can define the n -point connected correlation function as

$$\xi^{(n)}(\mathbf{y}_1, \dots, \mathbf{y}_n) \equiv \langle \delta(\mathbf{y}_1) \cdots \delta(\mathbf{y}_n) \rangle_c \equiv \frac{1}{i^n} \frac{\delta^n \ln Z[J]}{\delta J(\mathbf{y}_1) \cdots \delta J(\mathbf{y}_n)} \Big|_{J=0}. \tag{K.9}$$

In other words, if we know all n -point correlation functions, we can reconstruct the partition function as a Taylor expansion:

$$\ln Z[J] = \sum_{n=2}^{\infty} \frac{i^n}{n!} \int d^3y_1 \cdots \int d^3y_n \xi^{(n)}(\mathbf{y}_1, \dots, \mathbf{y}_n) J(\mathbf{y}_1) \cdots J(\mathbf{y}_n). \tag{K.10}$$

By substituting the source function in equation (K.8), the partition functional becomes

$$\begin{aligned}
\ln Z[J] &= \sum_{n=2}^{\infty} \frac{i^n}{n!} \int d^3y_1 \cdots \int d^3y_n \xi^{(n)}(\mathbf{y}_1, \dots, \mathbf{y}_n) \\
&\quad \times \left[\sum_{r_1=1}^N \phi_{r_1} W_R(\mathbf{x}_{r_1} - \mathbf{y}_1) \right] \cdots \left[\sum_{r_n=1}^N \phi_{r_n} W_R(\mathbf{x}_{r_n} - \mathbf{y}_n) \right]. \tag{K.11}
\end{aligned}$$

The strategy of calculating P_N is following. For given n -point correlation functions of density field, we can calculate the partition functional from equation (K.11). By using the partition functional, P_N becomes

$$P_N(\mathbf{x}_1, \dots, \mathbf{x}_N) = \int_{\nu\sigma_R}^{\infty} d\alpha_1 \cdots \int_{\nu\sigma_R}^{\infty} d\alpha_N \int_{-\infty}^{\infty} \frac{d\phi_1}{(2\pi)^3} \cdots \int_{-\infty}^{\infty} \frac{d\phi_N}{(2\pi)^3} e^{-i \sum_{r=1}^N \alpha_r \phi_r} Z[J], \quad (\text{K.12})$$

which involves only ordinary integration.

Matarrese et al. (1986) provide the general solution for P_N , from there the authors reach the general formula for the N -point correlation function of peaks. We however find that the formula in Matarrese et al. (1986) is too abstract to be directly adopted without justification from the explicit calculation. Therefore, we shall explicitly show the solution for $N = 1, 2$ and 3 in the following sections. For the notational simplicity we denote the smoothed n -th order correlation function as $\xi_R^{(n)}$:

$$\xi_R^{(n)}(\mathbf{x}_1, \dots, \mathbf{x}_n) = \left\{ \prod_{r=1}^n \int d^3 y_r w_R(\mathbf{x}_r - \mathbf{y}_r) \right\} \xi^{(n)}(\mathbf{y}_1, \dots, \mathbf{y}_n). \quad (\text{K.13})$$

K.1.1 Calculation of P_1

Let's consider $N = 1$ case. The generating functional becomes

$$\begin{aligned} \ln Z[J] &= \sum_{n=2}^{\infty} \frac{i^n}{n!} \int d^3 y_1 \cdots \int d^3 y_n \xi^{(n)}(\mathbf{y}_1, \dots, \mathbf{y}_n) \prod_{r=1}^n \phi_1 W_R(\mathbf{x}_1 - \mathbf{y}_r), \end{aligned} \quad (\text{K.14})$$

and the integration over y_i s are simply a smoothed correlation function:

$$\begin{aligned} &\int d^3 y_1 \cdots \int d^3 y_n \xi^{(n)}(\mathbf{y}_1, \dots, \mathbf{y}_n) \prod_{r=1}^n W_R(\mathbf{x}_1 - \mathbf{y}_r) \\ &= \xi_R^{(n)} \left(\begin{array}{c} \mathbf{x}_1, \dots, \mathbf{x}_1 \\ n \text{ times} \end{array} \right). \end{aligned} \quad (\text{K.15})$$

As we have only one argument in the smoothed correlation function, we simply denote it as $\xi_R^{(n)}$ without arguments. Then, the generating functional is simplified as

$$\ln Z[J] = \sum_{n=2}^{\infty} \frac{i^n}{n!} \phi_1^n \xi_R^{(n)}, \quad (\text{K.16})$$

and we can calculate P_1 from equation (K.12).

$$P_1(\mathbf{x}_1) = \frac{1}{(2\pi)} \int_{\nu\sigma_R}^{\infty} d\alpha_1 \int_{-\infty}^{\infty} d\phi_1 \exp \left[-i\alpha_1\phi_1 + \sum_{n=2}^{\infty} \frac{i^n}{n!} \phi_1^n \xi_R^{(n)} \right]. \quad (\text{K.17})$$

Let us first consider the ϕ_i integration.

$$I \equiv \int_{-\infty}^{\infty} d\phi_1 \exp \left[-i\alpha_1\phi_1 + \sum_{n=2}^{\infty} \frac{i^n}{n!} \phi_1^n \xi_R^{(n)} \right]. \quad (\text{K.18})$$

We first calculate the quadratic term

$$I_0 = \int_{-\infty}^{\infty} d\phi_1 \exp \left[-i\alpha_1\phi_1 - \frac{1}{2} \phi_1^2 \sigma_R^2 \right] = \sqrt{\frac{2\pi}{\sigma_R^2}} \exp \left[-\frac{\alpha_1^2}{2\sigma_R^2} \right], \quad (\text{K.19})$$

then, we calculate the other terms in I , which include ϕ_1^n ($n > 3$), by taking n -th derivative of α_1 on I_0 as following.

$$\begin{aligned} I &= \exp \left[\sum_{n=3}^{\infty} \frac{(-1)^n}{n!} \xi_R^{(n)} \frac{\partial^n}{\partial \alpha_1^n} \right] \int_{-\infty}^{\infty} d\phi_1 e^{-\phi_1^2 \sigma_R^2 / 2 - i\alpha_1 \phi_1} \\ &= \frac{\sqrt{2\pi}}{\sigma_R} \exp \left[\sum_{n=3}^{\infty} \frac{(-1)^n}{n!} \xi_R^{(n)} \frac{\partial^n}{\partial \alpha_1^n} \right] \exp \left[-\frac{\alpha_1^2}{2\sigma_R^2} \right] \end{aligned} \quad (\text{K.20})$$

Now, P_1 becomes

$$\begin{aligned} P_1(\mathbf{x}_1) &= \frac{1}{\sqrt{2\pi}\sigma_R} \int_{\nu\sigma_R}^{\infty} d\alpha_1 \exp \left[\sum_{n=3}^{\infty} \frac{(-1)^n}{n!} \xi_R^{(n)} \frac{\partial^n}{\partial \alpha_1^n} \right] \exp \left[-\frac{\alpha_1^2}{2\sigma_R^2} \right] \\ &= \frac{1}{\sqrt{2\pi}} \int_{\nu}^{\infty} d\alpha'_1 \exp \left[\sum_{n=3}^{\infty} \frac{(-1)^n}{n!} \frac{\xi_R^{(n)}}{\sigma_R^n} \frac{\partial^n}{\partial \alpha'^n_1} \right] \exp \left[-\frac{\alpha'^2_1}{2} \right], \end{aligned} \quad (\text{K.21})$$

where in the second equality, we change the variable $\alpha'_i = \alpha_i/\sigma_R$. We use the Hermite polynomial, $H_n(x)$ to simplify the equation further. Especially, following two properties of Hermite polynomial is useful for our purpose.

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2} \quad (\text{K.22})$$

$$\lim_{x \rightarrow \infty} H_n(x) = 2^n x^n \quad (\text{K.23})$$

From equation (K.22), we find

$$(-1)^n \frac{d^n}{d\alpha'^n_1} e^{-\alpha'^2_1/2} = 2^{-n/2} e^{-\alpha'^2_1/2} H_n \left(\frac{\alpha'_1}{\sqrt{2}} \right),$$

and when the argument is large, ($\alpha'_1 \gg 1$) we approximate the derivative as

$$(-1)^n \frac{d^n}{d\alpha'_1} e^{-\alpha'^2_1/2} \rightarrow e^{-\alpha'^2_1/2} (\alpha'_1)^n.$$

Therefore, by using the asymptotic formula of Gaussian integration

$$\int_a^\infty f(x) e^{-x^2/2} \simeq \frac{f(a)}{a} e^{-a^2/2} + \mathcal{O}\left(\frac{1}{a^2}\right), \quad (\text{K.24})$$

we further simplify P_1 as

$$P_1(\mathbf{x}_1) \simeq \frac{1}{\sqrt{2\pi\nu^2}} \exp \left[\sum_{n=3}^{\infty} \frac{\nu^n}{n!} \frac{\xi_R^{(n)}}{\sigma_R^n} \right] e^{-\nu^2/2}. \quad (\text{K.25})$$

for the high peak limit ($\nu \gg 1$).

K.1.2 Calculation of P_2

In this section, we focus only on $N = 2$ case. The generating functional becomes

$$\begin{aligned} \ln Z[J] &= \sum_{n=2}^{\infty} \frac{i^n}{n!} \int d^3 y_1 \cdots \int d^3 y_n \xi^{(n)}(\mathbf{y}_1, \cdots, \mathbf{y}_n) \\ &\quad \times \prod_{r=1}^n \left[\phi_1 W_R(\mathbf{x}_1 - \mathbf{y}_r) + \phi_2 W_R(\mathbf{x}_2 - \mathbf{y}_r) \right] \\ &= \sum_{n=2}^{\infty} \frac{i^n}{n!} \int d^3 y_1 \cdots \int d^3 y_n \xi^{(n)}(\mathbf{y}_1, \cdots, \mathbf{y}_n) \sum_{m=0}^n \binom{n}{m} \phi_1^m \phi_2^{n-m} \\ &\quad \times \prod_{r_1=1}^m W_R(\mathbf{x}_1 - \mathbf{y}_{r_1}) \prod_{r_2=m+1}^n W_R(\mathbf{x}_2 - \mathbf{y}_{r_2}), \end{aligned} \quad (\text{K.26})$$

where, in the second line, we use the binomial expansion and the symmetry of $\xi^{(n)}$, namely, the correlation function does not depend on the order of argument. We can replace the integration over y_i s as a smoothed correlation function.

$$\begin{aligned} &\int d^3 y_1 \cdots \int d^3 y_n \xi^{(n)}(\mathbf{y}_1, \cdots, \mathbf{y}_n) \prod_{r_1=1}^m W_R(\mathbf{x}_1 - \mathbf{y}_{r_1}) \prod_{r_2=m+1}^n W_R(\mathbf{x}_2 - \mathbf{y}_{r_2}) \\ &= \xi_R^{(n)} \left(\begin{array}{cc} \mathbf{x}_1, \cdots, \mathbf{x}_1, & \mathbf{x}_2, \cdots, \mathbf{x}_2 \\ m \text{ times} & n-m \text{ times} \end{array} \right) \end{aligned} \quad (\text{K.27})$$

For notational simplicity we denote such a smoothed correlation function as $\xi_{R,m}^{(n)}$. Then, the generating functional becomes

$$\ln Z[J] = \sum_{n=2}^{\infty} i^n \sum_{m=0}^n \frac{\phi_1^m \phi_2^{n-m}}{m!(n-m)!} \xi_{R,m}^{(n)}. \quad (\text{K.28})$$

We calculate P_2 from equation (K.12)

$$P_2(\mathbf{x}_1, \mathbf{x}_2) = \int_{\nu\sigma_R}^{\infty} d\alpha_1 \int_{\nu\sigma_R}^{\infty} d\alpha_2 \int_{-\infty}^{\infty} \frac{d\phi_1}{(2\pi)} \int_{-\infty}^{\infty} \frac{d\phi_2}{(2\pi)} e^{-i(\alpha_1\phi_1 + \alpha_2\phi_2)} Z[J]. \quad (\text{K.29})$$

Let's first consider the ϕ_i integration.

$$I = \int_{-\infty}^{\infty} d\phi_1 \int_{-\infty}^{\infty} d\phi_2 \exp \left[-i \sum_{r=1}^2 \alpha_r \phi_r + \sum_{n=2}^{\infty} i^n \sum_{m=0}^n \frac{\phi_1^m \phi_2^{n-m}}{m!(n-m)!} \xi_{R,m}^{(n)} \right] \quad (\text{K.30})$$

As we have done in Section K.1.1, we first calculate the quadratic integration,

$$\begin{aligned} I_0 &= \int_{-\infty}^{\infty} d\phi_1 \int_{-\infty}^{\infty} d\phi_2 \exp \left[-i \sum_{r=1}^2 \alpha_r \phi_r - \left(\frac{1}{2} \phi_1^2 \sigma_R^2 + \frac{1}{2} \phi_2^2 \sigma_R^2 \right) \right] \\ &= \frac{2\pi}{\sigma_R^2} \exp \left[-\frac{1}{2} \sum_{r=1}^2 \frac{\alpha_r^2}{\sigma_R^2} \right]. \end{aligned} \quad (\text{K.31})$$

Then, ϕ_i integration simplifies to be a sum of successive derivatives on I_0 as below.

$$\begin{aligned} I &= \frac{2\pi}{\sigma_R^2} \exp \left[(-1)^2 \xi_R^{(2)}(x_{12}) \frac{\partial^2}{\partial \alpha_1 \partial \alpha_2} + \sum_{n=3}^{\infty} (-1)^n \sum_{m=0}^n \frac{\xi_{R,m}^{(n)}}{m!(n-m)!} \frac{\partial^n}{\partial \alpha_1^m \partial \alpha_2^{n-m}} \right] \\ &\quad \times \exp \left[-\frac{1}{2} \sum_{r=1}^2 \frac{\alpha_r^2}{\sigma_R^2} \right] \end{aligned} \quad (\text{K.32})$$

Now, we have to calculate the α_i integration. For the notational simplicity, we define $\alpha'_i \equiv \alpha_i/\sigma_R$, and $w_m^{(n)}$ as following.

$$\begin{cases} w_m^{(2)} = \xi_R^{(2)}(x_{12})/\sigma_R^2 & (m=1) \\ w_m^{(2)} = 0 & (m=0 \text{ or } 2) \\ w_m^{(n)} = \xi_{R,m}^{(n)}/\sigma_R^n & (n>2) \end{cases} \quad (\text{K.33})$$

Now the two-point probability $P_2(\mathbf{x}_1, \mathbf{x}_2)$ becomes

$$\begin{aligned} P_2(\mathbf{x}_1, \mathbf{x}_2) &= \frac{1}{2\pi} \int_{\nu}^{\infty} d\alpha'_1 \int_{\nu}^{\infty} d\alpha'_2 \exp \left[\sum_{n=2}^{\infty} (-1)^n \sum_{m=0}^n \frac{w_m^{(n)}}{m!(n-m)!} \frac{\partial^n}{\partial \alpha'_1{}^m \partial \alpha'_2{}^{n-m}} \right] \\ &\quad \times \exp \left[-\frac{1}{2} \sum_{r=1}^2 \alpha_r'^2 \right], \end{aligned} \quad (\text{K.34})$$

Again, we use the Hermite polynomial, and take the high peak limit of $\nu \gg 1$

$$P_2(\mathbf{x}_1, \mathbf{x}_2) \simeq \frac{1}{2\pi\nu^2} \exp \left[\sum_{n=2}^{\infty} \sum_{m=0}^n \frac{w_m^{(n)} \nu^n}{m!(n-m)!} \right] e^{-\nu^2}$$

K.1.3 Calculation of P_3

Let's calculate for $P_3(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$:

$$P_3(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = \frac{1}{(2\pi)^3} \prod_{i=1}^3 \int_{\nu\sigma_R}^{\infty} d\alpha_i \int_{-\infty}^{\infty} d\phi_i e^{-i \sum_{r=1}^3 \alpha_r \phi_r} Z[J]. \quad (\text{K.35})$$

We first calculate the partition functional

$$\ln Z[J] = \sum_{n=2}^{\infty} \frac{i^n}{n!} \left(\prod_{i=1}^n \int d^3 y_i \right) \xi^{(n)}(\mathbf{y}_1, \dots, \mathbf{y}_n) \prod_{r=1}^n \left[\sum_{j=1}^3 \phi_j W_R(\mathbf{x}_j - \mathbf{y}_r) \right]. \quad (\text{K.36})$$

By using a multinomial expansion theorem,

$$\begin{aligned} & \prod_{r=1}^n \left[\sum_{j=1}^3 \phi_j W_R(\mathbf{x}_j - \mathbf{y}_r) \right] \\ &= \sum_{m_1=0}^n \sum_{m_2=0}^{n-m_1} \frac{n!}{m_1! m_2! (n-m_1-m_2)!} \phi_1^{m_1} \phi_2^{m_2} \phi_3^{n-m_1-m_2} \\ & \quad \times \prod_{r_1 \in [m_1]} W_R(\mathbf{x}_1 - \mathbf{y}_{r_1}) \prod_{r_2 \in [m_2]} W_R(\mathbf{x}_2 - \mathbf{y}_{r_2}) \prod_{r_3 \in [m_3]} W_R(\mathbf{x}_3 - \mathbf{y}_{r_3}) \end{aligned} \quad (\text{K.37})$$

where, $m_3 \equiv n - m_1 - m_2$, and $[m_i]$ stands for the set of indexes

$$[m_i] = \{a_k | 1 \leq a_k \leq n, k = 1, \dots, m_i\},$$

and empty when $m_i = 0$. Note that, $[m_i]$ are mutually exclusive, that is $[m_i] \cap [m_j] = \{\}$, and $[m_1] \cup [m_2] \cup [m_3] = \{1, 2, \dots, n\}$. For example, when $n=2$, it becomes

$$\begin{aligned} & \prod_{r=1}^2 \left[\sum_{j=1}^3 \phi_j W_R(\mathbf{x}_j - \mathbf{y}_r) \right] \\ &= \phi_1^2 W_R(x_1 - y_1) W_R(x_1 - y_2) + \phi_2^2 W_R(x_2 - y_1) W_R(x_2 - y_2) \\ & \quad + \phi_3^2 W_R(x_3 - y_1) W_R(x_3 - y_2) + \phi_1 \phi_2 W_R(x_1 - y_1) W_R(x_2 - y_2) \\ & \quad + \phi_1 \phi_2 W_R(x_1 - y_2) W_R(x_2 - y_1) + \phi_2 \phi_3 W_R(x_2 - y_1) W_R(x_3 - y_2) \\ & \quad + \phi_3 \phi_2 W_R(x_2 - y_2) W_R(x_3 - y_1) + \phi_1 \phi_3 W_R(x_1 - y_1) W_R(x_3 - y_2) \\ & \quad + \phi_1 \phi_3 W_R(x_1 - y_2) W_R(x_3 - y_1), \end{aligned} \quad (\text{K.38})$$

and we can write $n = 2$ component of $\ln Z[J]$ as

$$\begin{aligned}
& -\frac{1}{2} \int d^3 y_1 \int d^3 y_2 \xi^{(2)}(\mathbf{y}_1, \mathbf{y}_2) \times \text{eq. [K.38]} \\
& = -\frac{1}{2} \left[(\phi_1^2 + \phi_2^2 + \phi_3^2) \sigma_R^2 \right. \\
& \quad \left. + 2\phi_1 \phi_2 \xi_R^{(2)}(\mathbf{x}_1, \mathbf{x}_2) + 2\phi_2 \phi_3 \xi_R^{(2)}(\mathbf{x}_2, \mathbf{x}_3) + 2\phi_3 \phi_1 \xi_R^{(2)}(\mathbf{x}_3, \mathbf{x}_1) \right]. \tag{K.39}
\end{aligned}$$

Therefore, we rewrite $Z[J]$ as

$$\ln Z[J] = \sum_{n=2}^{\infty} \frac{i^n}{n!} \sum_{m_1=0}^n \sum_{m_2=0}^{n-m_1} \frac{n!}{m_1! m_2! m_3!} \phi_1^{m_1} \phi_2^{m_2} \phi_3^{m_3} \xi_{R, m_1 m_2 m_3}^{(n)}, \tag{K.40}$$

where

$$\xi_{R, m_1 m_2 m_3}^{(n)} = \xi_R^{(n)} \left(\begin{array}{ccc} \mathbf{x}_1, \dots, \mathbf{x}_1, & \mathbf{x}_2, \dots, \mathbf{x}_2, & \mathbf{x}_3, \dots, \mathbf{x}_3, \\ m_1 \text{ times} & m_2 \text{ times} & m_3 \text{ times} \end{array} \right).$$

For example, for $n = 2$ case we write

$$\begin{aligned}
& \left. \frac{i^n}{n!} \sum_{m_1=0}^n \sum_{m_2=0}^{n-m_1} \frac{n!}{m_1! m_2! m_3!} \phi_1^{m_1} \phi_2^{m_2} \phi_3^{m_3} \xi_{R, m_1 m_2 m_3}^{(n)} \right|_{n=2} \\
& = -\frac{1}{2} \left[\phi_1^2 \xi_{R, 200}^{(2)} + \phi_2^2 \xi_{R, 020}^{(2)} + \phi_3^2 \xi_{R, 002}^{(2)} + 2\phi_1 \phi_2 \xi_{R, 110}^{(2)} + 2\phi_2 \phi_3 \xi_{R, 011}^{(2)} + 2\phi_3 \phi_1 \xi_{R, 101}^{(2)} \right].
\end{aligned}$$

We follow the same procedure as we have calculated P_1 and P_2 in the previous section.

First, we calculate ϕ_i integration by using a quadratic integration of

$$\begin{aligned}
I_0 &= \int_{-\infty}^{\infty} d\phi_1 \int_{-\infty}^{\infty} d\phi_2 \int_{-\infty}^{\infty} d\phi_3 \exp \left[-i \sum_{r=1}^3 \alpha_r \phi_r - \frac{1}{2} \sum_{r=1}^3 \phi_r^2 \sigma_R^2 \right] \\
&= \frac{(2\pi)^{3/2}}{\sigma_R^3} \exp \left[-\frac{1}{2} \sum_{r=1}^3 \frac{\alpha_r^2}{\sigma_R^2} \right]. \tag{K.41}
\end{aligned}$$

Then, we can replace the ϕ_r -integration by applying successive differentiations of $(-i\alpha_r)$.

That is,

$$\begin{aligned}
I &= \exp \left[\sum_{i \neq j} \xi_R^{(2)}(\mathbf{x}_i, \mathbf{x}_j) \frac{\partial}{\partial \alpha_i} \frac{\partial}{\partial \alpha_j} \right. \\
& \quad \left. + \sum_{n=3}^{\infty} \sum_{m_1=0}^n \sum_{m_2=0}^{n-m_1} \frac{(-1)^n \xi_{R, m_1 m_2 m_3}^{(n)}}{m_1! m_2! m_3!} \frac{\partial^n}{\partial \alpha_1^{m_1} \partial \alpha_2^{m_2} \partial \alpha_3^{m_3}} \right] I_0. \tag{K.42}
\end{aligned}$$

As we do not apply the derivative operator to the terms appears in I_0 , we define $w_{m_1 m_2 m_3}^{(n)}$ as following.

$$\begin{cases} w_{m_1, m_2, m_3}^{(2)} = \xi_R^{(2)}(x_{ij})/\sigma_R^2 & (m_1 \neq 2, m_2 \neq 2, m_3 \neq 2) \\ w_{m_1, m_2, m_3}^{(2)} = 0 & (\text{otherwise}) \\ w_{m_1, m_2, m_3}^{(n)} = \xi_{R, m_1 m_2 m_3}^{(n)}/\sigma_R^n & (n > 2) \end{cases} \quad (\text{K.43})$$

By using the new notation, and changing the variable to $\alpha'_i \equiv \alpha_i/\sigma_R$, P_3 becomes

$$\begin{aligned} & P_3(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) \\ &= \frac{1}{(2\pi)^{3/2}} \prod_{r=1}^3 \int_{\nu}^{\infty} d\alpha'_r \\ & \times \exp \left[\sum_{n=2}^{\infty} \sum_{m_1, m_2} \frac{(-1)^n w_{m_1 m_2 m_3}^{(n)}}{m_1! m_2! m_3!} \frac{\partial^n}{\partial \alpha'^{m_1}_1 \partial \alpha'^{m_2}_2 \partial \alpha'^{m_3}_3} \right] \exp \left[-\frac{1}{2} \sum_{r=1}^3 \alpha'^2_r \right]. \end{aligned}$$

Using the Hermite polynomial, and, finally, we imposing the high peak condition of $\nu \gg 1$, we find

$$P_3(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) \simeq \frac{1}{(2\pi\nu^2)^{3/2}} \exp \left[\sum_{n=2}^{\infty} \sum_{m_1, m_2, m_3} \frac{w_{m_1 m_2 m_3}^{(n)} \nu^n}{m_1! m_2! m_3!} \right] e^{-\frac{3}{2}\nu^2}. \quad (\text{K.44})$$

K.2 The two point correlation function of peaks

We calculate the two point correlation function of peaks by substituting P_1 from equation (K.25) and P_2 from equation (K.35) into equation (K.3):

$$\begin{aligned} \xi_p(\mathbf{x}_1, \mathbf{x}_2) &= \frac{P_2(\mathbf{x}_1, \mathbf{x}_2)}{P_1^2} - 1 \\ &= \exp \left[\sum_{n=2}^{\infty} \sum_{m=0}^n \frac{w_m^{(n)} \nu^n}{m!(n-m)!} - 2 \sum_{n=3}^{\infty} \frac{\nu^n}{n!} \frac{\xi_R^{(n)}}{\sigma_R^n} \right] - 1, \end{aligned} \quad (\text{K.45})$$

where $w_m^{(n)}$ and $\xi_R^{(n)}$ are defined in equation (K.33) and equation (K.15), respectively. We further simplify the notation by using $w_0^{(n)} = w_n^{(n)} = \xi_R^{(n)}/\sigma_R^n$ for $n \geq 3$, and $w_0^{(2)} = w_2^{(2)} = 0$.

$$\begin{aligned}
\xi_p(\mathbf{x}_1, \mathbf{x}_2) &= \exp \left[\sum_{n=2}^{\infty} \sum_{m=0}^n \frac{w_m^{(n)} \nu^n}{m!(n-m)!} - \sum_{n=3}^{\infty} \frac{\nu^n w_0^{(n)}}{n!} - \sum_{n=3}^{\infty} \frac{\nu^n w_n^{(n)}}{n!} \right] - 1 \\
&= \exp \left[\sum_{n=2}^{\infty} \sum_{m=1}^{n-1} \frac{w_m^{(n)} \nu^n}{m!(n-m)!} \right] - 1 \\
&= \exp \left[\sum_{n=2}^{\infty} \sum_{m=1}^{n-1} \binom{n}{m} \frac{\nu^n}{n!} \frac{\xi_{R,m}^{(n)}}{\sigma_R^n} \right] - 1
\end{aligned} \tag{K.46}$$

This result was first derived in Grinstein & Wise (1986) with a following ansatz

$$n_p(\mathbf{x}) = C \exp \left[T \int d^3 y \delta(\mathbf{y}) W(\mathbf{x} - \mathbf{y}) \right]. \tag{K.47}$$

It was only after the full calculation of Matarrese et al. (1986) that $T = \nu/\sigma_R$ is identified.

Note that for Gaussian case, where all the higher order correlation function vanishes ($\xi^{(n)} = 0$ for $n > 3$) equation (K.46) coincides with the result in Politzer & Wise (1984):

$$\xi_p^G(\mathbf{x}_1, \mathbf{x}_2) = \exp \left[\frac{\xi_R^{(2)}(x_{12})}{\sigma_R^2} \nu^2 \right] - 1 \simeq \frac{\nu^2}{\sigma_R^2} \xi_R^{(2)}(x_{12}), \tag{K.48}$$

and, in the second equality, we can also reproduce the ‘Lagrangian bias’, $b_L = \nu/\sigma_R$, by taking the large scale limit ($\xi_R \ll 1$).

Recently, in the light of the scale dependent bias induced by primordial non-Gaussianity (see, Section I.3), Matarrese & Verde (2008) study the large scale limit of equation (K.46). For the large separation, we can expand the exponential in equation (K.46), and the two-point correlation function including the leading order correction due to the non-Gaussianity becomes

$$\xi_p(|\mathbf{x}_1 - \mathbf{x}_2|) \simeq \frac{\nu^2}{\sigma_R^2} \xi_R^{(2)}(\mathbf{x}_1, \mathbf{x}_2) + \frac{\nu^3}{\sigma_R^3} \xi_R^{(3)}(\mathbf{x}_1, \mathbf{x}_1, \mathbf{x}_2). \tag{K.49}$$

Fourier transform of the non-Gaussian correction term is related to the matter bispectrum as

$$\begin{aligned}
&\xi_R^{(3)}(\mathbf{x}_1, \mathbf{x}_1, \mathbf{x}_2) \\
&= \int \frac{d^3 q_1}{(2\pi)^3} \int \frac{d^3 q_2}{(2\pi)^3} \int \frac{d^3 q_3}{(2\pi)^3} \langle \tilde{\delta}_R(\mathbf{q}_1) \tilde{\delta}_R(\mathbf{q}_2) \tilde{\delta}_R(\mathbf{q}_3) \rangle e^{i\mathbf{q}_1 \cdot \mathbf{x}_1} e^{i\mathbf{q}_2 \cdot \mathbf{x}_1} e^{i\mathbf{q}_3 \cdot \mathbf{x}_2} \\
&= \int \frac{d^3 q_1}{(2\pi)^3} \int \frac{d^3 q_2}{(2\pi)^3} B_R(\mathbf{q}_1, \mathbf{q}_2, -(\mathbf{q}_1 + \mathbf{q}_2)) e^{i(\mathbf{q}_1 + \mathbf{q}_2) \cdot (\mathbf{x}_1 - \mathbf{x}_2)}.
\end{aligned} \tag{K.50}$$

Therefore, by taking Fourier transformation of equation (K.49), we find the power spectrum of peaks as

$$\begin{aligned}
P_p(k) &= \int d^3r \xi_p(r) e^{-i\mathbf{k}\cdot\mathbf{r}} \\
&= \frac{\nu^2}{\sigma_R^2} \int d^3r \xi_R^{(2)}(r) e^{-i\mathbf{k}\cdot\mathbf{r}} \\
&\quad + \frac{\nu^3}{\sigma_R^3} \int d^3r \int \frac{d^3q_1}{(2\pi)^3} \int \frac{d^3q_2}{(2\pi)^3} B_R(\mathbf{q}_1, \mathbf{q}_2, -(\mathbf{q}_1 + \mathbf{q}_2)) e^{i(\mathbf{q}_1 + \mathbf{q}_2 - \mathbf{k})\cdot\mathbf{r}} \\
&= \frac{\nu^2}{\sigma_R^2} P_R(k) + \frac{\nu^3}{\sigma_R^3} \int \frac{d^3q}{(2\pi)^3} B_R(\mathbf{q}, -\mathbf{k}, \mathbf{k} - \mathbf{q})
\end{aligned} \tag{K.51}$$

where $\mathbf{r} = \mathbf{x}_1 - \mathbf{x}_2$.

K.3 The three point correlation function of peaks

We calculate the three point function from the relation between P_3 and correlation functions in equation (K.4). By using the calculation of P_1 in equation (K.25) and P_3 in equation (K.44), we calculate P_3/P_1^3 (in the high peak limit $\nu \gg 1$, and on large scales,

$\xi_R^{(n)}(\mathbf{x}_1, \dots, \mathbf{x}_n) \ll 1$) up to four-point function order:

$$\begin{aligned}
& \frac{P_3(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)}{P_1^3} \\
&= \exp \left[\sum_{n=2}^{\infty} \sum_{m_1=0}^n \sum_{m_2=0}^{n-m_1} \frac{\nu^n w_{m_1 m_2 m_3}^{(n)}}{m_1! m_2! (n - m_1 - m_2)!} - 3 \sum_{n=3}^{\infty} \frac{\nu^n}{n!} \frac{\xi_R^{(n)}}{\sigma_R^n} \right] \\
&\simeq 1 + \sum_{n=2}^{\infty} \sum_{m_1=0}^n \sum_{m_2=0}^{n-m_1} \frac{\nu^n w_{m_1 m_2 m_3}^{(n)}}{m_1! m_2! (n - m_1 - m_2)!} - 3 \sum_{n=3}^{\infty} \frac{\nu^n}{n!} \frac{\xi_R^{(n)}}{\sigma_R^n} \\
&\simeq 1 + \frac{\nu^2}{\sigma_R^2} \left[\xi_R^{(2)}(x_{12}) + \xi_R^{(2)}(x_{23}) + \xi_R^{(2)}(x_{31}) \right] \\
&\quad + \frac{\nu^3}{\sigma_R^3} \left[\frac{\xi_{R,003}^{(3)}}{6} + \frac{\xi_{R,012}^{(3)}}{2} + \frac{\xi_{R,021}^{(3)}}{2} + \frac{\xi_{R,030}^{(3)}}{6} + \frac{\xi_{R,102}^{(3)}}{2} + \frac{\xi_{R,111}^{(3)}}{1} + \frac{\xi_{R,120}^{(3)}}{2} + \frac{\xi_{R,201}^{(3)}}{2} \right. \\
&\quad \left. + \frac{\xi_{R,210}^{(3)}}{2} + \frac{\xi_{R,300}^{(3)}}{6} \right] - 3 \frac{\nu^3}{\sigma_R^3} \frac{\xi_R^{(3)}}{6} \\
&\quad + \frac{\nu^4}{\sigma_R^4} \left[\frac{\xi_{R,004}^{(4)}}{24} + \frac{\xi_{R,013}^{(4)}}{6} + \frac{\xi_{R,022}^{(4)}}{4} + \frac{\xi_{R,031}^{(4)}}{6} + \frac{\xi_{R,040}^{(4)}}{24} + \frac{\xi_{R,103}^{(4)}}{6} + \frac{\xi_{R,112}^{(4)}}{2} + \frac{\xi_{R,121}^{(4)}}{2} \right. \\
&\quad \left. + \frac{\xi_{R,130}^{(4)}}{6} + \frac{\xi_{R,202}^{(4)}}{4} + \frac{\xi_{R,211}^{(4)}}{2} + \frac{\xi_{R,220}^{(4)}}{4} + \frac{\xi_{R,301}^{(4)}}{6} + \frac{\xi_{R,310}^{(4)}}{6} + \frac{\xi_{R,400}^{(4)}}{24} \right] - 3 \frac{\nu^4}{\sigma_R^4} \frac{\xi_R^{(4)}}{24} \\
&= 1 + \frac{\nu^2}{\sigma_R^2} \left[\xi_R^{(2)}(x_{12}) + \xi_R^{(2)}(x_{23}) + \xi_R^{(2)}(x_{31}) \right] \\
&\quad + \frac{\nu^3}{\sigma_R^3} \left[\xi_R^{(3)}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) + \xi_R^{(3)}(\mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_3) + \xi_R^{(3)}(\mathbf{x}_1, \mathbf{x}_1, \mathbf{x}_3) + \xi_R^{(3)}(\mathbf{x}_1, \mathbf{x}_1, \mathbf{x}_2) \right] \\
&\quad + \frac{\nu^4}{\sigma_R^4} \left[\frac{1}{2} \xi_R^{(4)}(\mathbf{x}_1, \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) + \frac{1}{2} \xi_R^{(4)}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_3) + \frac{1}{2} \xi_R^{(4)}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_2, \mathbf{x}_3) \right. \\
&\quad + \frac{1}{4} \xi_R^{(4)}(\mathbf{x}_2, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_3) + \frac{1}{3} \xi_R^{(4)}(\mathbf{x}_2, \mathbf{x}_2, \mathbf{x}_2, \mathbf{x}_3) + \frac{1}{4} \xi_R^{(4)}(\mathbf{x}_1, \mathbf{x}_1, \mathbf{x}_3, \mathbf{x}_3) \\
&\quad \left. + \frac{1}{3} \xi_R^{(4)}(\mathbf{x}_1, \mathbf{x}_1, \mathbf{x}_1, \mathbf{x}_3) + \frac{1}{4} \xi_R^{(4)}(\mathbf{x}_1, \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_2) + \frac{1}{3} \xi_R^{(4)}(\mathbf{x}_1, \mathbf{x}_1, \mathbf{x}_1, \mathbf{x}_2) \right]. \quad (\text{K.52})
\end{aligned}$$

In order to calculate the three-point correlation function of peaks, we need to subtract the two-point correlation functions from equation (K.52). As we expand P_3/P_1^3 up to four-point

function order, we also expand equation (K.46) in the same order as

$$\begin{aligned}
\xi_p(\mathbf{x}_1, \mathbf{x}_2) &\simeq \sum_{n=2}^{\infty} \sum_{m=0}^n \frac{\nu^n w_m^{(n)}}{m!(n-m)!} - 2 \sum_{n=3}^{\infty} \frac{\nu^n}{n!} \frac{\xi_R^{(n)}}{\sigma_R^n} \\
&= \frac{\nu^2}{\sigma_R^2} \xi_R^{(2)}(x_{12}) + \frac{\nu^3}{\sigma_R^3} \left[\frac{\xi_{R,1}^{(3)}}{2} + \frac{\xi_{R,2}^{(3)}}{2} \right] + \frac{\nu^4}{\sigma_R^4} \left[\frac{\xi_{R,1}^{(4)}}{6} + \frac{\xi_{R,2}^{(4)}}{4} + \frac{\xi_{R,3}^{(4)}}{6} \right] + \dots \\
&= \frac{\nu^2}{\sigma_R^2} \xi_R^{(2)}(x_{12}) + \frac{\nu^3}{\sigma_R^3} \xi_R^{(3)}(\mathbf{x}_1, \mathbf{x}_1, \mathbf{x}_2) \\
&\quad + \frac{\nu^4}{\sigma_R^4} \left[\frac{1}{3} \xi_R^{(4)}(\mathbf{x}_1, \mathbf{x}_1, \mathbf{x}_1, \mathbf{x}_2) + \frac{1}{4} \xi_R^{(4)}(\mathbf{x}_1, \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_2) \right] + \dots \tag{K.53}
\end{aligned}$$

Subtracting the two-point correlation function, we find that the three-point correlation function is

$$\begin{aligned}
\zeta_p(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) &= \frac{\nu^3}{\sigma_R^3} \xi_R^{(3)}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) + \frac{\nu^4}{\sigma_R^4} \left[\xi_R^{(2)}(x_{12}) \xi_R^{(2)}(x_{23}) + (\text{cyclic}) \right] \\
&\quad + \frac{1}{2} \frac{\nu^4}{\sigma_R^4} \left[\xi_R^{(4)}(\mathbf{x}_1, \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) + (\text{cyclic}) \right]. \tag{K.54}
\end{aligned}$$

Bispectrum is the Fourier transform of the three-point correlation function. As universe is statistically isotropic, the three-point correlation function only depends on the shape and size of triangles constructed by three points $(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$, which can be fully specified by using two vectors, $\mathbf{r} \equiv \mathbf{x}_1 - \mathbf{x}_3$ and $\mathbf{s} \equiv \mathbf{x}_2 - \mathbf{x}_3$. That is,

$$\zeta(\mathbf{r}, \mathbf{s}) = \zeta(r, s, |\mathbf{r} - \mathbf{s}|). \tag{K.55}$$

In order to see this, let's Fourier transform the three point correlation function.

$$\begin{aligned}
&\zeta(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) \\
&= \int \frac{d^3 q_1}{(2\pi)^3} \int \frac{d^3 q_2}{(2\pi)^3} \int \frac{d^3 q_3}{(2\pi)^3} \langle \delta(\mathbf{q}_1) \delta(\mathbf{q}_2) \delta(\mathbf{q}_3) \rangle e^{i\mathbf{q}_1 \cdot \mathbf{x}_1} e^{i\mathbf{q}_2 \cdot \mathbf{x}_2} e^{i\mathbf{q}_3 \cdot \mathbf{x}_3} \\
&= \int \frac{d^3 q_1}{(2\pi)^3} \int \frac{d^3 q_2}{(2\pi)^3} \int \frac{d^3 q_3}{(2\pi)^3} (2\pi)^3 B(q_1, q_2, q_3) \delta^D(\mathbf{q}_{123}) e^{i\mathbf{q}_1 \cdot \mathbf{x}_1} e^{i\mathbf{q}_2 \cdot \mathbf{x}_2} e^{i\mathbf{q}_3 \cdot \mathbf{x}_3} \\
&= \int \frac{d^3 q_1}{(2\pi)^3} \int \frac{d^3 q_2}{(2\pi)^3} B(q_1, q_2, |\mathbf{q}_1 + \mathbf{q}_2|) e^{i\mathbf{q}_1 \cdot (\mathbf{x}_1 - \mathbf{x}_3)} e^{i\mathbf{q}_2 \cdot (\mathbf{x}_2 - \mathbf{x}_3)} \\
&= \int \frac{d^3 q}{(2\pi)^3} \int \frac{d^3 q'}{(2\pi)^3} B(\mathbf{q}, \mathbf{q}') e^{i\mathbf{q} \cdot \mathbf{r}} e^{i\mathbf{q}' \cdot \mathbf{s}} = \zeta(\mathbf{r}, \mathbf{s}) \tag{K.56}
\end{aligned}$$

Therefore, we calculate the bispectrum by inverse-Fourier transform of the three-point correlation function:

$$B(\mathbf{q}, \mathbf{q}') = \int d^3 r \int d^3 s \zeta(\mathbf{r}, \mathbf{s}) e^{-i\mathbf{q} \cdot \mathbf{r}} e^{-i\mathbf{q}' \cdot \mathbf{s}} \tag{K.57}$$

Finally, let's think about the Fourier transformation of *connected* four point function,

$$\begin{aligned}
& \xi_R^{(4)}(\mathbf{x}_1, \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) \\
&= \int \frac{d^3 q_1}{(2\pi)^3} \int \frac{d^3 q_2}{(2\pi)^3} \int \frac{d^3 q_3}{(2\pi)^3} \int \frac{d^3 q_4}{(2\pi)^3} \langle \tilde{\delta}_R(\mathbf{q}_1) \tilde{\delta}_R(\mathbf{q}_2) \tilde{\delta}_R(\mathbf{q}_3) \tilde{\delta}_R(\mathbf{q}_4) \rangle_c \\
&\quad \times e^{i\mathbf{q}_1 \cdot \mathbf{x}_1} e^{i\mathbf{q}_2 \cdot \mathbf{x}_1} e^{i\mathbf{q}_3 \cdot \mathbf{x}_2} e^{i\mathbf{q}_4 \cdot \mathbf{x}_3} \\
&= \int \frac{d^3 q_1}{(2\pi)^3} \int \frac{d^3 q_2}{(2\pi)^3} \int \frac{d^3 q_3}{(2\pi)^3} \int \frac{d^3 q_4}{(2\pi)^3} (2\pi)^3 T_R(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3, \mathbf{q}_4) \delta^D(\mathbf{q}_{1234}) \\
&\quad \times e^{i\mathbf{q}_1 \cdot \mathbf{x}_1} e^{i\mathbf{q}_2 \cdot \mathbf{x}_1} e^{i\mathbf{q}_3 \cdot \mathbf{x}_2} e^{i\mathbf{q}_4 \cdot \mathbf{x}_3} \\
&= \int \frac{d^3 q_1}{(2\pi)^3} \int \frac{d^3 q_2}{(2\pi)^3} \int \frac{d^3 q_3}{(2\pi)^3} T_R(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3, -\mathbf{q}_{123}) e^{i\mathbf{q}_1 \cdot (\mathbf{x}_1 - \mathbf{x}_3)} e^{i\mathbf{q}_2 \cdot (\mathbf{x}_1 - \mathbf{x}_3)} e^{i\mathbf{q}_3 \cdot (\mathbf{x}_2 - \mathbf{x}_3)} \\
&= \int \frac{d^3 q_1}{(2\pi)^3} \int \frac{d^3 q_2}{(2\pi)^3} \int \frac{d^3 q_3}{(2\pi)^3} T_R(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3, -\mathbf{q}_{123}) e^{i(\mathbf{q}_1 + \mathbf{q}_2) \cdot \mathbf{r}} e^{i\mathbf{q}_3 \cdot \mathbf{s}} \tag{K.58}
\end{aligned}$$

Similarly,

$$\begin{aligned}
& \xi_R^{(4)}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_2, \mathbf{x}_3) \\
&= \int \frac{d^3 q_1}{(2\pi)^3} \int \frac{d^3 q_2}{(2\pi)^3} \int \frac{d^3 q_3}{(2\pi)^3} T_R(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3, -\mathbf{q}_{123}) e^{i\mathbf{q}_1 \cdot \mathbf{r}} e^{i(\mathbf{q}_2 + \mathbf{q}_3) \cdot \mathbf{s}} \tag{K.59}
\end{aligned}$$

$$\begin{aligned}
& \xi_R^{(4)}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_3) \\
&= \int \frac{d^3 q_1}{(2\pi)^3} \int \frac{d^3 q_2}{(2\pi)^3} \int \frac{d^3 q_3}{(2\pi)^3} T_R(\mathbf{q}_1, -\mathbf{q}_{123}, \mathbf{q}_2, \mathbf{q}_3) e^{i\mathbf{q}_1 \cdot (\mathbf{r} - \mathbf{s})} e^{-i(\mathbf{q}_2 + \mathbf{q}_3) \cdot \mathbf{s}} \tag{K.60}
\end{aligned}$$

By using the Fourier relations above, we can easily calculate the Fourier transfor-

mation of equation (K.54).

$$\begin{aligned}
& B_p(\mathbf{k}_1, \mathbf{k}_2) \\
&= \int d^3r \int d^3s \zeta_p(\mathbf{r}, \mathbf{s}) e^{-i\mathbf{k}_1 \cdot \mathbf{r}} e^{-i\mathbf{k}_2 \cdot \mathbf{s}} \\
&= \frac{\nu^3}{\sigma_R^3} B_R(\mathbf{k}_1, \mathbf{k}_2) + \frac{\nu^4}{\sigma_R^4} \left[P_R(k_1)P_R(k_2) + P_R(k_2)P_R(k_3) + P_R(k_3)P_R(k_1) \right] \\
&\quad + \frac{\nu^4}{2\sigma_R^4} \int d^3r \int d^3s \int \frac{d^3q_1}{(2\pi)^3} \int \frac{d^3q_2}{(2\pi)^3} \int \frac{d^3q_3}{(2\pi)^3} \\
&\quad \times \left[T_R(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3, -\mathbf{q}_{123}) e^{i\mathbf{q}_{12} \cdot \mathbf{r}} e^{i\mathbf{q}_3 \cdot \mathbf{s}} e^{-i\mathbf{k}_1 \cdot \mathbf{r}} e^{-i\mathbf{k}_2 \cdot \mathbf{s}} \right. \\
&\quad + T_R(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3, -\mathbf{q}_{123}) e^{i\mathbf{q}_1 \cdot \mathbf{r}} e^{i\mathbf{q}_{23} \cdot \mathbf{s}} e^{-i\mathbf{k}_1 \cdot \mathbf{r}} e^{-i\mathbf{k}_2 \cdot \mathbf{s}} \\
&\quad \left. + T_R(\mathbf{q}_1, -\mathbf{q}_{123}, \mathbf{q}_2, \mathbf{q}_3) e^{i\mathbf{q}_1 \cdot (\mathbf{r}-\mathbf{s})} e^{-i\mathbf{q}_{23} \cdot \mathbf{s}} e^{-i\mathbf{k}_1 \cdot \mathbf{r}} e^{-i\mathbf{k}_2 \cdot \mathbf{s}} \right] \\
&= \frac{\nu^3}{\sigma_R^3} B_R(\mathbf{k}_1, \mathbf{k}_2) + \frac{\nu^4}{\sigma_R^4} \left[P_R(k_1)P_R(k_2) + (2 \text{ cyclic}) \right] \\
&\quad + \frac{\nu^4}{2\sigma_R^4} \int \frac{d^3q}{(2\pi)^3} \left[T_R(\mathbf{q}, \mathbf{k}_1 - \mathbf{q}, \mathbf{k}_2, \mathbf{k}_3) + (2 \text{ cyclic}) \right], \tag{K.61}
\end{aligned}$$

where $\mathbf{k}_3 = -\mathbf{k}_2 - \mathbf{k}_1$.

Note that, even in the absence of the primordial non-Gaussianity, matter bispectrum $B_R(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)$ is non-zero due to the gravitational instability, and that is given by

$$B_R(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = 2F_2^{(s)}(\mathbf{k}_1, \mathbf{k}_2)P_R(\mathbf{k}_1)P_R(\mathbf{k}_2) + (\text{cyclic}). \tag{K.62}$$

Therefore, the galaxy bispectrum is given by

$$\begin{aligned}
B_p(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) &= \frac{\nu^3}{\sigma_R^3} \left[2F_2^{(s)}(\mathbf{k}_1, \mathbf{k}_2)P_R(\mathbf{k}_1)P_R(\mathbf{k}_2) + (\text{cyclic}) \right] \\
&\quad + \frac{\nu^4}{\sigma_R^4} \left[P_R(k_1)P_R(k_2) + (\text{cyclic}) \right]. \tag{K.63}
\end{aligned}$$

in the large scale where we expect that halo/galaxy bias is linear.

Appendix L

Mean Tangential Shear and its covariance

In this appendix, we derive the mean tangential shear [Eq. (6.9)] and its covariance [Eq. (6.16)].

L.1 Derivation of the mean tangential shear

One may write down the observed tangential shears at a given distance from a lens halo, $\boldsymbol{\theta}$, averaged over N_L lens halos as

$$\gamma_t^h(\boldsymbol{\theta}) = \frac{1}{N_L} \int d^2\hat{\mathbf{n}} \left[\sum_i^{N_L} \delta_D(\hat{\mathbf{n}} - \hat{\mathbf{n}}_i) \right] \gamma_t(\hat{\mathbf{n}} + \boldsymbol{\theta}), \quad (\text{L.1})$$

where δ_D is the delta function, and i denotes the location of lens halos. Note that we have not azimuthally averaged the tangential shears yet. The ensemble average of γ_t^h yields the number-weighted average of the tangential shear:

$$\langle \gamma_t^h \rangle(\boldsymbol{\theta}) = \frac{1}{N_L} \int d^2\hat{\mathbf{n}} \langle n_L(\hat{\mathbf{n}}) \gamma_t(\hat{\mathbf{n}} + \boldsymbol{\theta}) \rangle, \quad (\text{L.2})$$

where $n_L(\hat{\mathbf{n}})$ is the surface number density of lens halos at a given location on the sky, $\hat{\mathbf{n}}$. Expanding it into the perturbation, $n_L(\hat{\mathbf{n}}) = \bar{n}_L[1 + \delta_h(\hat{\mathbf{n}})]$, we obtain

$$\langle \gamma_t^h \rangle(\boldsymbol{\theta}) = \frac{1}{f_{\text{sky}}} \int \frac{d^2\hat{\mathbf{n}}}{4\pi} \langle \delta_h(\hat{\mathbf{n}}) \gamma_t(\hat{\mathbf{n}} + \boldsymbol{\theta}) \rangle, \quad (\text{L.3})$$

where $f_{\text{sky}} \equiv N_L/(4\pi\bar{n}_L)$ is a fraction of sky covered by the observation. From statistical isotropy of the universe, $\langle \delta_h(\hat{\mathbf{n}}) \gamma_t(\hat{\mathbf{n}} + \boldsymbol{\theta}) \rangle$ does not depend on $\hat{\mathbf{n}}$, and thus the integral over $\hat{\mathbf{n}}$ simply gives $4\pi f_{\text{sky}}$. Expanding δ_h and γ_t in Fourier space, we obtain

$$\begin{aligned} & \langle \gamma_t^h \rangle(\boldsymbol{\theta}) \\ &= - \int \frac{d^2\mathbf{l}}{(2\pi)^2} \frac{d^2\mathbf{l}'}{(2\pi)^2} e^{i\mathbf{l} \cdot \hat{\mathbf{n}}} e^{i\mathbf{l}' \cdot (\hat{\mathbf{n}} + \boldsymbol{\theta})} \cos[2(\phi - \varphi)] \langle \delta_h(\mathbf{l}) \kappa(\mathbf{l}') \rangle \\ &= - \int \frac{d^2\mathbf{l}}{(2\pi)^2} C_l^{h\kappa} \cos[2(\phi - \varphi)] e^{-i\mathbf{l} \cdot \boldsymbol{\theta}}, \end{aligned} \quad (\text{L.4})$$

where we have used $\langle \delta_h(1)\kappa(\mathbf{l}') \rangle = (2\pi)^2 C_l^{h\kappa} \delta_D(1+\mathbf{l}')$. Finally, we take the azimuthal average of $\langle \gamma_t^h \rangle(\boldsymbol{\theta})$ to find the averaged mean tangential shear:

$$\begin{aligned}
\langle \bar{\gamma}_t^h \rangle(\boldsymbol{\theta}) &= \int_0^{2\pi} \frac{d\phi}{2\pi} \langle \gamma_t^h \rangle(\boldsymbol{\theta}) \\
&= - \int \frac{d^2\mathbf{l}}{(2\pi)^2} C_l^{h\kappa} \int_0^{2\pi} \frac{d\phi}{2\pi} \cos[2(\phi - \varphi)] e^{-il\theta \cos(\phi - \varphi)} \\
&= \int \frac{d^2\mathbf{l}}{(2\pi)^2} C_l^{h\kappa} J_2(l\theta) \\
&= \int \frac{ldl}{2\pi} C_l^{h\kappa} J_2(l\theta). \tag{L.5}
\end{aligned}$$

This completes the derivation of Eq. (6.9).

L.2 Derivation of the covariance matrix of the mean tangential shear

To compute the covariance matrix of the tangential shears (not yet azimuthally averaged), we first compute

$$\begin{aligned}
&\langle \gamma_t^h(\boldsymbol{\theta}) \gamma_t^h(\boldsymbol{\theta}') \rangle \\
&= \frac{1}{N_L^2} \sum_{ij} \int d^2\hat{\mathbf{n}} \int d^2\hat{\mathbf{n}}' \\
&\quad \times \langle \delta_D(\hat{\mathbf{n}} - \hat{\mathbf{n}}_i) \delta_D(\hat{\mathbf{n}}' - \hat{\mathbf{n}}_j) \gamma_t(\hat{\mathbf{n}} + \boldsymbol{\theta}) \gamma_t(\hat{\mathbf{n}}' + \boldsymbol{\theta}') \rangle \\
&= \frac{1}{N_L^2} \int d^2\hat{\mathbf{n}} \int d^2\hat{\mathbf{n}}' \\
&\quad \times [\delta_D(\hat{\mathbf{n}} - \hat{\mathbf{n}}') \langle n_L(\hat{\mathbf{n}}) \gamma_t(\hat{\mathbf{n}} + \boldsymbol{\theta}) \gamma_t(\hat{\mathbf{n}}' + \boldsymbol{\theta}') \rangle \\
&\quad + \langle n_L(\hat{\mathbf{n}}) n_L(\hat{\mathbf{n}}') \gamma_t(\hat{\mathbf{n}} + \boldsymbol{\theta}) \gamma_t(\hat{\mathbf{n}}' + \boldsymbol{\theta}') \rangle]. \tag{L.6}
\end{aligned}$$

Here, the first term in the square bracket correlates two γ_t 's measured relative to the same lens halo (1-halo term), and the second correlates two γ_t 's relative to two lens halos (2-halo term). Again expanding n_L into the perturbation, $n_L(\hat{\mathbf{n}}) = \bar{n}_L[1 + \delta_h(\hat{\mathbf{n}})]$, we obtain

$$\begin{aligned}
&\langle \gamma_t(\boldsymbol{\theta}) \gamma_t(\boldsymbol{\theta}') \rangle \\
&= \frac{1}{f_{\text{sky}}} \frac{1}{N_L} \int \frac{d^2\hat{\mathbf{n}}}{4\pi} \langle \gamma_t(\hat{\mathbf{n}} + \boldsymbol{\theta}) \gamma_t(\hat{\mathbf{n}} + \boldsymbol{\theta}') \rangle \\
&\quad + \frac{1}{f_{\text{sky}}^2} \int \frac{d^2\hat{\mathbf{n}}}{4\pi} \int \frac{d^2\hat{\mathbf{n}}'}{4\pi} [\langle \gamma_t(\hat{\mathbf{n}} + \boldsymbol{\theta}) \gamma_t(\hat{\mathbf{n}}' + \boldsymbol{\theta}') \rangle \\
&\quad + \langle \delta_h(\hat{\mathbf{n}}) \delta_h(\hat{\mathbf{n}}') \gamma_t(\hat{\mathbf{n}} + \boldsymbol{\theta}) \gamma_t(\hat{\mathbf{n}}' + \boldsymbol{\theta}') \rangle]. \tag{L.7}
\end{aligned}$$

Here, we assume that δ_h and γ_t obey Gaussian statistics, i.e., $\langle \delta_h \gamma_t \gamma_t \rangle = 0$. This approximation is justified even in the presence of primordial non-Gaussianity, as non-Gaussianity is weak, and this approximation only affects the size of errorbars. Let us evaluate each term. With γ_t expanded in Fourier space, the first term (1-halo term) becomes

$$\begin{aligned} & \frac{1}{N_L} \frac{1}{f_{\text{sky}}} \int \frac{d^2 \hat{\mathbf{n}}}{4\pi} \langle \gamma_t(\hat{\mathbf{n}} + \boldsymbol{\theta}) \gamma_t(\hat{\mathbf{n}} + \boldsymbol{\theta}') \rangle \\ &= \frac{1}{N_L} \int \frac{d^2 \mathbf{l}}{(2\pi)^2} C_l^\kappa \cos[2(\phi - \varphi)] \cos[2(\phi' - \varphi)] e^{i\mathbf{l} \cdot (\boldsymbol{\theta} - \boldsymbol{\theta}')} \\ & \quad + \frac{\sigma_\gamma^2}{N_L n_S} \delta_D(\boldsymbol{\theta} - \boldsymbol{\theta}'), \end{aligned} \quad (\text{L.8})$$

where σ_γ is the r.m.s. shape noise (dimensionless), and n_S is the surface density of source (background) galaxies that are available for the shear measurement at a given location. By azimuthally averaging γ_t , we find

$$\begin{aligned} & \frac{1}{N_L} \int_0^{2\pi} \frac{d\phi}{2\pi} \int_0^{2\pi} \frac{d\phi'}{2\pi} \langle \gamma_t(\hat{\mathbf{n}} + \boldsymbol{\theta}) \gamma_t(\hat{\mathbf{n}} + \boldsymbol{\theta}') \rangle \\ &= \frac{1}{N_L} \int \frac{d^2 \mathbf{l}}{(2\pi)^2} C_l^\kappa J_2(l\theta) J_2(l\theta') + \frac{\sigma_\gamma^2}{N_L n_S} \frac{\delta_D(\theta - \theta')}{2\pi\theta}. \end{aligned} \quad (\text{L.9})$$

Here, C_l^κ is the angular power spectrum of $\kappa(\mathbf{l})$.

As for the second term (2-halo term), the first of the second term vanishes, as $\int d^2 \hat{\mathbf{n}} \gamma_t(\hat{\mathbf{n}} + \boldsymbol{\theta}) = 0$. The remaining non-vanishing term gives

$$\begin{aligned} & \frac{1}{f_{\text{sky}}^2} \int \frac{d^2 \hat{\mathbf{n}}}{4\pi} \int \frac{d^2 \hat{\mathbf{n}}'}{4\pi} \\ & \times [\langle \delta_h(\hat{\mathbf{n}}) \gamma_t(\hat{\mathbf{n}} + \boldsymbol{\theta}) \rangle \langle \delta_h(\hat{\mathbf{n}}') \gamma_t(\hat{\mathbf{n}}' + \boldsymbol{\theta}') \rangle \\ & \quad + \langle \delta_h(\hat{\mathbf{n}}) \gamma_t(\hat{\mathbf{n}}' + \boldsymbol{\theta}') \rangle \langle \delta_h(\hat{\mathbf{n}}') \gamma_t(\hat{\mathbf{n}} + \boldsymbol{\theta}) \rangle \\ & \quad + \langle \delta_h(\hat{\mathbf{n}}) \delta_h(\hat{\mathbf{n}}') \rangle \langle \gamma_t(\hat{\mathbf{n}} + \boldsymbol{\theta}) \gamma_t(\hat{\mathbf{n}}' + \boldsymbol{\theta}') \rangle] \\ &= \langle \gamma_t^h(\boldsymbol{\theta}) \rangle \langle \gamma_t^h(\boldsymbol{\theta}') \rangle \\ & \quad + \frac{1}{4\pi f_{\text{sky}}} \int \frac{d^2 \mathbf{l}}{(2\pi)^2} \cos[2(\phi - \varphi)] \cos[2(\phi' - \varphi)] e^{i\mathbf{l} \cdot (\boldsymbol{\theta} - \boldsymbol{\theta}')} \\ & \quad \times \left[(C_l^{h\kappa})^2 + C_l^h \left(C_l^\kappa + \frac{\sigma_\gamma^2}{n_S} \right) \right]. \end{aligned} \quad (\text{L.10})$$

Here, C_l^h is the angular power spectrum of $\delta_h(\mathbf{l})$. By azimuthally averaging γ_t in the above

equation, we find

$$\begin{aligned} & \langle \bar{\gamma}_t^h(\theta) \rangle \langle \bar{\gamma}_t^h(\theta') \rangle \\ & + \frac{1}{4\pi f_{\text{sky}}} \int \frac{d^2\mathbf{l}}{(2\pi)^2} J_2(l\theta) J_2(l\theta') \left[(C_l^{h\kappa})^2 + C_l^h \left(C_l^\kappa + \frac{\sigma_\gamma^2}{n_S} \right) \right], \end{aligned}$$

where we have used the identity

$$\frac{\delta_D(\theta - \theta')}{2\pi\theta} = \int \frac{ldl}{2\pi} J_2(l\theta) J_2(l\theta'). \quad (\text{L.11})$$

Collecting both the 1-halo and 2-halo terms, we finally obtain the covariance matrix of the azimuthally-averaged mean tangential shear:

$$\begin{aligned} & \langle \bar{\gamma}_t^h(\theta) \bar{\gamma}_t^h(\theta') \rangle - \langle \bar{\gamma}_t^h(\theta) \rangle \langle \bar{\gamma}_t^h(\theta') \rangle \\ & = \frac{1}{4\pi f_{\text{sky}}} \int \frac{ldl}{2\pi} J_2(l\theta) J_2(l\theta') \left[(C_l^{h\kappa})^2 + \left(C_l^h + \frac{1}{n_L} \right) \left(C_l^\kappa + \frac{\sigma_\gamma^2}{n_S} \right) \right]. \end{aligned}$$

This completes the derivation of Eq. (6.16).

Appendix M

On the accuracy of Limber's approximation

Throughout this paper we have repeatedly used Limber's approximation in order to relate the angular correlation function to the corresponding three dimensional power spectrum. In general, Limber's approximation is known to be accurate only for small angular scales, and only for the quantities which are integrated over a broad range of redshift.

However, the situations we have considered in this paper sometimes violate both of the conditions above: 1) We correlate the convergence field with galaxies within a very thin redshift slice, and 2) the non-Gaussianity signal we study in this paper appears only on very large scales.

Then, how accurate is Limber's approximation in this case? In this Appendix, we shall study in detail the validity and limitation of Limber's approximation, by comparing the main results of the paper to the result of *exact* calculations.

Consider a quantity $x_i(\hat{\mathbf{n}})$, which is projected on the sky. Here, $\hat{\mathbf{n}}$ is the unit vector pointing toward a given direction on the sky. This quantity is related to the three dimensional quantity $s_i(\mathbf{r}; z)$ by a projection kernel $W_i(z)$ as

$$x_i(\hat{\mathbf{n}}) = \int dz W_i(z) s_i[d_A(z)\hat{\mathbf{n}}; z]. \quad (\text{M.1})$$

Throughout this Appendix, we use $d_A(z)$ to denote $d_A(0; z)$.

Fourier transforming $s_i(\mathbf{r})$, one obtains

$$\begin{aligned} & s_i[d_A(z)\hat{\mathbf{n}}; z] \\ &= \int \frac{d^3k}{(2\pi)^3} s_i(\mathbf{k}, z) e^{i\mathbf{k} \cdot \hat{\mathbf{n}} d_A(z)} \\ &= 4\pi \sum_{l,m} i^l \int \frac{d^3k}{(2\pi)^3} s_i(\mathbf{k}, z) j_l[kd_A(z)] Y_{lm}^*(\hat{\mathbf{k}}) Y_{lm}(\hat{\mathbf{n}}). \end{aligned} \quad (\text{M.2})$$

In the third line, we have used Rayleigh's formula:

$$e^{i\mathbf{k} \cdot \hat{\mathbf{n}} r} = 4\pi \sum_{l,m} i^l j_l(kr) Y_{lm}^*(\hat{\mathbf{k}}) Y_{lm}(\hat{\mathbf{n}}).$$

By using Eq. (M.2), we rewrite Eq. (M.1) as

$$x_i(\hat{\mathbf{n}}) = 4\pi \sum_{l,m} i^l \int dz W_i(z) \times \int \frac{d^3 k}{(2\pi)^3} s_i(\mathbf{k}, z) j_l[kd_A(z)] Y_{lm}^*(\hat{\mathbf{k}}) Y_{lm}(\hat{\mathbf{n}}). \quad (\text{M.3})$$

Therefore, the coefficients of the spherical harmonics decomposition of $x_i(\hat{\mathbf{n}})$, $a_{lm}^{x_i}$, becomes

$$a_{lm}^{x_i} = 4\pi i^l \int dz W_i(z) \int \frac{d^3 k}{(2\pi)^3} s_i(\mathbf{k}, z) j_l[kd_A(z)] Y_{lm}^*(\hat{\mathbf{k}}). \quad (\text{M.4})$$

We calculate the angular power spectrum, $C_l^{x_i x_j}$, by taking an ensemble average of $\langle a_{lm}^{x_i} a_{lm}^{x_j*} \rangle$ as

$$\begin{aligned} C_l^{x_i x_j} &\equiv \langle a_{lm}^{x_i} a_{lm}^{x_j*} \rangle \\ &= (4\pi)^2 \int dz W_i(z) \int dz' W_j(z') \int \frac{d^3 k}{(2\pi)^3} P^{s_i s_j}(\mathbf{k}; z, z') \\ &\quad \times j_l[kd_A(z)] j_l[kd_A(z')] Y_{lm}^*(\hat{\mathbf{k}}) Y_{lm}(\hat{\mathbf{k}}), \end{aligned} \quad (\text{M.5})$$

where we have used the definition of the power spectrum:

$$\langle s_i(\mathbf{k}, z) s_j^*(\mathbf{k}', z) \rangle = (2\pi)^3 \delta(\mathbf{k} - \mathbf{k}') P^{s_i s_j}(\mathbf{k}; z, z').$$

Now, by assuming statistical isotropy of the universe, we write $P^{s_i s_j}(\mathbf{k}; z, z') = P^{s_i s_j}(k; z, z')$, and do the angular integration of $\hat{\mathbf{k}}$ by using the orthonormality condition of spherical harmonics:

$$\int d\hat{\mathbf{k}} Y_{lm}(\hat{\mathbf{k}}) Y_{lm}^*(\hat{\mathbf{k}}) = 1.$$

We then obtain the angular power spectrum given by

$$\begin{aligned} C_l^{x_i x_j} &= \int dz W_i(z) \int dz' W_j(z') \\ &\quad \times \left\{ \frac{2}{\pi} \int k^2 dk P^{s_i s_j}(k; z, z') j_l[kd_A(z)] j_l[kd_A(z')] \right\}. \end{aligned} \quad (\text{M.6})$$

This is the *exact* relation.

What determines the form of $W_i(z)$? For a projected galaxy distribution projected on the sky, this kernel is simply a normalized galaxy distribution function in redshift space. In this paper, we consider the delta function-like distribution, i.e.,

$$W_g(z) = \delta^D(z - z_L). \quad (\text{M.7})$$

Using Eq. (M.6) with the delta function kernel above yields Eq. (6.34):

$$C_l^h = \frac{2}{\pi} \int dk k^2 P_g(k, z_L) j_l^2[kd_A(z_L)]. \quad (\text{M.8})$$

Again, this is still the exact result. As the form of $W_g(z)$ we have considered here (i.e., a delta function) is a sharply peaked function, we cannot use Limber's approximation given below. This is the reason why we have used the exact result for C_l^h .

In order to get the expression for Limber's approximation, we assume that $P^{s_i s_j}(k)$ is a slowly-varying function of k . Then, by using the identity

$$\frac{2}{\pi} \int k^2 dk j_l(kr) j_l(kr') = \frac{\delta^D(r - r')}{r^2}, \quad (\text{M.9})$$

we approximate the k integral of Eq. (M.6) as

$$\begin{aligned} & \frac{2}{\pi} \int k^2 dk P^{s_i s_j}(k) j_l(kr) j_l(kr') \\ & \approx \frac{\delta^D(r - r')}{r^2} P^{s_i s_j} \left(k = \frac{l + 1/2}{r} \right). \end{aligned} \quad (\text{M.10})$$

By using this approximation, we finally get

$$C_l^{x_i x_j} \approx \int dz W_i(z) W_j(z) \frac{H(z)}{d_A^2(z)} P^{s_i s_j} \left(k = \frac{l + 1/2}{r}; z \right), \quad (\text{M.11})$$

which is the result known as Limber's approximation.

One important application of Limber's approximation is the statistics involving weak gravitational lensing. The lensing kernel for the convergence field, $W_\kappa(z)$, can be calculated by integrating the lens equation:

$$W_\kappa(z) = \frac{\rho_0}{\Sigma_c(z; z_S) H(z)}, \quad (\text{M.12})$$

where $\Sigma_c(z; z_S)$ is the critical surface density defined in Eq. (6.14). The exact result for the galaxy-convergence angular cross power spectrum is

$$\begin{aligned} C_l^{h\kappa}(z_L) &= \frac{2}{\pi} \int_0^{z_S} dz \frac{\rho_0}{\Sigma_c(z; z_S) H(z)} \\ &\times \int dk k^2 P_{hm}(k, z_L, z) j_l[kd_A(z_L)] j_l[kd_A(z)], \end{aligned} \quad (\text{M.13})$$

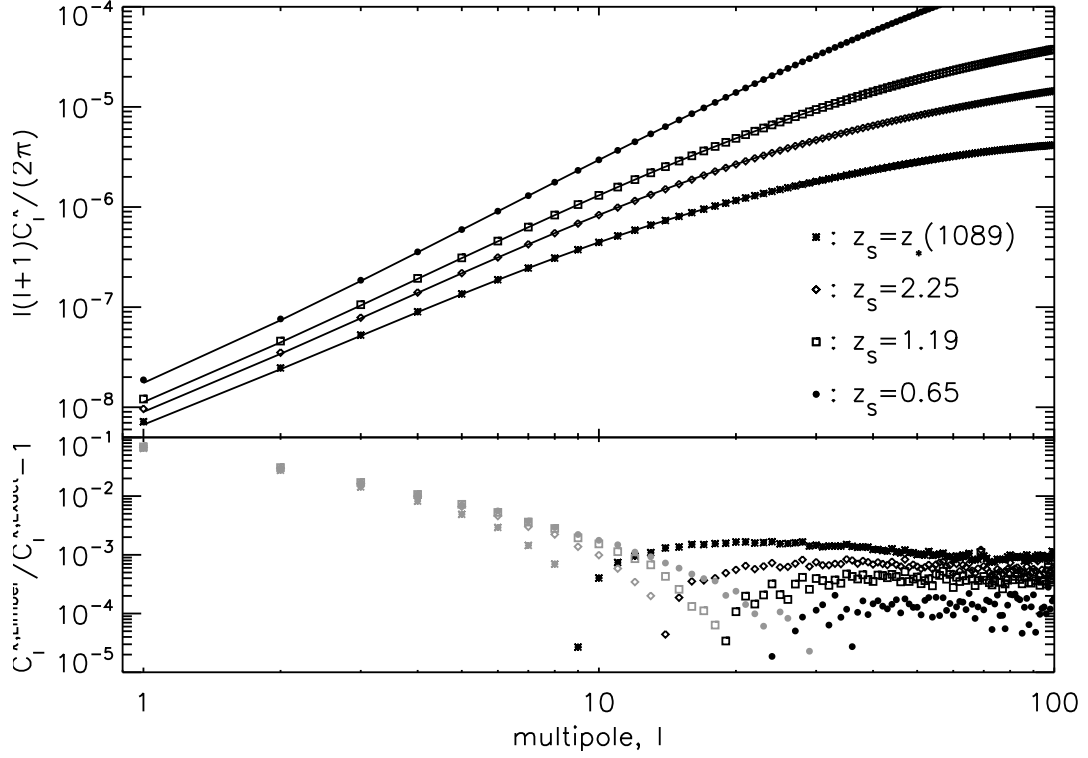


Figure M.1: Top: Convergence-convergence angular power spectrum from two different methods: the exact calculation (Eq. M.14, symbols) and Limber's approximation (Eq. 6.33, solid lines). Bottom: Fractional differences between Limber's approximation and the exact integration. Symbols are the same as the top panel. Grey symbols show the absolute values of negative values.

and the exact result for the convergence-convergence angular power spectrum is

$$\begin{aligned}
 C_l^\kappa(z_S) &= \frac{2}{\pi} \int_0^{z_S} dz \int_0^{z_S} dz' \frac{\rho_0^2}{\Sigma_c(z; z_S) H(z) \Sigma_c(z'; z_S) H(z')} \\
 &\times \int dk k^2 P_m(k, z, z') j_l[kd_A(z)] j_l[kd_A(z')].
 \end{aligned} \tag{M.14}$$

First, we compare the exact convergence-convergence angular power spectrum to Limber's approximation. Fig. M.1 shows that Limber's approximation works very well for all four source redshifts we study in the paper: $z_S = 0.65, 1.19, 2.25$, and 1089.0 . For $l > 10$, the error caused by Limber's approximation is always much smaller than 1%.

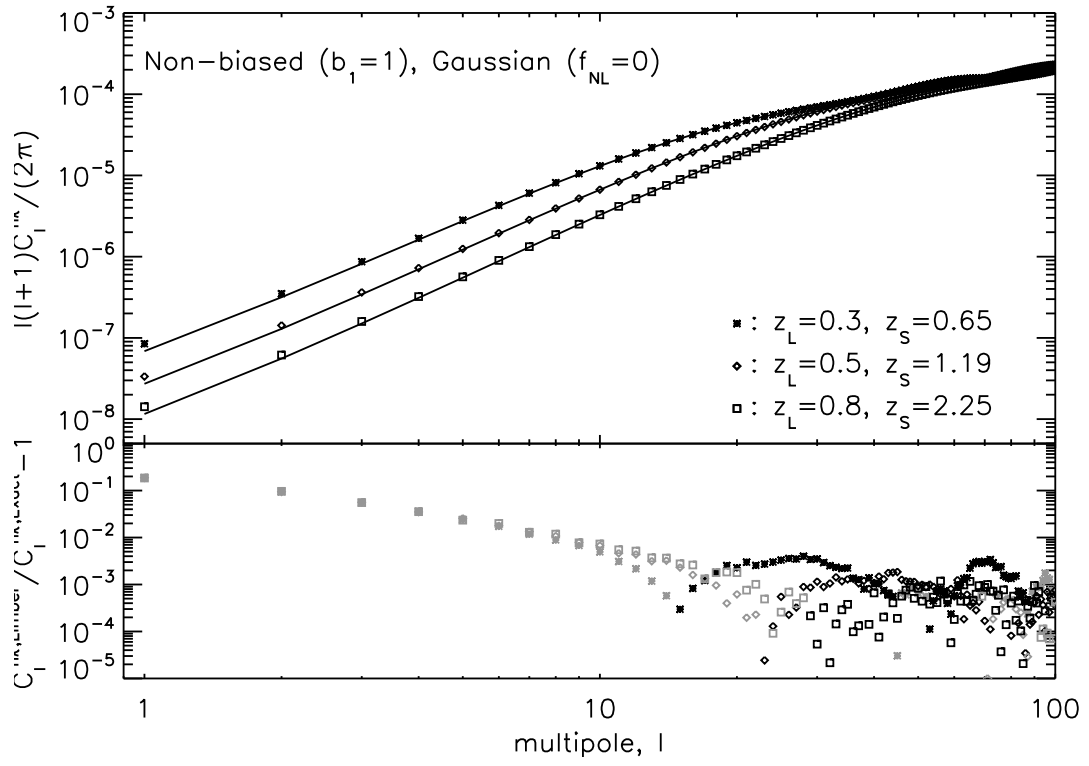


Figure M.2: Same as Fig. M.1, but for the galaxy-convergence cross angular power spectrum with $f_{\text{NL}} = 0$ and $b_1 = 1$.

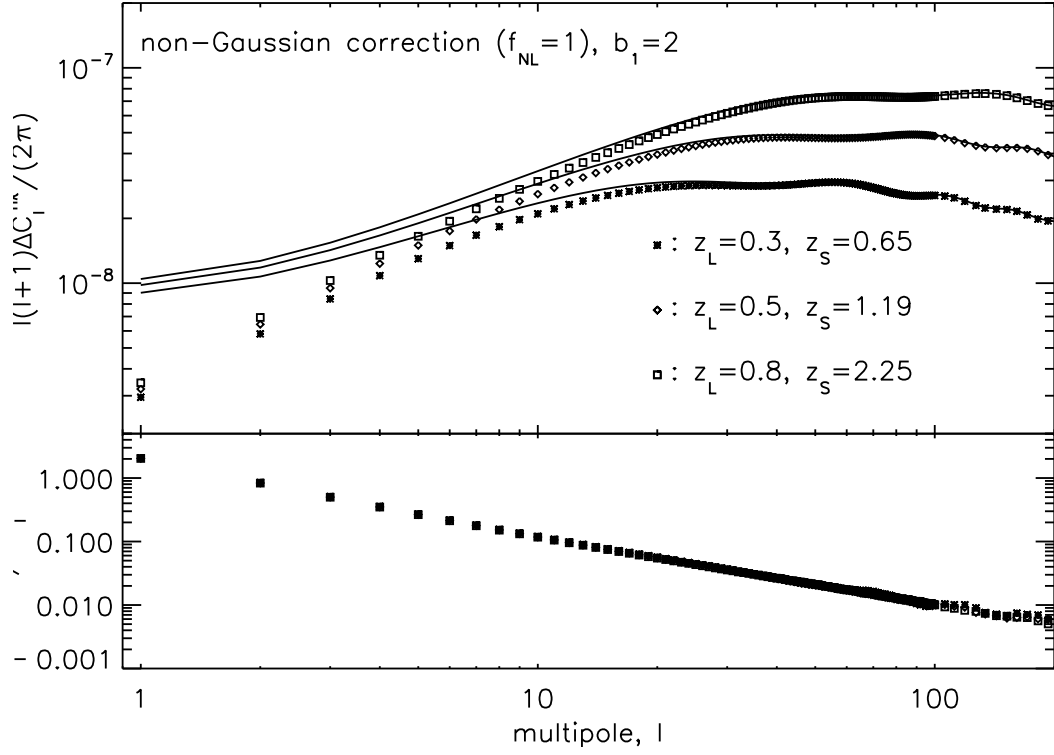


Figure M.3: Same as Fig. M.1, but for the non-Gaussian correction (i.e., the term proportional to $\Delta b(k)$) to the galaxy-convergence cross angular power spectrum. We show the corrections with $f_{\text{NL}} = 1$ and $b_1 = 2$.

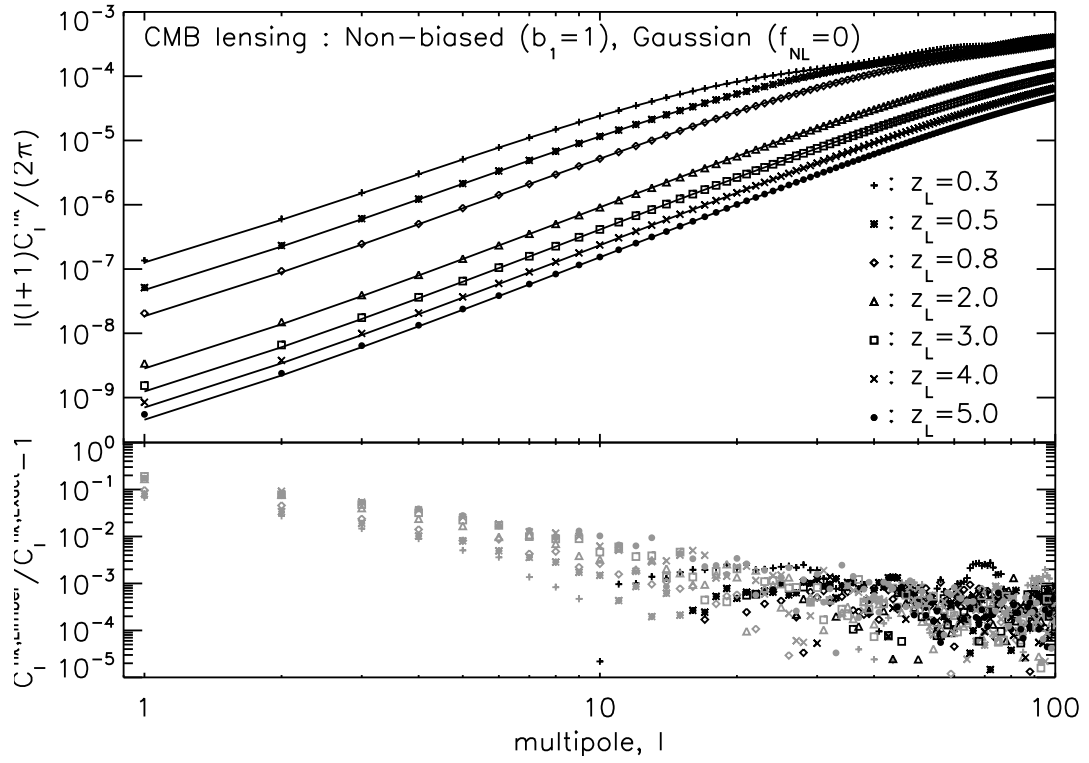


Figure M.4: Same as Fig. M.2, but for the galaxy-CMB lensing.

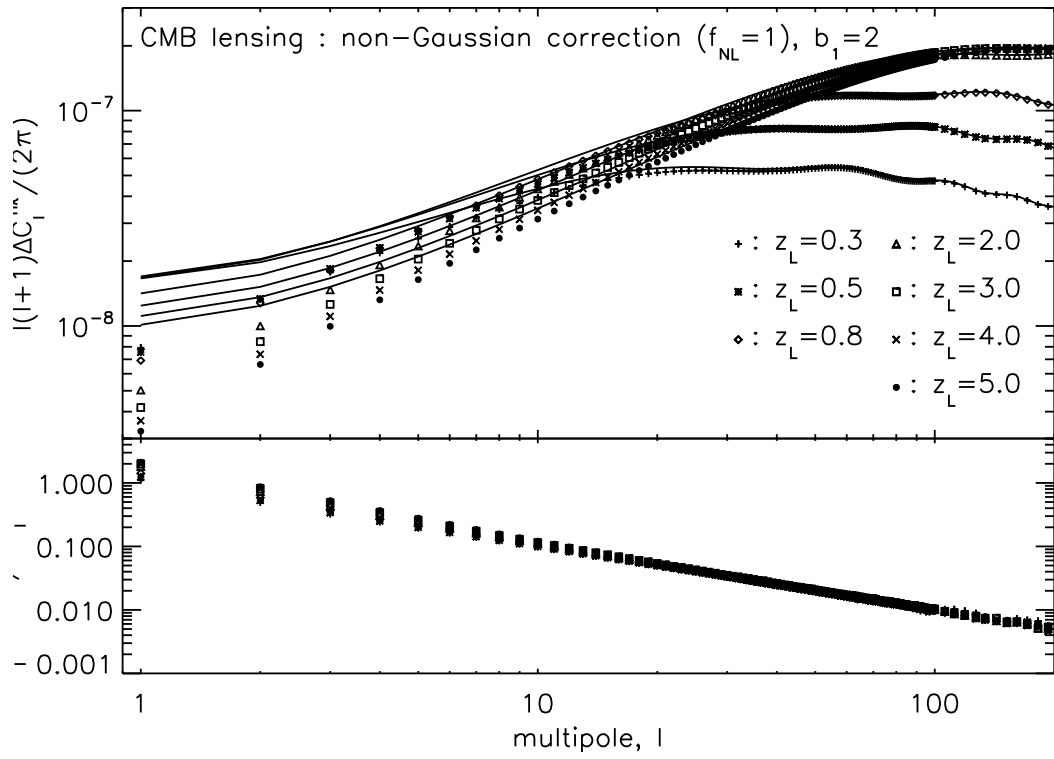


Figure M.5: Same as Fig. M.3, but for the galaxy-CMB lensing.

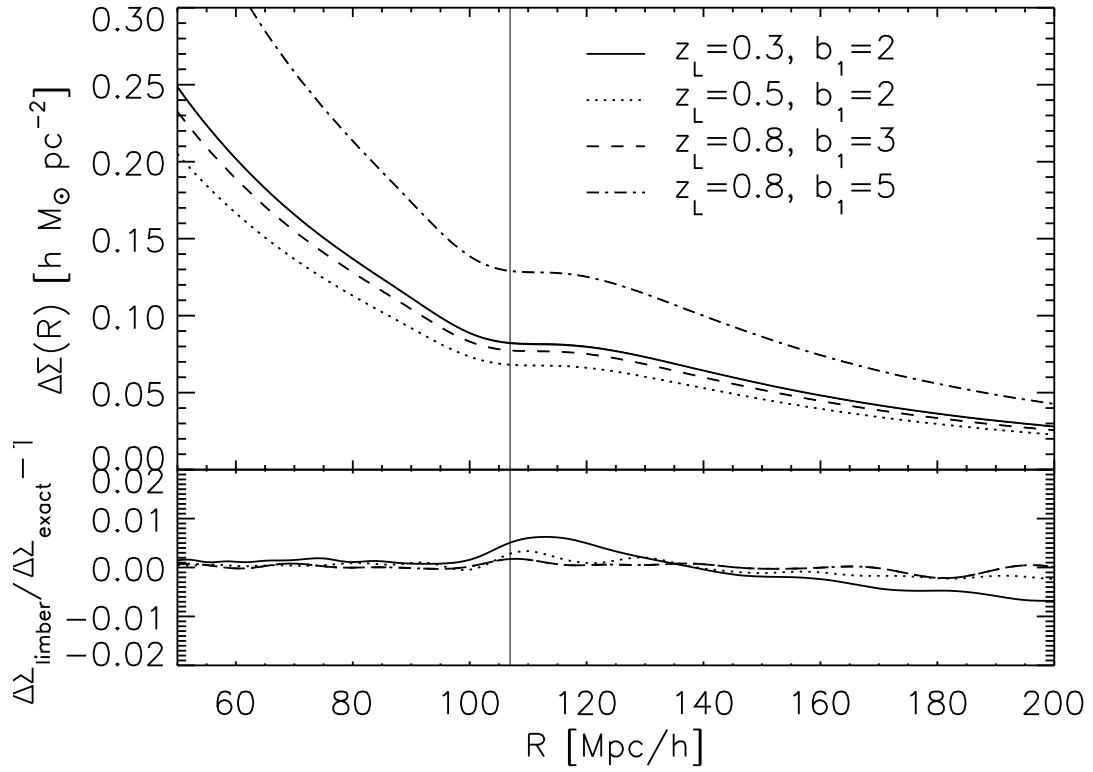


Figure M.6: Top: Same as Fig. 6.3, but also showing the exact result (Eq. M.13, thick lines) on top of the result from Limber's approximation (Eq. 6.30, thin lines). Bottom: Fractional difference of Limber's approximation relative to the exact result.

Then, we compare the galaxy-convergence cross angular power spectra. Fig. M.2 and Fig. M.3 show the comparison between the exact galaxy-convergence cross power spectrum (Eq. M.13, symbols) and their Limber approximation (Eq. 6.30, solid lines) for three galaxy-galaxy lensing cases we study in Sec. 6.1: $(z_L, z_S) = (0.3, 0.65)$, $(0.5, 1.19)$, and $(0.8, 2.25)$.

For the Gaussian term (Fig. M.2), Limber's approximation is accurate at $l > 10$ with the errors less than 1%. On the other hand, Limber's approximation to the non-Gaussian correction term (Fig. M.3) has a sizable error, at the level of 10%, at $l \sim 10$. The error goes down to the 1% level only at $l \sim 100$. One needs to keep this in mind when comparing Limber's approximation with observations. We find that Limber's approximation underpredicts the Gaussian term at $l \lesssim 20$, while it overpredicts the non-Gaussian corrections at all multipoles.

The story is basically the same for the galaxy-CMB lensing cross power spectrum. Fig. M.4 (Gaussian term) and Fig. M.5 (non-Gaussian correction) show the comparison between the exact galaxy-convergence cross power spectrum (Eq. M.13, solid lines) and their Limber approximation (Eq. 6.30, dashed lines) for seven lens redshifts we study in Sec. 6.3: $z_L = 0.3, 0.5, 0.8, 2, 3, 4$, and 5 . Again, for small scales, $l > 10$, Limber's approximation works better than 1% for the Gaussian term, while it overpredicts the non-Gaussian correction at the level of 10% at $l \sim 10$ and 1% at $l \sim 100$.

What about the effect on the mean tangential shear, $\Delta\Sigma(R)$? Fig. M.6 compares the Gaussian term of $\Delta\Sigma(R)$ from the exact integration and that from Limber's approximation. On the top panel of Fig. M.6, we show the baryonic feature computed with Limber's approximation (thin lines, the same as those in Fig. 6.3) as well as that computed with the exact integration (thick lines). They are indistinguishable by eyes. The bottom panel shows the fractional differences between the two. We find that Limber's approximation is better than 0.5% for $R < 180 h^{-1}$ Mpc; thus, the baryonic feature in $\Delta\Sigma$ is not an artifact caused by Limber's approximation.

However, Limber's approximation becomes worse and worse as we go to larger R . Fig. M.8 shows $\Delta\Sigma$ on large scales. For the lens redshifts that we have studied here, the error is at most 5% for $R < 500 h^{-1}$ Mpc, and the error is the largest for the lowest z_L , as a given R at a lower redshift corresponds to a larger angular separation on the sky.

While Limber's approximation underpredicts the Gaussian term on large scales, it overpredicts the non-Gaussian correction terms. Fig. M.8 shows the fractional differences of

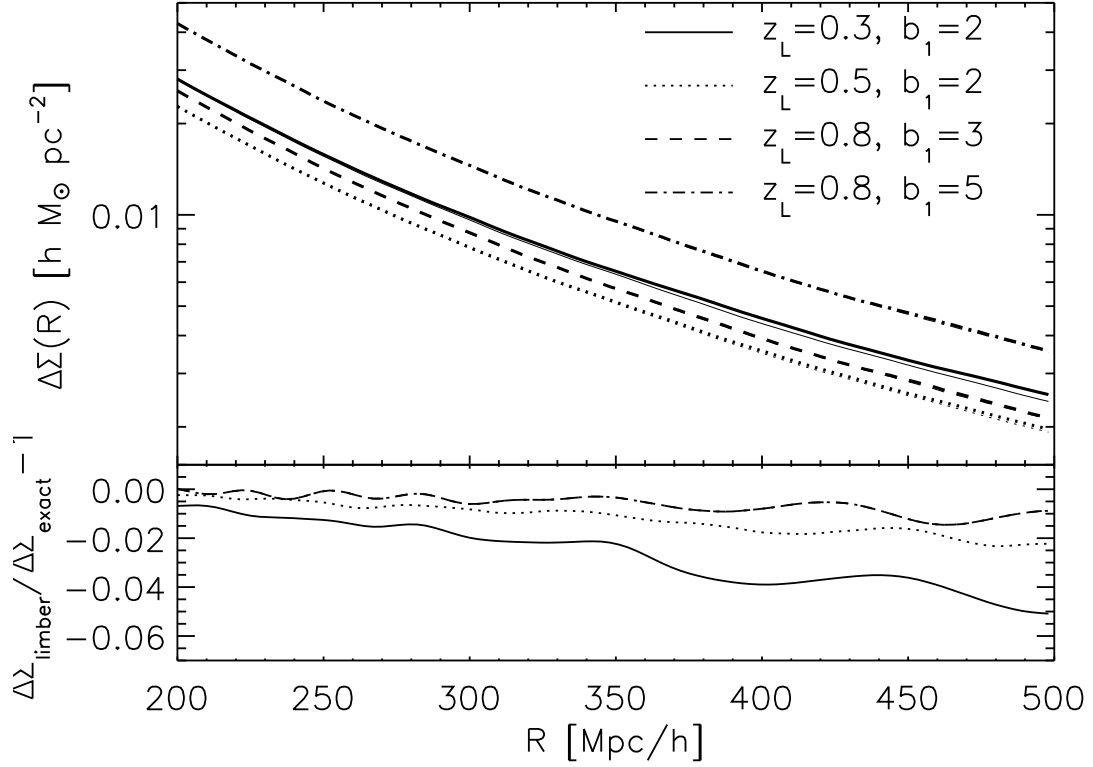


Figure M.7: Same as Fig. M.6, but for larger R . Thick lines are the results of the exact integration, while the thin lines are Limber's approximation. The Limber approximation overpredicts $\Delta\Sigma(R)$ for large R , but the error is at most 5% for $R < 500 h^{-1}$ Mpc. The error is the largest for the lowest z_L , as a physical separation R at a lower redshift corresponds to a larger angular separation on the sky.

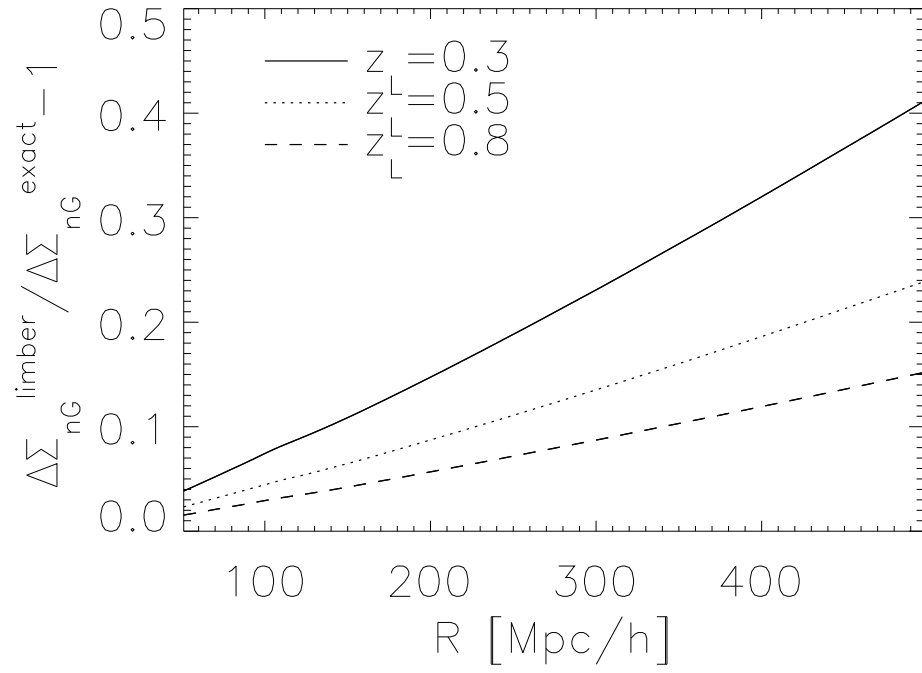


Figure M.8: Fractional differences in the non-Gaussian correction terms, $\Delta\Sigma_{nG}$, from Limber's approximation and the exact integration. Using Limber's approximation, we overpredict the non-Gaussian correction by $\sim 20\%$ at $R = 300 h^{-1}$ Mpc for $z_L = 0.3$.

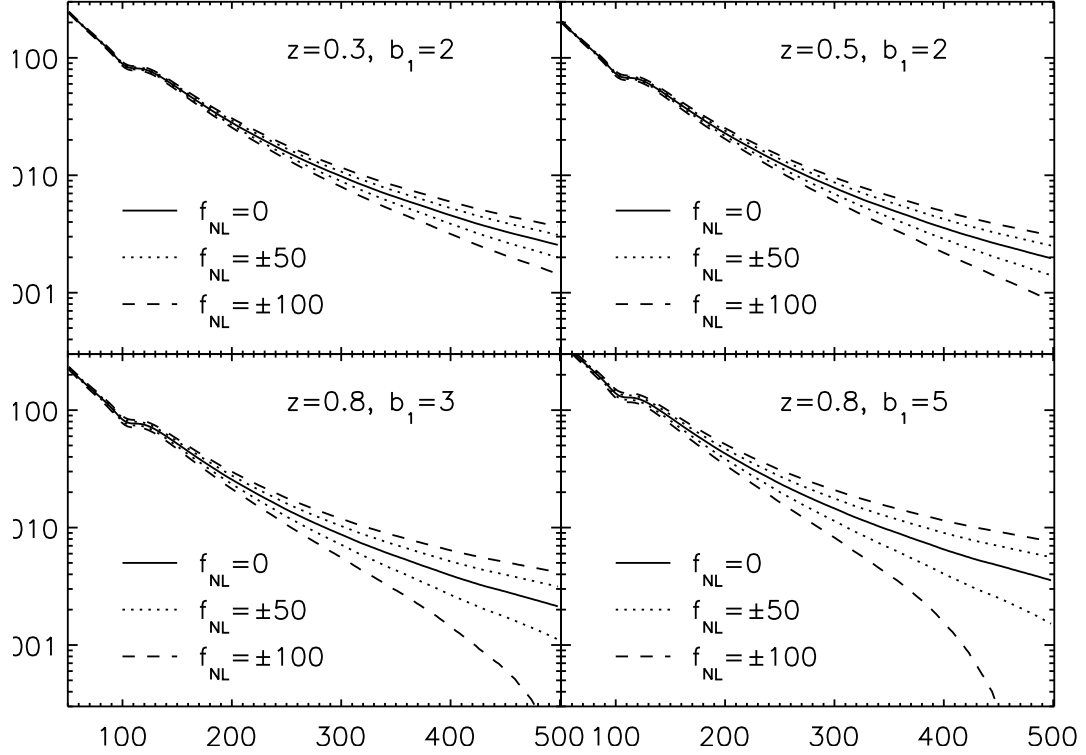


Figure M.9: Same as Fig. 6.4, but with the exact integration instead of Limber's approximation.

the non-Gaussian correction terms, $\Delta\Sigma_{\text{nG}}$, between Limber's approximation and the exact calculation as a function of separation R for three lens redshifts: $z_L = 0.3, 0.5$, and 0.8 . This figure shows that the error caused by Limber's approximation can be substantial on $\Delta\Sigma_{\text{nG}}$.

As Limber's approximation to $\Delta\Sigma(R)$ can be quite inaccurate on very large scales, we show the exact calculations of $\Delta\Sigma(R)$ in Fig. M.9. (Limber's approximation is given in Fig. 6.4.)

Finally, we note that the definition of the tangential shear we have used (Eq. 6.2) is valid only on the flat sky (as noted in the footnote there), and thus the prediction for $\Delta\Sigma$ on very large scales probably needs to be revisited with the exact definition of the tangential shears on the full sky using the spin-2 harmonics. This is beyond of the scope of

this dissertation.

Appendix N

Cosmic Microwave Background lensing reconstruction: Quadratic Estimator with the flat-sky approximation

In this appendix, we summarize the of Cosmic Microwave Background (CMB) lensing reconstruction method in Hu & Okamoto (2002). The method summarized here is valid for small angular scales as we use the flat-sky approximation. For general Full-sky treatment with spin-2 spherical harmonics, see Okamoto & Hu (2003).

The intervening matter distribution of the large scale structure of the Universe lenses CMB photon along the direction $\hat{\mathbf{n}}$ with lensing potential $\phi(\hat{\mathbf{n}})$ given by

$$\phi(\hat{\mathbf{n}}) = 2 \int d\eta \frac{d_A(\eta_*)}{d_A(\eta_*)d_A(\eta)} \Phi(d_A(\eta)\hat{\mathbf{n}}, \eta), \quad (\text{N.1})$$

where $d_A(\eta)$ is the comoving angular diameter distance corresponding to the comoving distance η , and η_* is the comoving distance to the last scattering surface. Here Φ is the gravitational potential. As a result, temperature ($\Theta(\hat{\mathbf{n}}) \equiv \Delta T(\hat{\mathbf{n}})/T$) anisotropy and Stokes parameters $Q(\hat{\mathbf{n}})$ and $U(\hat{\mathbf{n}})$ of CMB is re-mapped as

$$\Theta(\hat{\mathbf{n}}) = \tilde{\Theta}(\hat{\mathbf{n}} + \nabla\phi(\hat{\mathbf{n}})) \quad (\text{N.2})$$

$$(Q \pm iU)(\hat{\mathbf{n}}) = (\tilde{Q} \pm i\tilde{U})(\hat{\mathbf{n}} + \nabla\phi(\hat{\mathbf{n}})). \quad (\text{N.3})$$

Following the notation of Hu & Okamoto (2002), we denote the unlensed power spectra of $x = \Theta$ (temperature fluctuation), E -, and B -mode polarization as

$$\langle \tilde{x}^*(\mathbf{l}) \tilde{x}(\mathbf{l}') \rangle \equiv (2\pi)^2 \delta(\mathbf{l} - \mathbf{l}') \tilde{C}_l^{xx'} \quad (\text{N.4})$$

$$\langle \phi^*(\mathbf{L}) \phi(\mathbf{L}') \rangle \equiv (2\pi)^2 \delta(\mathbf{L} - \mathbf{L}') L^{-2} C_L^{dd}. \quad (\text{N.5})$$

Note that, as we assume that parity is conserved,

$$\tilde{C}_l^{\Theta B} = \tilde{C}_l^{EB} = 0.$$

On the other hands, observed power spectra are denoted without tilde as

$$\langle x^*(\mathbf{l}) x(\mathbf{l}') \rangle \equiv (2\pi)^2 \delta(\mathbf{l} - \mathbf{l}') C_l^{xx'}. \quad (\text{N.6})$$

These ‘observed’ power spectra include all the sources of uncertainties such as instrumental white noise and foreground contaminations. In order to quantify the uncertainty of the observed power spectra, we use the Gaussian random detector noise of Knox (1995):

$$C_{l,noise}^{\Theta\Theta} = \left(\frac{T_{CMB}}{\Delta_T} \right)^{-2} e^{l(l+1)\sigma^2/8 \ln 2} \quad (\text{N.7})$$

$$C_{l,noise}^{BB} = C_{l,noise}^{EE} = \left(\frac{T_{CMB}}{\Delta_p} \right)^{-2} e^{l(l+1)\sigma^2/8 \ln 2}, \quad (\text{N.8})$$

where Δ_T and Δ_p is in the unit of [$\mu\text{K rad}$]. For *Planck*, we use $\Delta_T = 35.4 \mu\text{K arcmin}$, and $\Delta_p = 63.1 \mu\text{K arcmin}$, and set the FWHM of beam $\sigma = 7 \text{ arcmin}$, which comes from the CMBPol mission study of Zaldarriaga et al. (2008). The ”nearly perfect” experiment referred by Hu & Okamoto (2002) has white noise of $\Delta_T = 1 \mu\text{K arcmin}$, and $\Delta_p = \sqrt{2} \mu\text{K arcmin}$, and FWHM of $\sigma = 4 \text{ arcmin}$.

In order to quantify the effect of weak lensing, we first Taylor expand equation (N.2) and equation (N.2):

$$\Theta(\hat{\mathbf{n}}) = \tilde{\Theta}(\hat{\mathbf{n}}) + (\nabla \tilde{\Theta}(\hat{\mathbf{n}})) \cdot (\nabla \phi(\hat{\mathbf{n}})) + \dots \quad (\text{N.9})$$

$$(Q \pm iU)(\hat{\mathbf{n}}) = (\tilde{Q} \pm i\tilde{U})(\hat{\mathbf{n}}) + (\nabla(\tilde{Q} \pm i\tilde{U})(\hat{\mathbf{n}})) \cdot (\nabla \phi(\hat{\mathbf{n}})) + \dots, \quad (\text{N.10})$$

and Fourier transform them¹. Then, the re-mapping effect of gravitational lensing in Fourier space becomes a mode coupling between $\tilde{\Phi}$, \tilde{E} , \tilde{B} and lensing potential ϕ :

$$\Theta(\mathbf{l}) = \tilde{\Theta}(\mathbf{l}) - \int \frac{d^2 l_1}{(2\pi)^2} \int d^2 l_2 \delta^D(\mathbf{l} - \mathbf{l}_2) \mathbf{l}_1 \cdot \mathbf{l}_2 \phi(\mathbf{l}_2) \tilde{\Theta}(\mathbf{l}_1) \quad (\text{N.13})$$

$$\begin{aligned} E(\mathbf{l}) \pm iB(\mathbf{l}) &= \tilde{E}(\mathbf{l}) \pm i\tilde{B}(\mathbf{l}) - \int \frac{d^2 l_1}{(2\pi)^2} \int d^2 l_2 \delta^D(\mathbf{l} - \mathbf{l}_2) \mathbf{l}_1 \cdot \mathbf{l}_2 \phi(\mathbf{l}_2) \\ &\quad \times (\tilde{E}(\mathbf{l}_1) \pm i\tilde{B}(\mathbf{l}_1)) e^{\pm 2i(\varphi_{l_1} - \varphi_l)}. \end{aligned} \quad (\text{N.14})$$

Decomposing the equations for each Fourier mode of Θ , E , and B , the change due to weak

¹Fourier transform with the flat sky approximation is defined as

$$\Theta(\hat{\mathbf{n}}) = \int \frac{d^2 l}{(2\pi)^2} \Theta(\mathbf{l}) e^{i\mathbf{l} \cdot \mathbf{n}} \quad (\text{N.11})$$

$$(Q \pm iU)(\hat{\mathbf{n}}) = - \int \frac{d^2 l}{(2\pi)^2} [E(\mathbf{l}) \pm iB(\mathbf{l})] e^{2i\varphi_l} e^{i\mathbf{l} \cdot \mathbf{n}}, \quad (\text{N.12})$$

where $\varphi_l \equiv \cos^{-1}(\hat{\mathbf{x}} \cdot \hat{\mathbf{l}})$ is an azimuthal angle of \mathbf{l} .

lensing is (Hu, 2000):

$$\Theta(\mathbf{l}) = \tilde{\Theta}(\mathbf{l}) + \int \frac{d^2 l'}{(2\pi)^2} W(\mathbf{l}, \mathbf{l}') \tilde{\Theta}(\mathbf{l}') \quad (\text{N.15})$$

$$E(\mathbf{l}) = \tilde{E}(\mathbf{l}) + \int \frac{d^2 l'}{(2\pi)^2} W(\mathbf{l}, \mathbf{l}') \left[\tilde{E}(\mathbf{l}') \cos(2\varphi_{l'l}) - \tilde{B}(\mathbf{l}') \sin(2\varphi_{l'l}) \right] \quad (\text{N.16})$$

$$B(\mathbf{l}) = \tilde{B}(\mathbf{l}) + \int \frac{d^2 l'}{(2\pi)^2} W(\mathbf{l}, \mathbf{l}') \left[\tilde{B}(\mathbf{l}') \cos(2\varphi_{l'l}) + \tilde{E}(\mathbf{l}') \sin(2\varphi_{l'l}) \right], \quad (\text{N.17})$$

where

$$W(\mathbf{l}, \mathbf{l}') \equiv -\mathbf{l}' \cdot (\mathbf{l} - \mathbf{l}') \phi(\mathbf{l} - \mathbf{l}').$$

Note that lensed E -mode and B -mode polarization are mixed. That is why the most of lens reconstruction signal comes from the estimator using EB cross-correlation (see, Figure N.2).

How do we reconstruct the lensing potential from the observed CMB anisotropy? The key is that weak lensing also mixes different wave-modes and the mode-mixing strength are proportional to the Fourier transform of lensing potential $\phi(\mathbf{l})$. Let us first quantify the coupling strength. For fixed lensing potential, we define the mode coupling strength as

$$\langle x(\mathbf{l}) x'(\mathbf{l}') \rangle_{CMB} = f_\alpha(\mathbf{l}, \mathbf{l}') \phi(\mathbf{L}). \quad (\text{N.18})$$

Here, $\mathbf{L} = \mathbf{l} + \mathbf{l}'$, and x is one of anisotropy variables, $\{\Theta, E, B\}$, and α denotes the xx' pairing. We calculate the coupling strength f_α by using equations (N.15), (N.16) and (N.17)²:

$$f_{\Theta\Theta}(\mathbf{l}_1, \mathbf{l}_2) = \tilde{C}_{l_1}^{\Theta\Theta}(\mathbf{L} \cdot \mathbf{l}_1) + \tilde{C}_{l_2}^{\Theta\Theta}(\mathbf{L} \cdot \mathbf{l}_2) \quad (\text{N.19})$$

$$f_{\Theta E}(\mathbf{l}_1, \mathbf{l}_2) = \tilde{C}_{l_1}^{\Theta E} \cos(2\varphi_{l_1 l_2})(\mathbf{L} \cdot \mathbf{l}_1) + \tilde{C}_{l_2}^{\Theta E}(\mathbf{L} \cdot \mathbf{l}_2) \quad (\text{N.20})$$

$$f_{\Theta B}(\mathbf{l}_1, \mathbf{l}_2) = \tilde{C}_{l_1}^{\Theta E} \sin(2\varphi_{l_1 l_2})(\mathbf{L} \cdot \mathbf{l}_1) \quad (\text{N.21})$$

$$f_{EE}(\mathbf{l}_1, \mathbf{l}_2) = \left[\tilde{C}_{l_1}^{EE}(\mathbf{L} \cdot \mathbf{l}_1) + \tilde{C}_{l_2}^{EE}(\mathbf{L} \cdot \mathbf{l}_2) \right] \cos(2\varphi_{l_1 l_2}) \quad (\text{N.22})$$

$$f_{EB}(\mathbf{l}_1, \mathbf{l}_2) = \left[\tilde{C}_{l_1}^{EE}(\mathbf{L} \cdot \mathbf{l}_1) - \tilde{C}_{l_2}^{BB}(\mathbf{L} \cdot \mathbf{l}_2) \right] \sin(2\varphi_{l_1 l_2}) \quad (\text{N.23})$$

$$f_{BB}(\mathbf{l}_1, \mathbf{l}_2) = \left[\tilde{C}_{l_1}^{BB}(\mathbf{L} \cdot \mathbf{l}_1) + \tilde{C}_{l_2}^{BB}(\mathbf{L} \cdot \mathbf{l}_2) \right] \cos(2\varphi_{l_1 l_2}) \quad (\text{N.24})$$

However, we cannot use equation (N.18) as an estimator of ϕ , as $\langle \phi \rangle = 0$ due to the statistical isotropy. Instead, we estimate the deflection field $d(\hat{\mathbf{n}})$ by taking a weighted average over

²There is a typo in Table 1 of Hu & Okamoto (2002). For $\alpha = \Theta E$, $\cos(\varphi_{l_1 l_2})$ in the equation has to be $\cos(2\varphi_{l_1 l_2})$.

multipole moments. The estimate suggested by Hu & Okamoto (2002) is

$$d_\alpha(\mathbf{L}) = \frac{A_\alpha(L)}{L} \int \frac{d^2 l_1}{(2\pi)^2} x(\mathbf{l}_1) x'(\mathbf{l}_2) F_\alpha(\mathbf{l}_1, \mathbf{l}_2). \quad (\text{N.25})$$

Here, we introduce a normalization factor

$$A_\alpha(L) = L^2 \left[\int \frac{d^2 l_1}{(2\pi)^2} f_\alpha(\mathbf{l}_1, \mathbf{l}_2) F_\alpha(\mathbf{l}_1, \mathbf{l}_2) \right]^{-1}, \quad (\text{N.26})$$

so that

$$\langle d_\alpha(\mathbf{L}) \rangle_{CMB} \equiv L \phi(\mathbf{L}) \quad (\text{N.27})$$

is satisfied. By minimizing the variance $\langle d_\alpha(L) d_\alpha(L) \rangle$, Hu & Okamoto (2002) has calculate the minimum variance filter $F_\alpha(\mathbf{l}_1, \mathbf{l}_2)$ as:

$$F_\alpha(\mathbf{l}_1, \mathbf{l}_2) = \frac{C_{l_1}^{x'x'} C_{l_2}^{xx} f_\alpha(\mathbf{l}_1, \mathbf{l}_2) - C_{l_1}^{xx'} C_{l_2}^{xx'} f_\alpha(\mathbf{l}_2, \mathbf{l}_1)}{C_{l_1}^{xx} C_{l_2}^{x'x'} C_{l_1}^{x'x'} C_{l_2}^{xx} - (C_{l_1}^{xx'} C_{l_2}^{xx'})^2}, \quad (\text{N.28})$$

For $\alpha = \Theta\Theta$, EE and BB , F_α is reduced to

$$F_\alpha(\mathbf{l}_1, \mathbf{l}_2) = \frac{f_\alpha(\mathbf{l}_1, \mathbf{l}_2)}{2 C_{l_1}^{xx} C_{l_2}^{xx}}, \quad (\text{N.29})$$

and if $\tilde{C}_l^{xx'} = 0$ as in the case of $\alpha = \Theta B$ and EB ,

$$F_\alpha(\mathbf{l}_1, \mathbf{l}_2) = \frac{f_\alpha(\mathbf{l}_1, \mathbf{l}_2)}{C_{l_1}^{xx} C_{l_2}^{x'x'}}. \quad (\text{N.30})$$

What about the noise matrix of lensing reconstruction? The noise matrix for lensing reconstruction,

$$\langle d_\alpha^*(\mathbf{L}) d_\beta(\mathbf{L}') \rangle = (2\pi)^2 [C_L^{dd} + N_{\alpha\beta}(L)] \delta^D(\mathbf{L} - \mathbf{L}'), \quad (\text{N.31})$$

is calculated as

$$\begin{aligned} N_{\alpha\beta}(L) &= \frac{A_\alpha(L) A_\beta(L)}{L^2} \int \frac{d^2 l_1}{(2\pi)^2} F_\alpha(\mathbf{l}_1, \mathbf{l}_2) \\ &\times \left[F_\beta(\mathbf{l}_1, \mathbf{l}_2) C_{l_1}^{x_\alpha x_\beta} C_{l_2}^{x'_\alpha x'_\beta} + F_\beta(\mathbf{l}_2, \mathbf{l}_1) C_{l_1}^{x_\alpha x'_\beta} C_{l_2}^{x'_\alpha x_\beta} \right]. \end{aligned} \quad (\text{N.32})$$

Note that for the minimum variance filter,

$$N_{\alpha\alpha}(L) = A_\alpha(L).$$

Now we have five different estimators where $\alpha = \Theta\Theta$, ΘE , ΘB , EE , EB , and $N_{\alpha\beta}$ is a 5 by 5 matrix. We linearly combine these estimators

$$d_{\text{mv}}(\mathbf{L}) = \sum_\alpha w_\alpha(L) d_\alpha(\mathbf{L}) \quad (\text{N.33})$$

with coefficient determined by

$$w_\alpha = \frac{\sum_\beta (N^{-1})_{\alpha\beta}}{\text{Tr}(N^{-1})}.$$

Then, final noise power spectrum for lensing reconstruction is

$$N_{\text{mv}}(L) = \frac{1}{\sum_{\alpha\beta} (N^{-1})_{\alpha\beta}}. \quad (\text{N.34})$$

Finally, as our analysis in Chapter 6 is based on the convergence field $\kappa \equiv \nabla^2 \phi / 2$, the noise power spectrum in equation (N.34) has to be properly rescaled as following. The deflection field d is defined to be $d = l\phi$ in equation (N.27), therefore, the convergence power spectrum is related to the deflection field power spectrum as

$$C_l^{\kappa\kappa} = \frac{1}{4} l^4 C_l^{\phi\phi} = \frac{1}{4} l^2 C_l^{dd}. \quad (\text{N.35})$$

The noise power spectrum of convergence field scales in the same way:

$$N_l^\kappa = \frac{1}{4} l^2 N_{\text{mv}}(l). \quad (\text{N.36})$$

Figure N.1 and N.2 show two examples of noise power spectra. In those figures, we show the noise power spectra of each of five estimators, $d_{\Theta\Theta}$, $d_{\Theta E}$, $d_{\Theta B}$, d_{EE} , d_{EB} as well as that of the minimum variance linear combination d_{mv} (*orange solid line*).

Figure N.1 shows the noise power spectra for *Planck* mission. As Planck satellite is expected to observe EE and ΘE correlation but probably not to observe the EB and BB correlation, even minimum variance noise power spectrum exceeds the lensing potential power spectrum. Therefore, we cannot measure the lensing potential power spectrum for a single wavenumber for Planck. Instead, we may have to constrain a few parameters which depends on the whole shape of the lensing potential power spectrum.

On the other hand, Figure N.2 shows the noise power spectra for the *nearly perfect* experiment suggested by Hu & Okamoto (2002). For this case, the instrumental noise is so small that we can detect the B -mode polarization from CMB lensing down to $l \sim 1000$ (Figure 2 of Hu & Okamoto (2002)), and we can detect the lensing potential for individual wave mode until $l \sim 1000$. Note that the best estimator is d_{EB} provide the best estimator. It is because B -mode polarization in this analysis is solely from weak lensing of E -mode polarization as we ignore the primordial tensor mode, Since we need to first reconstruct the lensing potential map in order to cross correlate them with the lens galaxies, we use latter case for our analysis in Chapter 6.

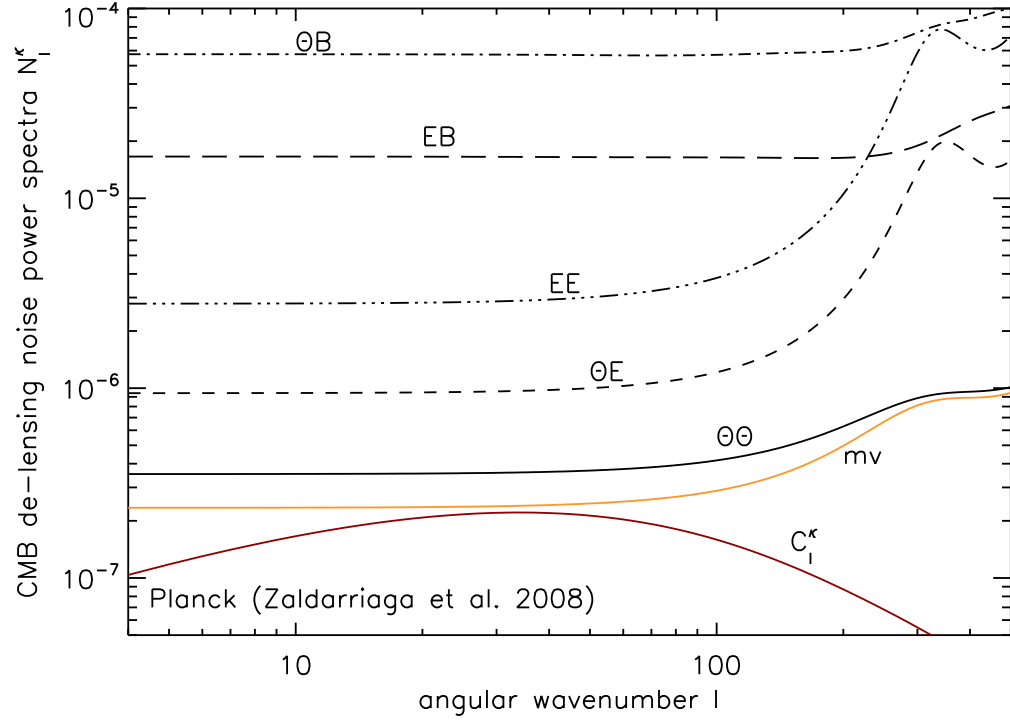


Figure N.1: Noise power spectra of Cosmic Microwave Background lensing reconstruction. We show the noise power spectrum of $d_{\Theta\Theta}$ (solid line), $d_{\Theta E}$ (dashed line), $d_{\Theta B}$ (dot-dashed line), d_{EE} (dots-dashed line), d_{EB} (long-dashed line), and minimum variance estimator d_{mv} (orange solid line). For comparison, we also show the convergence power spectrum C_l^κ (red line). Noise power spectrum is calculated for Planck satellite: $\Delta_T = 35.4 \mu\text{Karcmin}$, $\Delta_P = 63.1 \mu\text{Karcmin}$, and $\sigma = 7 \text{ arcmin}$ as described in Appendix A of Zaldarriaga et al. (2008).

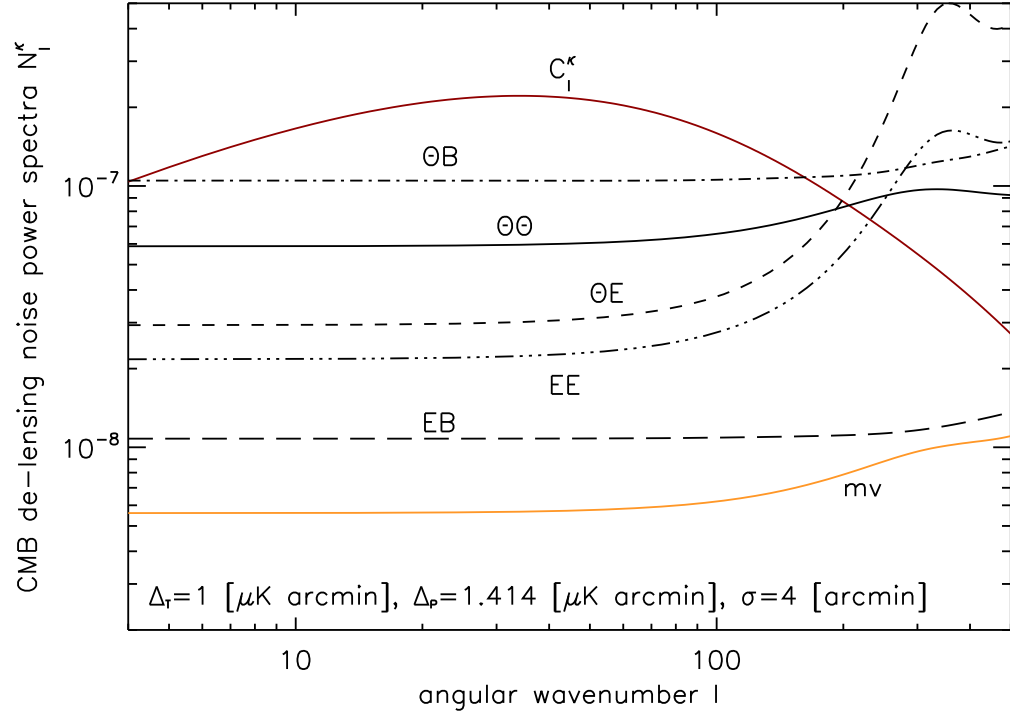


Figure N.2: Same as Figure N.1, but for the *nearly perfect* experiment quoted by Hu & Okamoto (2002). We use $\Delta_T = 1 \mu\text{K arcmin}$, $\Delta_P = \sqrt{2} \mu\text{K arcmin}$, and $\sigma = 4 \text{ arcmin}$. In Chapter 6, we estimate the noise power spectrum (N_l^κ) by minimum variance estimator (orange solid line).

Bibliography

- Acquaviva, V., Bartolo, N., Matarrese, S., & Riotto, A. 2003, Nucl. Phys., B667, 119
- Acquaviva, V., Hajian, A., Spergel, D. N., & Das, S. 2008, Phys. Rev., D78, 043514
- Afshordi, N. & Tolley, A. J. 2008, Phys. Rev. D, 78, 123507
- Albrecht, A. & Steinhardt, P. J. 1982, Phys. Rev. Lett., 48, 1220
- Angulo, R. E., Baugh, C. M., Frenk, C. S., & Lacey, C. G. 2008, Mon. Not. R. Astron. Soc., 383, 755
- Babich, D., Creminelli, P., & Zaldarriaga, M. 2004, JCAP, 0408, 009
- Ballinger, W. E., Peacock, J. A., & Heavens, A. F. 1996, Mon. Not. R. Astron. Soc., 282, 877
- Bardeen, J. M., Bond, J. R., Kaiser, N., & Szalay, A. S. 1986, Astrophys. J., 304, 15
- Bardeen, J. M., Steinhardt, P. J., & Turner, M. S. 1983, Phys. Rev. D, 28, 679
- Bartelmann, M. & Schneider, P. 2001, Phys. Rep., 340, 291
- Bartolo, N., Komatsu, E., Matarrese, S., & Riotto, A. 2004, Phys. Rept., 402, 103
- Benson, A. J., Bower, R. G., Frenk, C. S., Lacey, C. G., Baugh, C. M., & Cole, S. 2003, Astrophys. J., 599, 38
- Bernardeau, F., Colombi, S., Gaztañaga, E., & Scoccimarro, R. 2002, Phys. Rep., 367, 1
- Bertschinger, E. 1995, ArXiv Astrophysics e-prints
- Blake, C. & Glazebrook, K. 2003, Astrophys. J., 594, 665
- Bond, J. R., Cole, S., Efstathiou, G., & Kaiser, N. 1991, Astrophys. J., 379, 440
- Bond, J. R., Efstathiou, G., Lubin, P. M., & Meinhold, P. R. 1991, Phys. Rev. Lett., 66, 2179
- Boubekeur, L. & Lyth, D. H. 2006, Phys. Rev. D, 73, 021301

- Bouchet, F. R., Colombi, S., Hivon, E., & Juszkievicz, R. 1995, *Astron. Astrophys.*, 296, 575
- Boughn, S. & Crittenden, R. 2004, *Nature*, 427, 45
- Bower, R. G., Benson, A. J., Malbon, R., Helly, J. C., Frenk, C. S., Baugh, C. M., Cole, S., & Lacey, C. G. 2006, *Mon. Not. R. Astron. Soc.*, 370, 645
- Brainerd, T. G., Blandford, R. D., & Smail, I. 1996, *Astrophys. J.*, 466, 623
- Buchbinder, E. I., Khoury, J., & Ovrut, B. A. 2008, *Phys. Rev. Lett.*, 100, 171302
- Byrnes, C. T., Sasaki, M., & Wands, D. 2006, *Phys. Rev. D*, 74, 123519
- Carbone, C., Verde, L., & Matarrese, S. 2008, *Astrophys. J. Lett.*, 684, L1
- Chabrier, G. 2003, *Publ. Astron. Soc. Pac.*, 115, 763
- Chen, X., Huang, M.-x., Kachru, S., & Shiu, G. 2007, *JCAP*, 0701, 002
- Cheung, C., Creminelli, P., Fitzpatrick, A. L., Kaplan, J., & Senatore, L. 2008, *JHEP*, 03, 014
- Cole, S. & Kaiser, N. 1989, *Mon. Not. R. Astron. Soc.*, 237, 1127
- Cole, S., Lacey, C. G., Baugh, C. M., & Frenk, C. S. 2000, *Mon. Not. R. Astron. Soc.*, 319, 168
- Cole, S., et al. 2005, *Mon. Not. R. Astron. Soc.*, 362, 505
- Cooray, A. & Sheth, R. 2002, *Phys. Rep.*, 372, 1
- Copeland, E. J., Sami, M., & Tsujikawa, S. 2006, *Int. J. Mod. Phys.*, D15, 1753
- Creminelli, P. & Zaldarriaga, M. 2004, *J. Cosmol. Astropart. Phys.*, 10, 6
- Crocce, M., Pueblas, S., & Scoccimarro, R. 2006, *Mon. Not. R. Astron. Soc.*, 373, 369
- Crocce, M. & Scoccimarro, R. 2008, *Phys. Rev. D*, 77, 023533
- Croton, D. J., et al. 2006, *Mon. Not. R. Astron. Soc.*, 365, 11
- Dalal, N., Doré, O., Huterer, D., & Shirokov, A. 2008, *Phys. Rev. D*, 77, 123514
- de Lai, A. A. & Starkman, G. D. 1998, *Mon. Not. R. Astron. Soc.*, 299, 977

- De Lucia, G. & Blaizot, J. 2007, *Mon. Not. R. Astron. Soc.*, 375, 2
- Desjacques, V. & Seljak, U. 2010, *Phys. Rev. D*, 81, 023006
- Desjacques, V., Seljak, U., & Iliev, I. T. 2009, *Mon. Not. R. Astron. Soc.*, 396, 85
- Dunkley, J., et al. 2009, *Astrophys. J.*, submitted arXiv:0811.4280v1
- Efstathiou, G., Davis, M., White, S. D. M., & Frenk, C. S. 1985, *Astrophys. J. Suppl.*, 57, 241
- Eisenstein, D. J. & Hu, W. 1998, *Astrophys. J.*, 496, 605
- Eisenstein, D. J., Seo, H.-J., & White, M. 2007, *Astrophys. J.*, 664, 660
- Eisenstein, D. J., et al. 2005, *Astrophys. J.*, 633, 560
- Falk, T., Rangarajan, R., & Srednicki, M. 1993, *Astrophys. J. Lett.*, 403, L1
- Feldman, H. A., Kaiser, N., & Peacock, J. A. 1994, *Astrophys. J.*, 426, 23
- Frieman, J. A., Turner, M. S., & Huterer, D. 2008, *Ann. Rev. Astron. Astrophys.*, 46, 385
- Fry, J. N. 1984, *Astrophys. J.*, 279, 499
- . 1996, *Astrophys. J. Lett.*, 461, L65+
- Fry, J. N. & Gaztanaga, E. 1993, *Astrophys. J.*, 413, 447
- Gabrielli, A., Labini, F. S., Joyce, M., & Pietronero, L. 2005, *Statistical physics for cosmic structure* (Berlin Heidelberg, DE: Springer-Verlag)
- Gangui, A., Lucchin, F., Matarrese, S., & Mollerach, S. 1994, *Astrophys. J.*, 430, 447
- Giannantonio, T. & Porciani, C. 2010, *Phys. Rev. D*, 81, 063530
- Goroff, M. H., Grinstein, B., Rey, S., & Wise, M. B. 1986, *Astrophys. J.*, 311, 6
- Grinstein, B. & Wise, M. B. 1986, *Astrophys. J.*, 310, 19
- Grossi, M., Verde, L., Carbone, C., Dolag, K., Branchini, E., Iannuzzi, F., Matarrese, S., & Moscardini, L. 2009, *ArXiv e-prints*
- Guth, A. H. 1981, *Phys. Rev. D*, 23, 347
- Guth, A. H. & Pi, S. Y. 1982, *Phys. Rev. Lett.*, 49, 1110

- Guzik, J. & Seljak, U. 2002, *Mon. Not. R. Astron. Soc.*, 335, 311
- Hamilton, A. J. S. 1998, in *Astrophysics and Space Science Library*, Vol. 231, *The Evolving Universe*, ed. D. Hamilton, 185–+
- Hamilton, A. J. S., Kumar, P., Lu, E., & Matthews, A. 1991, *Astrophys. J. Lett.*, 374, L1
- Hawking, S. W. 1982, *Phys. Lett.*, B115, 295
- Heavens, A. F., Matarrese, S., & Verde, L. 1998, *Mon. Not. R. Astron. Soc.*, 301, 797
- Heitmann, K., et al. 2007, *ArXiv e-prints*, 706
- Hill, G. J., Gebhardt, K., Komatsu, E., & MacQueen, P. J. 2004, 743, 224
- Hill, G. J. et al. 2008
- Hirata, C. M. & Seljak, U. 2003, *Phys. Rev.*, D68, 083002
- Hockney, R. W. & Eastwood, J. W. 1988, *Computer simulation using particles*, ed. Hockney, R. W. & Eastwood, J. W.
- Hu, W. 2000, *Phys. Rev. D*, 62, 43007
- Hu, W. & Haiman, Z. 2003, *Phys. Rev. D*, 68, 063004
- Hu, W. & Jain, B. 2004, *Phys. Rev. D*, 70, 043009
- Hu, W. & Okamoto, T. 2002, *Astrophys. J.*, 574, 566
- Huang, M.-X. & Shiu, G. 2006, *Phys. Rev. D*, 74, 121301
- Huff, E., Schulz, A. E., White, M., Schlegel, D. J., & Warren, M. S. 2007, *Astroparticle Physics*, 26, 351
- Jackson, J. C. 1972, *Mon. Not. R. Astron. Soc.*, 156, 1P
- Jain, B. & Bertschinger, E. 1994, *Astrophys. J.*, 431, 495
- Jeong, D. & Komatsu, E. 2006, *Astrophys. J.*, 651, 619
- . 2009, *Astrophys. J.*, 691, 569
- Jing, Y. P. 2005, *Astrophys. J.*, 620, 559
- Johnston, D. E., et al. 2007, *ArXiv e-prints*

- Kaiser, N. 1984, *Astrophys. J. Lett.*, 284, L9
- . 1987, *Mon. Not. R. Astron. Soc.*, 227, 1
- . 1992, *Astrophys. J.*, 388, 272
- Kennicutt, Jr., R. C. 1983, *Astrophys. J.*, 272, 54
- Kitayama, T. & Suto, Y. 1996, *Astrophys. J.*, 469, 480
- Knox, L. 1995, *Phys. Rev. D*, 52, 4307
- Kogo, N. & Komatsu, E. 2006, *Phys. Rev.*, D73, 083007
- Komatsu, E. 2001, *ArXiv Astrophysics e-prints*, ph.D. thesis at Tohoku University (astro-ph/0206039)
- Komatsu, E. & Spergel, D. N. 2001, *Phys. Rev. D*, 63, 63002
- Komatsu, E., Wandelt, B. D., Spergel, D. N., Banday, A. J., & Górski, K. M. 2002, *Astrophys. J.*, 566, 19
- Komatsu, E., et al. 2003, *Astrophys. J. Suppl.*, 148, 119
- . 2009, *Astrophys. J. Suppl.*, 180, 330
- Komatsu, E. et al. 2009
- Komatsu, E., et al. 2010, *ArXiv e-prints*
- Lacey, C. & Cole, S. 1993, *Mon. Not. R. Astron. Soc.*, 262, 627
- Lewis, A. & Challinor, A. 2006, *Phys. Rep.*, 429, 1
- Limber, D. N. 1954, *Astrophys. J.*, 119, 655
- Linde, A. D. 1982, *Phys. Lett.*, B108, 389
- LoVerde, M., Miller, A., Shandera, S., & Verde, L. 2008, *JCAP*, 0804, 014
- Lucchin, F. & Matarrese, S. 1988, *Astrophys. J.*, 330, 535
- Lukić, Z., Heitmann, K., Habib, S., Bashinsky, S., & Ricker, P. M. 2007, *Astrophys. J.*, 671, 1160
- Makino, N., Sasaki, M., & Suto, Y. 1992, *Phys. Rev. D*, 46, 585

- Maldacena, J. M. 2003, JHEP, 05, 013
- Mandelbaum, R., Seljak, U., Cool, R. J., Blanton, M., Hirata, C. M., & Brinkmann, J. 2006a, Mon. Not. R. Astron. Soc., 372, 758
- Mandelbaum, R., Seljak, U., Kauffmann, G., Hirata, C. M., & Brinkmann, J. 2006b, Mon. Not. R. Astron. Soc., 368, 715
- Mandelbaum, R., et al. 2005, Mon. Not. R. Astron. Soc., 361, 1287
- Matarrese, S., Lucchin, F., & Bonometto, S. A. 1986, Astrophys. J. Lett., 310, L21
- Matarrese, S. & Pietroni, M. 2007, Journal of Cosmology and Astro-Particle Physics, 6, 26
- Matarrese, S. & Verde, L. 2008, Astrophys. J. Lett., 677, L77
- Matarrese, S., Verde, L., & Heavens, A. F. 1997, Mon. Not. R. Astron. Soc., 290, 651
- Matarrese, S., Verde, L., & Jimenez, R. 2000, Astrophys. J., 541, 10
- Matsubara, T. 1995, Progress of Theoretical Physics, 94, 1151
- . 2008, Phys. Rev. D, 77, 063530
- Matsubara, T. & Szalay, A. S. 2003, Physical Review Letters, 90, 021302
- McDonald, P. 2006, Phys. Rev. D, 74, 103512
- . 2007, Phys. Rev. D, 75, 043514
- . 2008, Phys. Rev. D, 78, 123519
- McKay, T. A., et al. 2001, ArXiv Astrophysics e-prints
- Meiksin, A., White, M., & Peacock, J. A. 1999, Mon. Not. R. Astron. Soc., 304, 851
- Mo, H. J., Jing, Y. P., & White, S. D. M. 1997, Mon. Not. R. Astron. Soc., 284, 189
- Mo, H. J. & White, S. D. M. 1996, Mon. Not. R. Astron. Soc., 282, 347
- Mukhanov, V. F., Feldman, H. A., & Brandenberger, R. H. 1992, Phys. Rept., 215, 203
- Nishimichi, T., et al. 2007, Publ. Astron. Soc. Jpn., 59, 1049
- Nishioka, H. & Yamamoto, K. 2000, Astrophys. J. Suppl., 128, 1

- Noh, H. & Hwang, J. 2008, *Phys. Rev. D*, 77, 123533
- Noh, H., Jeong, D., & Hwang, J. 2009, *Physical Review Letters*, 103, 021301
- Okamoto, T. & Hu, W. 2002, *Phys. Rev. D*, 66, 063008
- . 2003, *Phys. Rev. D*, 67, 83002
- Peacock, J. A. & Dodds, S. J. 1996, *Mon. Not. R. Astron. Soc.*, 280, L19
- Peebles, P. J. E. 1980, *The large-scale structure of the universe* (Research supported by the National Science Foundation. Princeton, N.J., Princeton University Press, 1980. 435 p.)
- Percival, W. J., Cole, S., Eisenstein, D. J., Nichol, R. C., Peacock, J. A., Pope, A. C., & Szalay, A. S. 2007, *Mon. Not. R. Astron. Soc.*, 381, 1053
- Perlmutter, S., et al. 1999, *Astrophys. J.*, 517, 565
- Pillepich, A., Porciani, C., & Hahn, O. 2008, *ArXiv e-prints*
- Politzer, H. D. & Wise, M. B. 1984, *Astrophys. J. Lett.*, 285, L1
- Press, W. H. & Schechter, P. 1974, *Astrophys. J.*, 187, 425
- Pueblas, S. & Scoccimarro, R. 2009, *Phys. Rev. D*, 80, 043504
- Reid, B. A., et al. 2010, *Mon. Not. R. Astron. Soc.*, 404, 60
- Riess, A. G., et al. 1998, *Astron. J.*, 116, 1009
- Ryu, D., Ostriker, J. P., Kang, H., & Cen, R. 1993, *Astrophys. J.*, 414, 1
- Saito, S., Takada, M., & Taruya, A. 2009, *Phys. Rev. D*, 80, 083528
- Salopek, D. S. & Bond, J. R. 1990, *Phys. Rev. D*, 42, 3936
- Sanchez, A. G., Baugh, C. M., & Angulo, R. 2008, *ArXiv e-prints*, 804
- Sato, K. 1981, *Mon. Not. R. Astron. Soc.*, 195, 467
- Scherrer, R. J. & Bertschinger, E. 1991, *Astrophys. J.*, 381, 349
- Schlegel, D., White, M., & Eisenstein, D. 2009, *ArXiv e-prints*
- Scoccimarro, R. 1998, *Mon. Not. R. Astron. Soc.*, 299, 1097

- . 2004, *Phys. Rev. D*, 70, 083007
- Scoccimarro, R., Colombi, S., Fry, J. N., Frieman, J. A., Hivon, E., & Melott, A. 1998, *Astrophys. J.*, 496, 586
- Scoccimarro, R., Feldman, H. A., Fry, J. N., & Frieman, J. A. 2001a, *Astrophys. J.*, 546, 652
- Scoccimarro, R. & Frieman, J. A. 1996, *Astrophys. J.*, 473, 620
- Scoccimarro, R., Sefusatti, E., & Zaldarriaga, M. 2004, *Phys. Rev. D*, 69, 103513
- Scoccimarro, R., Sheth, R. K., Hui, L., & Jain, B. 2001b, *Astrophys. J.*, 546, 20
- Seery, D. & Lidsey, J. E. 2005, *JCAP*, 0506, 003
- Sefusatti, E. 2009, *Phys. Rev. D*, 80, 123002
- Sefusatti, E., Crocce, M., Pueblas, S., & Scoccimarro, R. 2006, *Phys. Rev. D*, 74, 023522
- Sefusatti, E. & Komatsu, E. 2007, *Phys. Rev. D*, 76, 083004
- Sefusatti, E., Vale, C., Kadota, K., & Frieman, J. 2007, *Astrophys. J.*, 658, 669
- Seljak, U. 2000, *Mon. Not. R. Astron. Soc.*, 318, 203
- Seljak, U., et al. 2005, *Phys. Rev. D*, 71, 103515
- Seo, H.-J. & Eisenstein, D. J. 2003, *Astrophys. J.*, 598, 720
- Seo, H.-J. & Eisenstein, D. J. 2005, *Astrophys. J.*, 633, 575
- Seo, H.-J., Siegel, E. R., Eisenstein, D. J., & White, M. 2008, *ArXiv e-prints*, 805
- Sheldon, E. S., et al. 2004, *Astron. J.*, 127, 2544
- . 2007, *ArXiv e-prints*
- Sheth, R. K., Mo, H. J., & Tormen, G. 2001, *Mon. Not. R. Astron. Soc.*, 323, 1
- Sheth, R. K. & Tormen, G. 1999, *Mon. Not. R. Astron. Soc.*, 308, 119
- Shoji, M., Jeong, D., & Komatsu, E. 2009, *Astrophys. J.*, 693, 1404
- Shoji, M. & Komatsu, E. 2009, *Astrophys. J.*, 700, 705

- Slosar, A., Hirata, C., Seljak, U., Ho, S., & Padmanabhan, N. 2008, *Journal of Cosmology and Astro-Particle Physics*, 8, 31
- Smith, R. E., Scoccimarro, R., & Sheth, R. K. 2007, *Phys. Rev. D*, 75, 063512
- . 2008a, *Phys. Rev. D*, 77, 043525
- Smith, R. E., Sheth, R. K., & Scoccimarro, R. 2008b, *Phys. Rev. D*, 78, 023523
- Smith, R. E., et al. 2003, *Mon. Not. R. Astron. Soc.*, 341, 1311
- Springel, V. 2005, *Mon. Not. R. Astron. Soc.*, 364, 1105
- Springel, V., Yoshida, N., & White, S. D. M. 2001, *New Astronomy*, 6, 79
- Springel, V., et al. 2005, *Nature*, 435, 629
- Starobinsky, A. A. 1982, *Phys. Lett.*, B117, 175
- Stebbins, A. 1996, *ArXiv Astrophysics e-prints*
- Suto, Y. & Sasaki, M. 1991, *Physical Review Letters*, 66, 264
- Takada, M., Komatsu, E., & Futamase, T. 2006, *Phys. Rev. D*, 73, 083520
- Takahashi, R. 2008, *Progress of Theoretical Physics*, 120, 549
- Takahashi, R., et al. 2008, *ArXiv e-prints*, 802
- Taruya, A. & Hiramatsu, T. 2008, *Astrophys. J.*, 674, 617
- Taruya, A., Koyama, K., & Matsubara, T. 2008, *Phys. Rev.*, D78, 123534
- Tegmark, M. 1997, *Physical Review Letters*, 79, 3806
- Tegmark, M. & Peebles, P. J. E. 1998, *Astrophys. J. Lett.*, 500, L79+
- Tegmark, M., et al. 2004, *Astrophys. J.*, 606, 702
- . 2006, *Phys. Rev. D*, 74, 123507
- Tyson, J. A., Valdes, F., Jarvis, J. F., & Mills, Jr., A. P. 1984, *Astrophys. J. Lett.*, 281, L59
- Valageas, P. 2007, *Astron. Astrophys.*, 465, 725

- Verde, L., Heavens, A. F., Matarrese, S., & Moscardini, L. 1998, *Mon. Not. R. Astron. Soc.*, 300, 747
- Verde, L., Wang, L., Heavens, A. F., & Kamionkowski, M. 2000, *Mon. Not. R. Astron. Soc.*, 313, 141
- Vishniac, E. T. 1983, *Mon. Not. R. Astron. Soc.*, 203, 345
- Wagner, C., Müller, V., & Steinmetz, M. 2008, *Astron. Astrophys.*, 487, 63
- Wang, L. & Kamionkowski, M. 2000, *Phys. Rev. D*, 61, 63504
- Weinberg, S. 2008, *Cosmology* (Oxford, UK: Oxford University Press)
- White, M. 2005, *Astroparticle Physics*, 24, 334
- Yamamoto, K., Nishioka, H., & Suto, Y. 1999, *Astrophys. J.*, 527, 488
- Yoshikawa, K., Taruya, A., Jing, Y. P., & Suto, Y. 2001, *Astrophys. J.*, 558, 520
- Zaldarriaga, M., et al. 2008, *ArXiv e-prints*
- Zel'dovich, Y. B. 1970, *Astron. Astrophys.*, 5, 84
- Zhan, H. 2006, *Journal of Cosmology and Astro-Particle Physics*, 8, 8

Vita

Donghui Jeong was born in a small town in the middle of South Korea as a elder son of K.-B. Jeong and N.-S. Lee. He graduated in 1998 from Choong-Book Science High School, and enrolled at Korea Advanced Institute of Science and Technology (KAIST). In 2001, he started studying cosmology with an insightful instruction of Dr. Ewan Stewart. Donghui graduated in 2002 with a Bachelor of Science degree in Physics (Mathematics minor). That fall he enrolled at KAIST as a master course graduate student and working with Dr. Ewan Stewart on the theoretical ideas of how to generate baryon/anti-baryon asymmetry in the early Universe. Donghui received a Master of Science degree in Physics from KAIST in 2004. After spending one more year in KAIST as a doctoral student, he decided to study our Universe based upon solid observation rather than mathematical completeness. In the fall of 2005 he enrolled at the Department of Astronomy of the University of Texas at Austin. There, he worked with Dr. Eiichiro Komatsu on various theoretical topics related to the high redshift galaxy surveys. He completed his Ph.D. in the summer of 2010. He will be taking up a position as a Robinson postdoctoral scholar in theoretical cosmology at California Institute of Technology in the fall of 2010.

Permanent address: 2688 Yongam-dong, Sangdang-gu, Cheongju
ChoongcheongBook-do, South Korea 360-812

This dissertation was typeset with \LaTeX^\dagger by the author.

[†] \LaTeX is a document preparation system developed by Leslie Lamport as a special version of Donald Knuth's \TeX Program.