

Theoretical Considerations in Large-Scale Structure Analyses

by

Michael Shengbo Wang



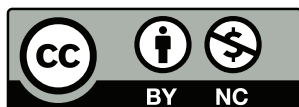
January 2021

The thesis is submitted in partial fulfilment of the requirements for
the award of the degree of
Doctor of Philosophy
of the
UNIVERSITY OF PORTSMOUTH.

Whilst registered as a candidate for the above degree, I have not been registered for any other research award. The results and conclusions embodied in this thesis are the work of the named candidate and have not been submitted for any other academic award.

The contents of this thesis were examined in a viva voce held on 31 March 2021 by **Prof Andrew Jaffe** and **Prof Adam Amara**, to whom special thanks are owed. All research underlying this thesis has passed ethics review at the University of Portsmouth with certification code A416-53F4-4B64-8B71-CB54-EB17-4A93-4D3D.

This thesis has an estimated word count of 42852 for the abstract, main matter and appendices. Captions and footnotes are included in this count. Wherever possible, the typesetting of this thesis adheres to the ISO 4 and ISO 80000 standards.



The contents of this work are licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>).

Repository version 2021.05.20.f0e3f36

© Mike Shengbo Wang 2021. Typeset with X_YL^AT_EX in the *memoir* class.

ABSTRACT

The study of the large-scale structure of the Universe is an integral part of modern cosmology, which is entering a golden era with the arrival of new generations of ambitious cosmological surveys. This thesis examines different cross-sections of the interface where theoretical predictions are confronted with observational measurements, and proposes a number of novel methodologies that can improve the constraints of cosmological models from the large-scale clustering of galaxies tracing the underlying matter distribution:

- 1) a hybrid-basis Fourier analysis approach, which balances the accuracy required for modelling the anisotropic galaxy distribution in wide and deep surveys against the computational cost of processing an unprecedented amount of cosmological data;
- 2) forward modelling from measurements of the tracer luminosity function to constraints on the amplitude of relativistic corrections, which can give rise to scale-dependent modifications to clustering statistics on scales close to the Hubble horizon;
- 3) statistical techniques within the existing survey analysis framework to mitigate the impact on cosmological inference from non-Gaussian likelihoods and the parameter dependence of covariance matrices.

A special focus in designing these new techniques for galaxy clustering analysis is the detection of local primordial non-Gaussianity f_{NL} , which parametrises initial conditions of structure formation as set by inflationary models and imprints a scale-dependent signature in the bias of tracers on very large scales. These new approaches to analysing large-scale structure observations are also broadly applicable to constraining other standard cosmological parameters and should be particularly beneficial to forthcoming galaxy surveys such as the Dark Energy Spectroscopic Instrument and the *Euclid* mission, the scientific objectives of which include drastically tighter bounds on primordial non-Gaussianity as well as the test of gravitational theories on cosmological scales.

“ *I design to speak of the Physical, Metaphysical and Mathematical—of the Material and Spiritual Universe:—of its Essence, its Origin, its Creation, its Present Condition and its Destiny. I shall be so rash, moreover, as to challenge the conclusions, and thus, in effect, to question the sagacity, of many of the greatest and most justly revered of men.* ”

— Edgar Allan Poe, *Eureka: A Prose Poem*

Contents

List of Tables	vi
List of Figures	vii
Abbreviations	x
Notations	xii
Preface	xvi
1 Introduction	2
1.1 Statistics of cosmic random fields	5
1.2 Anisotropic galaxy clustering	11
1.3 Primordial non-Gaussianity signature	17
1.4 Cosmological likelihood inference	21
1.5 Motivation and outline	25
2 Galaxy Clustering Analysis	26
2.1 The Feldman–Kaiser–Peacock formalism	26
2.2 Local plane-parallel estimators	30
2.3 Window convolution of models	33
3 Power Spectrum Gaussianisation and Covariance Decomposition	37
3.1 Distribution of power spectrum measurements	39
3.2 Beyond standard Gaussian likelihoods	42
3.3 Testing with Monte Carlo simulations	47
3.4 Likelihood construction for survey application	58
3.5 Summary and discussion	60
4 Impact of Relativistic Effects in Large-Scale Clustering	62
4.1 Relativistic effects in galaxy clustering	63
4.2 Luminosity function model fitting	65

4.3	Constraints on relativistic corrections	67
4.4	Scale-dependent modifications	73
4.5	Summary and discussion	76
5	Hybrid-Basis Fourier Analysis of Anisotropic Clustering	78
5.1	Spherical Fourier analysis	80
5.2	Hybrid-basis likelihood inference	85
5.3	Application to N -body simulations	90
5.4	Summary and discussion	100
6	Conclusion	103
6.1	Summary review	104
6.2	Future work and outlook	106
A	Cosmological Background	110
B	Linearised Cosmological Perturbations	114
C	Non-Linear Structure Formation	119
	Bibliography	123

List of Tables

3.1	Different forms of the likelihood function for power spectrum measurements with or without data Gaussianisation and/or covariance matrix marginalisation and estimate rescaling.	50
3.2	Maximum likelihood estimates and uncertainties of f_{NL} from various likelihood functions considered in Table 3.1.	53
3.3	Posterior median estimates and uncertainties of f_{NL} from various likelihood functions considered in Table 3.1.	54
3.4	KL divergence values with respect to the true f_{NL} posterior for the posteriors of various likelihood functions considered in Table 3.1.	57
4.1	Posterior median estimates of the PLE model parameters for the eBOSS QSO LF measurements.	67
5.1	Posterior median estimates and uncertainties of f_{NL} and b_1 from the hybrid-basis and Cartesian power spectrum likelihood analyses of halo mock catalogues.	98

List of Figures

Except where noted otherwise, figures appearing in this thesis are original and licensed under the same terms (see copyright page). Some figures from prior publications by the author are reprinted or reproduced in accordance with either the Oxford University Press Standard Licence (hereafter ‘the Licence’) or the IOP Publishing Author Rights Policy (hereafter ‘the Policy’).

1.1	Cosmic concordance amongst multiple probes on the joint $\Omega_{m,0}-\Omega_{\Lambda,0}$ constraints. [Licensed under APS/SciPris™ number RNP/20/OCT/032209]	4
1.2	Schematic representation of the decomposition of an N -point correlator.	7
1.3	Poisson sampling of a continuous number density field.	10
1.4	Simulated galaxy distribution in real space and in redshift space. [Permission to reuse granted by copyright owner]	14
1.5	Measurements of the galaxy clustering power spectrum multipoles and the best-fitting model from BOSS Data Release 12. [Licensed under CC BY]	15
1.6	Squeezed triangular configuration and bispectrum shape for local PNG. [Licensed under CC BY]	19
1.7	Scale-dependent signature of local PNG in the clustering power spectrum.	20
2.1	Geometric configuration of the local plane-parallel picture.	30
3.1	Solution of the Box–Cox transformation parameter as a function of the fiducial gamma distribution shape parameter. [Reprinted under the Licence]	44
3.2	Sample distribution of simulated power spectrum measurements compared with the gamma distribution approximation and the normal distribution assumption.	48
3.3	Comparison of directly sampled covariance matrices and the rescaled fiducial estimates from simulated power spectrum data.	49
3.4	Comparison of the exact true likelihood, the standard likelihood and the novel likelihood for f_{NL} inferred from an ensemble of simulated power spectrum data.	51

3.5	Comparison of the exact true likelihood, the standard likelihood and the novel likelihood for f_{NL} inferred from some individual power spectrum data realisations.	52
3.6	Comparison of the exact true likelihood, the standard likelihood and the novel likelihood for the Hubble parameter H_0 inferred from an ensemble of simulated power spectrum data.	53
3.7	Histogram of the posterior median estimates and credible bounds for f_{NL} produced from the exact true likelihood, the standard likelihood and the novel likelihood for an ensemble of data realisations.	55
3.8	Q–Q plots comparing f_{NL} posterior distributions obtained from the different likelihood functions with or without data Gaussianisation and/or covariance estimate rescaling. [Reproduced under the Licence]	56
3.9	Q–Q plots of f_{NL} posterior distributions for comparing the relative impact of covariance matrix estimation and the evaluation of gamma distribution parameters using estimated band power variance. [Reproduced under the Licence]	59
4.1	Best-fitting PLE models for eBOSS QSO LF measurements in different redshift bins. [Reprinted under the Licence]	68
4.2	Derived measurements of the quasar comoving number density below the absolute-magnitude threshold $\bar{m} = -25.0$ as a function of redshift from the best-fitting eBOSS QSO LF in this work. [Reprinted under the Licence]	69
4.3	Derived measurements of evolution bias and magnification bias at the absolute-magnitude threshold $\bar{m} = -25.0$ from the best-fitting eBOSS QSO LF in this work. [Reprinted under the Licence]	70
4.4	Constraints on evolution bias and magnification bias at redshift $z = 2.0$ for absolute-magnitude threshold $\bar{m} = -25.0$ from the eBOSS QSO LF measurements under the PLE model. [Reprinted under the Licence]	71
4.5	Derived measurements of relativistic correction parameters at the absolute-magnitude threshold $\bar{m} = -25.0$ from the best-fitting eBOSS QSO LF in this work. [Reprinted under the Licence]	72
4.6	Quasar clustering power spectrum monopole and quadrupole, with and without scale-dependent modifications, at redshift $z = 0.87$ and 3.75 with the absolute-magnitude threshold $\bar{m} = -25.0$. [Reproduced under the Licence]	74
4.7	Scale-dependent modifications to the quasar clustering power spectrum monopole and quadrupole as a fraction of the Kaiser RSD model at wavenumber $k = 0.001 h \text{ Mpc}^{-1}$. [Reproduced under the Licence]	75
5.1	BOSS Data Release 12 CMASS NGC angular mask. [Reprinted under the Policy]	91

5.2	Normalised SFB autocorrelator at the fiducial cosmological model used to generate the halo mock catalogues. [Reprinted under the Policy]	92
5.3	Dimensionless angular and radial SFB coupling coefficients in the spherical Fourier analysis of halo mock catalogues. [Reprinted under the Policy]	93
5.4	Window function multipoles in configuration space for the Cartesian power spectrum analysis of halo mock catalogues. [Reproduced under the Policy] . . .	93
5.5	Estimated correlation matrices of the power spectrum monopole for the Cartesian power spectrum analysis of halo mock catalogues. [Reprinted under the Policy]	94
5.6	Proxy condition number of the covariance matrix with data compression in the spherical Fourier analysis of halo mock catalogues. [Reproduced under the Policy]	95
5.7	Absolute-value estimates of the cross-correlation between spherical- and Cartesian-basis data in the clustering analysis of halo mock catalogues. [Reprinted under the Policy]	96
5.8	Joint posterior distribution of f_{NL} and b_1 from the hybrid-basis likelihood, including its low- k and high- k components based on the spherical- and Cartesian-basis data, in the Fourier analysis of halo mock catalogues. [Reprinted under the Policy]	97
5.9	Joint posterior distribution of f_{NL} and b_1 from the hybrid-basis and Cartesian power spectrum likelihoods for a subset of halo mock catalogues in the full-sky set-up.	97
5.10(a)	Full-sky $f_{\text{NL}}-b_1$ constraints from the hybrid-basis and Cartesian power spectrum likelihood analyses of halo mock catalogues. [Reprinted under the Policy]	99
5.10(b)	Partial-sky $f_{\text{NL}}-b_1$ constraints from the hybrid-basis and Cartesian power spectrum likelihood analyses of halo mock catalogues. [Reprinted under the Policy]	99
5.11(a)	Comparison of inferred models of the window-convolved power spectrum monopole in the hybrid-basis and Cartesian power spectrum analyses of halo mock catalogues in the full-sky set-up. [Reprinted under the Policy]	100
5.11(b)	Comparison of inferred models of the window-convolved power spectrum monopole in the hybrid-basis and Cartesian power spectrum analyses of halo mock catalogues in the partial-sky set-up. [Reprinted under the Policy]	100

Abbreviations

This is an exhaustive list of scientific abbreviations, including initialisms and acronyms, appearing in this thesis. Wherever applicable, an up-to-date web address is provided.

2LPT	second-order Lagrangian perturbation theory
2LPTic	2LPT Initial Conditions (↗ cosmo.nyu.edu/roman/2LPT)
2MASS	2-Micron All-Sky Redshift Survey (↗ old.ipac.caltech.edu/2mass)
2dFGRS	2-degree Field Galaxy Redshift Survey (↗ 2dfgrs.net)
6dFGRS	6-degree Field Galaxy Redshift Survey (↗ 6dfgs.net)
AHF	AMIGA Halo Finder (↗ popia.ft.uam.es/AHF)
AMIGA	Adaptive Mesh Investigations of Galaxy Assembly (↗ popia.ft.uam.es/AMIGA)
AP	Alcock–Paczyński (test/effect)
BAO	baryon acoustic oscillation
BBN	Big Bang nucleosynthesis
BOSS	Baryon Oscillation Spectroscopic Survey (↗ sdss.org/surveys/boss)
CAMB	Code for Anisotropies in the Microwave Background (↗ camb.info)
CDM	cold dark matter
CfA	Center for Astrophysics (Redshift Survey) (↗ tdc-harvard.edu/zcat)
CLASS	Cosmic Linear Anisotropy Solving System (↗ class-code.net)
CMASS	constant-mass (sample) (of BOSS)
CMB	cosmic microwave background
COBE	<i>Cosmic Background Explorer</i> (↗ lambda.gsfc.nasa.gov/product/cobe/)
DE	dark energy
DES	Dark Energy Survey (↗ darkenergysurvey.org)
DESI	Dark Energy Spectroscopic Instrument (↗ desi.lbl.gov)
DM	dark matter
eBOSS	extended BOSS (↗ sdss.org/surveys/eboss)
ELG	emission line galaxy
Euclid	<i>Euclid</i> (mission/satellite) (↗ euclid-ec.org)

FFT	fast Fourier transform
FFTL_{og}	FFT of a discrete periodic sequence of logarithmically spaced points (<a href="http://jila.colorado.edu/~ajsh/FFTL<sub>og</sub>">↗ jila.colorado.edu/~ajsh/FFTL_{og})
FKP	Feldman–Kaiser–Peacock (formalism/weight)
FLRW	Friedmann–Lemaître–Robertson–Walker (metric/spacetime)
GADGET	GAxaxies with Dark matter and Gas intEracT (↗ wwwmpa.mpa-garching.mpg.de/gadget)
GR	general relativity
GRF	Gaussian random field
HEALPix	Hierarchical Equal Area isoLatitude Pixelization (↗ github.com/healpy/healpy)
IRAS	<i>Infrared Astronomical Satellite</i> (↗ irsa.ipac.caltech.edu/Missions/iras.html)
ISW	integrated SW (effect)
KL	Kullback–Leibler (divergence)
ΛCDM	Lambda cold dark matter (model)
LF	luminosity function
LRG	luminous red galaxy
LSS	large-scale structure
Ly-α	Lyman-alpha (forest/line)
MCFIT	Multiplicatively Convolutional Fast Integral Transforms (↗ github.com/eelregit/mcfit)
MCMC	Markov chain Monte Carlo
NGC	North Galactic Cap
PDF	probability density function/functional
Planck	<i>Planck</i> (mission/satellite) (↗ sci.esa.int/web/planck)
PLE	pure luminosity evolution (model)
PNG	primordial non-Gaussianity
PSCz	Point Source Catalog Redshift (Survey) (of <i>IRAS</i>)
Q–Q	quantile–quantile (plot)
QSO	quasi-stellar object/quasar
RSD	redshift-space distortion
SDSS	Sloan Digital Sky Survey (↗ sdss.org)
SFB	spherical Fourier–Bessel (transform/basis/coefficients)
SN	supernova
SPHEREx	Spectro-Photometer for the History of the Universe, Epoch of Reionization, and Ices Explorer (↗ spherex.caltech.edu)
SW	Sachs–Wolfe (effect)
TNS	Taruya–Nishimichi–Saito (model)
UMF	universal mass function
WMAP	<i>Wilkinson Microwave Anisotropy Probe</i> (↗ wmap.gsfc.nasa.gov)

Notations

This is an exhaustive list of notations appearing in this thesis, most of which have broadly consistent usage in the literature; however, universal mathematical notations and notations locally defined, such as placeholder variables, may not be included. Note that the bullet point ‘ \bullet ’ is a placeholder for a generic quantity, and a caron ‘ \smile ’ (not to be confused with a breve ‘ \breve ’) is sometimes used to denote a transformed quantity. Some symbols may have more than one usage, which can be inferred unambiguously from the context.

$\langle \bullet \rangle$	(angle brackets/delimited Guillemets for) ensemble expectation
$\breve{\bullet}$	(a breve for) value in a fiducial cosmological model
$\hat{\bullet}$	(a wide hat/circumflex for) estimator/estimate for a quantity (only used when it needs to be distinguished from the quantity itself)
$\dot{\bullet}$	(an overdot for) cosmic-time derivative ($\partial \bullet / \partial t$)
$\bar{\bullet}$	(an overbar/macron for) background/mean/averaged quantity
$\tilde{\bullet}$	(a tilde for) filtered/windowed quantity
\bullet_0	(a subscript ‘0’ for) value at the current redshift $z = 0$
\bullet_s	(a subscript ‘s’ for) redshift-space quantity
$\delta \bullet$	(a letter ‘ δ ’ for) perturbation quantity
\bullet'	(a thin prime for) derivative of a function with respect to its argument
\bullet'	(a thick prime for) conformal-time derivative ($\partial \bullet / \partial \tau$)
a	scale factor
A_s	primordial scalar power spectrum amplitude
α	ratio of the weighted number count of the survey galaxy catalogue to that of the synthetic random catalogue
α_{AP}	AP dilation parameter
α_{\parallel}	line-of-sight AP scaling parameter
α_{\perp}	transverse AP scaling parameter
B	bispectrum
b	(Eulerian) bias
b_1	linear/scale-independent (Eulerian) bias
b_e	evolution bias
$b_1^{(\text{L})}$	linear Lagrangian bias

c	speed of light in vacuum
c_s	speed of sound
\mathbf{d}	local/pairwise line of sight
D_1	linear growth factor
D_A	angular diameter distance
D_C	radial/line-of-sight comoving distance
D_H	Hubble distance
$D_{\text{KL}}(f \parallel g)$	Kullback–Leibler divergence of the PDF g from the PDF f
D_L	luminosity distance
D_M	transverse comoving distance
D_μ	SFB clustering mode
D_σ	fingers-of-God damping factor
Δ	comoving density contrast
δ	over-density
δ_{ab}	Kronecker delta tensor
δ_c	(spherical-collapse) critical over-density
$\delta^{(D)}$	Dirac delta distribution
$\delta^{(K)}$	Kronecker delta function
δ_m	matter density contrast
\mathbb{E}	expectation
ϵ_{AP}	AP deformation parameter
erfc	complementary error function
f	linear growth rate
${}_2F_1$	ordinary hypergeometric function
f_{NL}	(local-type) primordial non-Gaussianity parameter
f_{sky}	(effective) survey sky fraction
Φ	Bardeen potential
ϕ	radial selection function/luminosity function
φ	spherical azimuthal angle
G	gravitational constant
g_{ab}	metric tensor
G_{ab}	Einstein tensor/trace-reversed Ricci curvature tensor
Γ	gamma function
Γ_d	multivariate gamma function
H	(physical) Hubble parameter
h	reduced Hubble parameter of today
\mathcal{H}	conformal Hubble parameter
η_{ab}	Minkowski metric tensor
\mathbf{J}	Jacobian matrix
j_ℓ	spherical Bessel function of the first kind of order ℓ
K	spatial curvature
\mathbf{k}	Fourier-space coordinates/comoving wavevector

κ	condition number
\mathcal{L}	likelihood function
\mathcal{L}_ℓ	Legendre polynomial of degree ℓ
Λ	cosmological constant
M	halo mass
m	absolute magnitude
\bar{m}	absolute-magnitude threshold
μ	cosine of the angle between a wavevector and the line of sight
n	number density
\bar{n}	background/mean number density
n_h	halo mass function/UMF
\hat{n}	global line of sight
n_s	primordial scalar spectral index
ν	peak height/significance (with respect to field variance)
Ω_I	density parameter for matter–energy component I
P	power spectrum/pressure
P_ℓ	power spectrum multipole of degree ℓ
P_m	matter power spectrum
P_N	N th-order polyspectrum
p	tracer-dependent parameter for scale-dependent bias due to PNG (see eq. 1.60)
P_{shot}	shot noise power
π	prior (probability density)
\mathcal{P}	posterior (probability density)
\mathbb{P}	probability
\mathbb{P}	probability density
Ψ	Bardeen potential
R	Ricci curvature scalar
\mathbf{r}	configuration-space coordinates/real-space comoving coordinates/comoving separation vector
R_{ab}	Ricci curvature tensor
ρ	density
ρ_c	critical density
S	bispectrum shape function
\mathbf{s}	redshift-space comoving coordinates/redshift-space comoving separation vector
s	magnification bias (not to be confused with redshift-space separation $s = \mathbf{s} $)
σ_8	root-mean-square linear density fluctuation within $8 h^{-1}$ Mpc radius
σ_p	pairwise line-of-sight velocity dispersion
σ_v	one-dimensional velocity dispersion
T	transfer function

t	cosmic time
T_{ab}	energy–momentum tensor
τ	conformal time
Θ	Heaviside function
θ	(set of) (cosmological) model parameter(s)
ϑ	spherical polar angle
u	$\mathcal{H}^{-1}v$, comoving radial RSD displacement
\mathbf{v}	peculiar velocity
Var	variance
W	window function
w	(clustering) weight/equation-of-state parameter
ξ	correlation function
ξ_N	N -point correlator
$Y_{\ell m}$	spherical harmonic function of degree ℓ and order m
z	redshift
ζ	uniform-density curvature perturbation
ζ_N	connected N -point correlator



Preface

This thesis is the culmination of both independent and collaborative research efforts undertaken during my doctoral programme.



Related works

The methodology and findings in this thesis have mostly been disseminated in earlier peer-reviewed publications of mine, which are listed below in chronological order:

- [1] **MS Wang**, WJ Percival, S Avila, R Crittenden and D Bianchi, 2019. ‘Cosmological inference from galaxy-clustering power spectrum: Gaussianization and covariance decomposition’, *Mon. Not. Roy. Astron. Soc.* **486** (1), 951 [1811.08155] [Paper I] (incorporated into chapter 3);
- [2] **MS Wang**, F Beutler and D Bacon, 2020. ‘Impact of relativistic effects on the primordial non-Gaussianity signature in the large-scale clustering of quasars’, *Mon. Not. Roy. Astron. Soc.* **499** (2), 2598 [2007.01802] [Paper II] (incorporated into chapter 4);
- [3] **MS Wang**, S Avila, D Bianchi, R Crittenden and WJ Percival, 2020. ‘Hybrid-basis inference for large-scale galaxy clustering: combining spherical and Cartesian Fourier analyses’, *J. Cosmol. Astropart. Phys.* **2020** (10), 022 [2007.14962] [Paper III] (incorporated into chapter 5).

All data and derived data products underlying this thesis and earlier related works can be found in the publications above or made available upon reasonable request. My academic authorship information can be found using either the ORCID record  0000-0002-2652-4043 or the arXiv public author identifier  arxiv.org/a/wang_m_4.

In the course of my doctoral research, two open-source software packages have been developed and released:

- 1) **HARMONIA** ( MikeSWang/Harmonia), which offers an analysis pipeline for combining spherical and Cartesian Fourier analyses of large-scale galaxy clustering measurements [ascl:2009.022];
- 2) **HORIZONGROUND** ( MikeSWang/HorizonGRound), which derives relativistic corrections to the clustering power spectrum, including uncertainties, from measurements of the tracer luminosity function [ascl:2008.027].

More details about these packages can be found in the linked repositories above and in the corresponding publications.

Acknowledgements

Private communications and software packages specific to individual research projects that constitute this thesis are separately acknowledged in the aforementioned publications. Some figures make use of the following Python packages: NUMPY [4], SCIPY [5], ASTROPY [6, 7], NBODYKIT [8] and MATPLOTLIB [9]. I am funded through the postgraduate bursary scheme by the Faculty of Technology at the University of Portsmouth. Numerical computations in this thesis are performed on the Sciama High Performance Compute cluster which is supported by the Institute of Cosmology and Gravitation (ICG), the South East Physics Network (SEPnet) and the University of Portsmouth.

I am very much indebted to my doctoral advisers **Prof Robert Crittenden**, **Prof Will Percival**, **Dr Santiago Avila** and **Dr Davide Bianchi**. It has been a privilege to have such an outstanding supervisory team who have offered me much of their expertise, wisdom, understanding and encouragement. I am grateful to Will, Davide and Santi for continuing to work closely with me after moving away from the UK, and to Rob for stepping in to guide me. I am also indebted to my collaborators **Dr Florian Beutler** and **Prof David Bacon** who have helped broaden my research scope and given me plenty of their time and support. I would also like to thank the **administrative and computing staff** at the ICG for their dedication to work and care for others. It is thanks to all of them that my doctoral programme has been an overwhelmingly positive journey.

It is perhaps only natural that moments at one's crossroads in life are oft imbued with poignance. I express my gratitude here to **friends and colleagues**, whose names are listed separately at the end of this thesis, for having gifted me many a fond memory in this island city; by extension, I am grateful to all **friends and 'family'** over the last ten years in the UK. Finally, to **my parents** who have for me made so much sacrifice, I am sorry—for I seem destined to live so far away from you in much of my life. It is fitting to end here with the words from *Odes 1.11* by Horace—

*Tu ne quaesieris, scire nefas, quem mihi, quem tibi finem di dederint, Leuconæ, nec
Babylonios temptaris numeros. ut melius, quidquid erit, pati. seu pluris hiemes seu
tribuit Iuppiter ultimam... dum loquimur, fugerit invida atas: carpe diem, quam
minimum credula postero.*

Mike (Shèngbó) Wáng (王圣博)
Portsmouth, Winter 2020

Theoretical Considerations
in
Large-Scale Structure Analyses



Introduction

The study of the large-scale structure (LSS) of the Universe is a firm cornerstone of modern cosmology. By recording the angular position of galaxies in the sky and measuring the redshift of their electromagnetic spectrum, one could build a three-dimensional map of cosmic structure on the largest accessible scales. Looking back to a century ago, the beginning of this field seemed rather humble, when the phrase ‘the large-scale distribution of matter’ was understood merely in the context of the Milky Way galaxy and its stars [10]. Yet to date, millions of galaxies spanning over billions of years of cosmic history have been observed, notably by the recently completed Baryon Oscillation Spectroscopic Survey (BOSS) and its extension eBOSS, both of which are part of the Sloan Digital Sky Survey (SDSS) [11]. The rapid technological advances in observational capabilities and ongoing developments in the theoretical foundation have led some cosmologists to refer to the current period as ‘the golden age of cosmology’ [12].

The objective of LSS cosmology, however, is not to simply catalogue all the observable galaxies, for that might lead to accusations of ‘stamp collecting’.¹ Instead, with the Universe itself being the ultimate laboratory, its quest is to solve, in some sense, an initial value problem: What were the *initial conditions*, most likely of a quantum nature, for structure formation? With gravity being the dominant long-range fundamental force, what are the *physical processes* governing structure formation?

To contextualise LSS within modern cosmology, one could not fail to connect it to the so-called ‘standard model’, known as Lambda cold dark matter (Λ CDM), which assumes Einstein’s theory of general relativity (GR) with the matter–energy content of the Universe at present divided into three major components: dark energy (DE), which is associated with the cosmological constant Λ , occupies around 70 %; cold dark matter accounts for about 25 %; and ordinary baryonic matter only 5 % [14].² Other components include a negligible amount of radiation and particle species such as neutrinos. The cosmological principle, which states that the spatial distribution of matter–energy is statistically homogeneous and isotropic on

¹ ☞ “All science is either physics or stamp collecting.” A quote attributed to Ernest Rutherford (and often misattributed to William Thomson, Lord Kelvin) [13].

² Strictly speaking, baryonic matter is solely composed of quarks. In astronomy, the term is used loosely to describe normal atomic matter.

large scales, also underpins this standard model; under these assumptions, the background space-time is described by the Friedmann–Lemaître–Robertson–Walker (FLRW) metric, allowing for a universe with an evolving physical size as well as non-zero spatial curvature [15–19]. Moreover, the theory of inflation, which postulates a brief period of exponential expansion in the very early Universe [20], is commonly incorporated into the standard model, as it offers a plausible mechanism for seeding the observed classical, macroscopic structure with quantum fluctuations.

Astronomical experiments in the past few decades have been undeniably fruitful in producing evidence supporting the Λ CDM model, whose predictions give a remarkably concordant account, with a minimal set of just six cosmological parameters, for some key observations:

- the existence and properties of the cosmic microwave background (CMB) [21, 22], which exhibits an exquisite black-body spectrum, as confirmed and measured by multiple generations of space missions including the *Cosmic Background Explorer (COBE)*, the *Wilkinson Microwave Anisotropy Probe (WMAP)* and *Planck*;
- the large-scale distribution of galaxies (i.e. LSS), with a preferred clustering scale at the wavelength of baryon acoustic oscillations (BAOs) at the photon–baryon decoupling epoch, as first detected by SDSS and the 2-degree Field Galaxy Redshift Survey (2dFGRS) [23, 24];³
- the accelerating recession of distant galaxies as measured with standardisable Type Ia supernovæ (SNe) [26, 27]; and
- the abundance of light elements produced in the primordial Universe through Big Bang nucleosynthesis (BBN) [28–31].

In particular, LSS clustering observations offer strong evidence, both complementary to and independent of CMB anisotropy and polarisation as well as SNe, for the existence of dark energy and dark matter [23, 32, 33]. They also paint a picture of dominantly adiabatic, Gaussian initial fluctuations which favour an inflationary origin of cosmic structure passively formed through gravitational instability, as opposed to an active, causal mechanism [34–36]. In combination with other probes, the use of BAOs as a ‘standard ruler’ as well as galaxy–void cross-correlation can also break degeneracy amongst Λ CDM model parameters and thus drastically improve cosmological constraints [37]. In Figure 1.1, an example of cosmic concordance amongst different probes is shown for the joint constraints on the dark energy and matter density parameters, which are discussed in detail in appendix A.

Yet with the arrival of precision cosmology that has achieved model constraints near the 1% level, discordance has recently become more apparent and tantalising, as demonstrated by current tensions in the Hubble parameter H_0 and the matter density fluctuation parameter σ_8 [see e.g. 38–40]. This all adds to the existing unsolved mysteries in the Λ CDM paradigm, which perhaps could be categorised, with the view of LSS as an *initial value problem*, as follows—

- initial conditions: How close are primordial fluctuations to being Gaussian? What type of primordial non-Gaussianity (PNG) could precision LSS measurements reveal? PNG is crucial in differentiating models of the very early Universe and probing high-energy physics at scales forever beyond the reach of terrestrial experiments.

³ There were hints in earlier results reported by 2dFGRS [25].

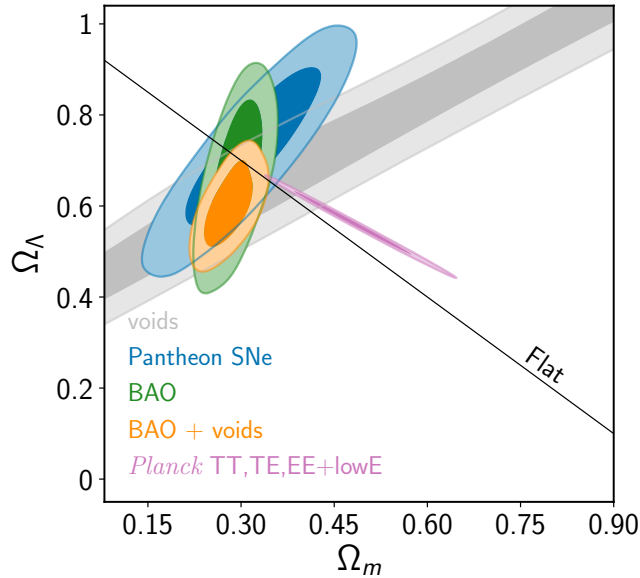


Figure 1.1. Cosmic concordance: marginalised constraints on the matter and dark energy density parameters, Ω_m and Ω_Λ , at the current epoch. Each set of contours shows the 68 % and 95 % credible regions from the following cosmological probes: *Planck* temperature and *E*-mode polarisation (purple contours); Pantheon SNe (blue contours); BOSS galaxy BAO (green contours), voids (grey contours), and the combination of the two (orange contours). Figure taken from Fig. 4 in ref. [37] (credit: S Nadathur et al.).

- matter–energy contents and space-time geometry: What is the material nature of dark energy or dark matter, and is dark energy dynamical or really a cosmological constant? How strong is the evidence for, or against, the cosmological principle? What more could be learnt about exotic particle species, including massive neutrinos, and their contributions to structure formation?
- gravitational evolution: Are there any discernible deviations from Einstein’s GR on cosmological scales? Could modified gravity theories supplant the need for dark energy, and how do they alter the evolution of cosmic structure?

In the so-called ‘big data’ era, LSS possesses some unique advantages in helping answer some of these questions: its data sets are intrinsically three-dimensional, in contrast to the case of CMB which is a two-dimensional surface; they also span a wide redshift range, with a pool of various tracers including low-redshift galaxies, luminous red galaxies (LRGs), emission line galaxies (ELGs), the Lyman-alpha ($\text{Ly-}\alpha$) forest and quasi-stellar objects (QSOs), also known as quasars. With CMB measurements gradually saturating the cosmic variance bound, LSS offers some powerful probes such as the Alcock–Paczyński (AP) test [41], closely related to the BAO measurements, and redshift-space distortions (RSDs) [42], which arise as the peculiar motion of galaxies affects their apparent distance converted from spectral redshift from the observer. In particular, the BAO and AP probes yield cosmological distance measurements, which trace the expansion history of the late-time Universe and place constraints on cosmological parameters; redshift-space distortions, a seemingly unwelcome observational effect at first glance, actually provide an indispensable tool for measuring the growth of structure and constraining the matter density parameter [42, 43], and thus qualify as an important test of gravity [see e.g. 44–46].



Having given a broad overview of LSS and made emphatic arguments for its rôle in modern cosmology, this introductory chapter will now lay down some theoretical background for the focus of this thesis: anisotropic galaxy clustering on very large scales and the statistical inference of cosmological parameters. The following sections will discuss the statistics of cosmic random fields, anisotropies in the observed galaxy distribution, the LSS signature of primordial non-Gaussianity, and lastly, parameter estimation from cosmological likelihoods.

Probability and statistical theory, linear algebra and mathematical methods in general are prerequisites for this thesis, and some familiarity with quantum field theory is helpful but not necessary. The reader is also assumed to have basic knowledge of general relativity, especially in the context of the FLRW space-time; however, many results quoted in the text can be referred to in the appendices, which contain more detailed discussions and derivations.

1.1 Statistics of cosmic random fields

Although cosmologists are privileged to have the cosmos as the ultimate laboratory, there is inherent stochasticity in the observations of our Universe, which only arises as a particular realisation in an infinite ensemble of possibilities due to the probabilistic nature of its quantum origin. This means that cosmological measurements are susceptible to *cosmic variance* and a statistical interpretation is essential, no matter how well-designed the experiment might be.⁴ At the same time, there is a growing demand for sophisticated statistical methods for data compression and transformation, as large data sets now steer the progress of modern precision cosmology.

Cosmological observables are usually modelled as (three-dimensional) *random fields*, a map that assigns a random variable to each point \mathbf{r} in configuration space.⁵ Catalogues of galaxies gradually built up in the last century have shown that their distribution is not entirely random in the Universe, but instead it weaves out an intricate web of interconnected structures with a pattern of over- and under-densities such as haloes and voids. Therefore a key cosmological observable in LSS is the *over-density* field δ for the galaxy number density n ,

$$\delta(\mathbf{r}) = \frac{n(\mathbf{r}) - \bar{n}}{\bar{n}}, \quad (1.1)$$

where the *background* number density $\bar{n} = \langle n \rangle$ is the *ensemble expectation* $\langle \cdot \rangle$ of the number density. By the cosmological principle, \bar{n} is spatially invariant; and by definition, the expectation of the over-density field is identically zero, $\langle \delta \rangle = 0$.

Gaussian random fields. Since the simplest inflationary models predict Gaussian initial perturbations, which under linear evolution preserve their Gaussian statistics, cosmic random

⁴ ☞ “If your experiment needs statistics, you ought to have done a better experiment.” A quote again attributed to Ernest Rutherford with a few variants, none of which though has a verifiable source.

⁵ A random variable is not a variable per se, but a map from an event (a set of outcomes) to a value (usually the real numbers). This technical distinction is not made in this thesis for simplicity, and a ‘random variable’ also refers to the value it takes.

fields are often treated as Gaussian random fields (GRFs) on large, linear scales [47, 48]. For any number N of points at $\{\mathbf{r}_i\}_{i=1}^N$ in a GRF δ , the distribution of $(\delta(\mathbf{r}_1), \dots, \delta(\mathbf{r}_N))$ is always multivariate normal.⁶ The field itself has the following probability density (PDF),

$$\mathbb{p}[\delta] = A^{-1/2} \exp\left[-\frac{1}{2} \int d^3\mathbf{r}_1 d^3\mathbf{r}_2 \delta(\mathbf{r}_1) \xi(\mathbf{r}_1, \mathbf{r}_2)^{-1} \delta(\mathbf{r}_2)\right], \quad (1.2)$$

where ξ is the *correlation function*,

$$\xi(\mathbf{r}_1, \mathbf{r}_2) = \langle \delta(\mathbf{r}_1) \delta(\mathbf{r}_2) \rangle, \quad (1.3)$$

and the normalisation constant is given by

$$A^{1/2} = \int D\delta \exp\left[-\frac{1}{2} \int d^3\mathbf{r}_1 d^3\mathbf{r}_2 \delta(\mathbf{r}_1) \xi(\mathbf{r}_1, \mathbf{r}_2)^{-1} \delta(\mathbf{r}_2)\right]. \quad (1.4)$$

Here $D\delta = \lim_{N \rightarrow \infty} \prod_{i=1}^N d\delta(\mathbf{r}_i) / (2\pi)^N$ is the Wiener measure, and A can be regarded as the determinant of ξ^{-1} as an operator over the field configurations of δ [49].

Correlators. The notion of a correlation function can be generalised to the *N -point correlator* by considering higher-order moments of a random field δ ,

$$\xi_N(\mathbf{r}_1, \dots, \mathbf{r}_N) = \langle \delta(\mathbf{r}_1) \cdots \delta(\mathbf{r}_N) \rangle, \quad (1.5)$$

which can be decomposed into a *connected*, or *irreducible*, component ζ_N (with $\langle \cdot \rangle$ labelled by a subscript ‘c’) and lower-order moments,

$$\xi_N(\mathbf{r}_1, \dots, \mathbf{r}_N) = \underbrace{\langle \delta(\mathbf{r}_1) \cdots \delta(\mathbf{r}_N) \rangle}_c + \sum_{\mathcal{S}_N} \prod_{S_j \in \mathcal{S}_N} \zeta_{|S_j|}. \quad (1.6)$$

Here the sum is over all non-trivial partitions \mathcal{S}_N of $\{\mathbf{r}_i\}_{i=1}^N$, and ξ_N or ζ_N is totally symmetric in its arguments [10, 48]. In Figure 1.2, a schematical representation of the decomposition for the case with $N = 4$ is shown. Sometimes, the N -point correlator is instead defined as $1 + \xi_N(\mathbf{r}_1, \dots, \mathbf{r}_N) = \langle \prod_{i=1}^N [1 + \delta(\mathbf{r}_i)] \rangle$, and ζ_N is known as the *reduced* N -point correlator [e.g. 47]. It is evident from the PDF (eq. 1.2) that the correlation function ξ completely specifies the probability distribution of a GRF δ ; indeed, any ζ_N vanishes identically, a result of Isserlis’ theorem [50].⁷ By expressing ξ in terms of the number densities, one finds the probability of having a pair of galaxies at \mathbf{r}_1 and \mathbf{r}_2 to be

$$\langle \delta N_1 \delta N_2 \rangle = [1 + \xi(\mathbf{r}_1, \mathbf{r}_2)] \bar{n}^2 \delta V_1 \delta V_2, \quad (1.7)$$

where $\delta N_{1,2} = n(\mathbf{r}_{1,2}) \delta V_{1,2}$ is the probability of having a galaxy in an infinitesimal volume $\delta V_{1,2}$ at $\mathbf{r}_{1,2}$. Therefore ξ could be interpreted as the *excess* probability of finding a pair of galaxies at two locations when there is *clustering*, compared to the case where galaxies are uniformly randomly distributed, i.e. $\langle \delta N_1 \delta N_2 \rangle = \bar{n}^2 \delta V_1 \delta V_2$ [10].

⁶ A subtle distinction is made between the words ‘Gaussian’ and ‘normal’ in this thesis: the former is used as a descriptive statistical property or refers to the shape of a function in its parameter space; and the latter is solely reserved for probability distributions.

⁷ Perhaps more commonly quoted in cosmology is Wick’s theorem, which is an analogous result proved for the normal ordering of quantum operators [51].

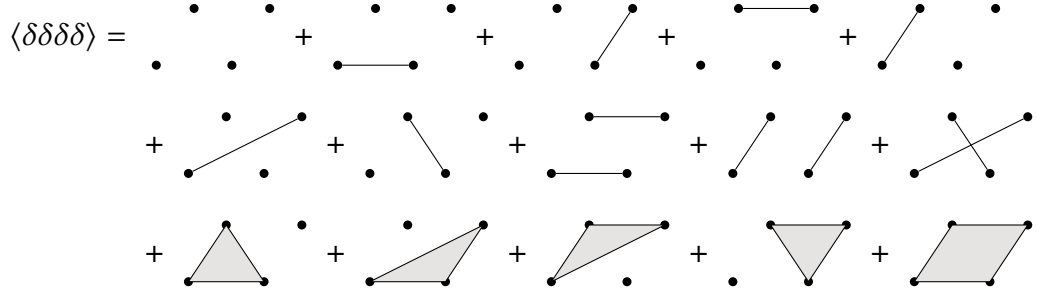


Figure 1.2. A schematic representation of the decomposition of an N -point correlator (in either configuration or Fourier space) into connected lower-order components for the case of $N = 4$ (see eq. 1.6). Each node represents the field at a given location; a connected component is a connected graph; and each summand is a product of connected graphs. Figure concept taken from Fig. 4 in ref. [48] (credit: F Bernardeau et al.).

Homogeneity and isotropy. If the cosmological principle applies to the random field δ , its statistical properties must satisfy homogeneity and isotropy. This means that any N -point correlator ξ_N must be invariant under spatial translation and rotation, which severely restricts its form. Though nominally a function of N coordinates, ξ_N depends on only $(N - 1)$ of them effectively; to see this, one could note that once one vertex of an N -gon is fixed by translation, its shape and size are then determined by just $(N - 1)$ sides. For the correlation function ξ , one would then find that it only depends a single scalar by rotational invariance,

$$\xi(\mathbf{r}_1, \mathbf{r}_2) \equiv \xi(\mathbf{r}_2 - \mathbf{r}_1) \equiv \xi(r_{12}), \quad (1.8)$$

where $r_{12} = |\mathbf{r}_2 - \mathbf{r}_1|$ is the pairwise separation. Analogous to equation (1.7), if one considers the conditional probability of finding a second galaxy given the first, then $(1 + \xi)$ could be interpreted as the probability density profile of galaxies around a given galaxy.

Fourier modes. Just as in music theory where a sound can be decomposed into tones at different pitches, cosmological perturbations such as the over-density field δ can also be decomposed into *Fourier modes* with different wavevectors \mathbf{k} by performing the Fourier transform,

$$\delta(\mathbf{k}) = \int d^3\mathbf{r} e^{-i\mathbf{k}\cdot\mathbf{r}} \delta(\mathbf{r}), \quad (1.9)$$

where the inverse mode wavenumber k^{-1} corresponds to the spatial scale of the fluctuation. Here the same notation δ is used to denote both the function $\delta(\mathbf{r})$ and its Fourier transform $\delta(\mathbf{k})$, and any ambiguity is dispelled by the context or the explicit argument. There are several advantages of a Fourier-space description of random fields, both mathematical and physical: for instance, the behaviour of the field at different scales can be decomposed and understood on a mode-by-mode basis, an idea related to the study of harmonic analysis; in relation to inflation and quantum field theory in general, the degrees of freedom of free fields decouple into simple harmonic oscillators in momentum space, which is the Fourier space dual to the position space. To recover the field value in configuration space, the inverse Fourier transform is used,

$$\delta(\mathbf{r}) = \int \frac{d^3\mathbf{k}}{(2\pi)^3} e^{i\mathbf{k}\cdot\mathbf{r}} \delta(\mathbf{k}). \quad (1.10)$$

In principle, the domain of integration in the Fourier transform above is the entire configuration space; in practice, however, one can only observe a patch of the Universe or run simulations in a finite volume. Therefore the periodic boundary condition is often imposed on a cubic volume V of side length L . The Fourier modes are then discretised and the inverse Fourier transform becomes a series expansion,

$$\delta(\mathbf{r}) = L^{-3} \sum_{\mathbf{k}} \delta(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{r}}, \quad (1.11)$$

where the summation is over all wavevectors \mathbf{k} whose components must be integer multiples of the fundamental wavenumber $k_f = 2\pi/L$. The set of plane waves $\{e^{i\mathbf{k}\cdot\mathbf{r}}\}_{\mathbf{k}}$ form an orthogonal basis in the field space; the Dirac delta distribution $\delta^{(D)}$ and the Kronecker delta function $\delta^{(K)}$ are represented by

$$\delta^{(D)}(\mathbf{r}) = L^{-3} \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}} \quad \text{and} \quad \delta_{\mathbf{k}_1\mathbf{k}_2}^{(K)} = L^{-3} \int_V d^3\mathbf{r} e^{-i(\mathbf{k}_1-\mathbf{k}_2)\cdot\mathbf{r}}. \quad (1.12)$$

Although the periodic boundary condition provides a convenient set-up especially for running N -body simulations, it must be emphasised that the Universe is not expected to be periodic in reality: indeed, periodic boundary conditions may introduce artefacts such as edge effects (e.g. the Gibbs phenomenon when there is a discontinuity at the boundary [52, 53]) as well as deviation from the inverse-square force law on large scales comparable to the period [e.g. 54]. Therefore the periodic volume should be set much larger than the scales of interest to mitigate these artefacts.

Polyspectra. Similar to correlation functions in configuration space, one could consider the N -point correlator $\langle\delta(\mathbf{k}_1) \cdots \delta(\mathbf{k}_N)\rangle$ in Fourier space. Since the Fourier transform is a linear integral transform, the Gaussianity of the field $\delta(\mathbf{r})$ also holds for its Fourier mode $\delta(\mathbf{k})$, and again by Isserlis' theorem, $\langle\delta(\mathbf{k}_1) \cdots \delta(\mathbf{k}_N)\rangle$ either vanishes identically for odd N or can be reduced into two-point correlators for even N . More significant is the constraint imposed by homogeneity: the wavevectors must obey $\sum_{i=1}^N \mathbf{k}_i = \mathbf{0}$ if $\langle\delta(\mathbf{k}_1) \cdots \delta(\mathbf{k}_N)\rangle$ is non-zero, which follows from the fact that the phase factor $\exp(-i\mathbf{a} \cdot \sum_{i=1}^N \mathbf{k}_i)$ introduced by any global translation, $\mathbf{r} \mapsto \mathbf{r} + \mathbf{a}$, must be unity. This means that in general,

$$\langle\delta(\mathbf{k}_1) \cdots \delta(\mathbf{k}_N)\rangle = (2\pi)^3 \delta^{(D)}(\mathbf{k}_1 + \cdots + \mathbf{k}_N) P_N(\mathbf{k}_1, \dots, \mathbf{k}_N), \quad (1.13)$$

where P_N is the N th-order *polyspectrum*. The *power spectrum* $P(\mathbf{k})$ for $N = 2$ is the Fourier transform of the correlation function $\xi(\mathbf{r})$,⁸

$$P(\mathbf{k}) = \int d^3\mathbf{r} e^{-i\mathbf{k}\cdot\mathbf{r}} \xi(\mathbf{r}), \quad \xi(\mathbf{r}) = \int \frac{d^3\mathbf{k}}{(2\pi)^3} e^{i\mathbf{k}\cdot\mathbf{r}} P(\mathbf{k}), \quad (1.14)$$

which follows from their definitions or the Wiener–Khinchin theorem [55, 56]. If the Fourier modes are discretised under the periodic boundary condition imposed on a cubic volume V , then

$$P(\mathbf{k}) = V^{-1} \langle\delta(\mathbf{k}) \delta^*(\mathbf{k})\rangle. \quad (1.15)$$

⁸ Note that despite the same notation, the coordinate variable \mathbf{r} in the correlation function ξ denotes the separation vector between a galaxy pair and not the configuration-space coordinates. Formally, this \mathbf{r} belongs to the affine space associated with the configuration space under homogeneity.

Either statistic, ξ or P , encodes all the information in a GRF. Under statistical isotropy, $P(\mathbf{k}) \equiv P(k)$ depends only on a wavenumber—this homogeneous and isotropic power spectrum is what an inflationary model typically predicts and specifies as the initial condition for the growth of perturbations during structure formation.

Ergodicity. The idealised notion of an ensemble of possible universes mentioned earlier alludes to a frequentists' view of probability where independent observations could be repeated, which is far removed from the reality of a single observable Universe. Thanks to the 'fair sample' hypothesis which assumes that cosmic random fields are *ergodic*, well-separated regions in the Universe unlikely to have had any causal contact could be regarded as independent samples of an ensemble [10, 47]. As a result, the ensemble expectation of a quantity F dependent on the random field δ ,

$$\langle F \rangle = \int \mathcal{D}\delta \mathbb{p}[\delta] F[\delta], \quad (1.16)$$

may be replaced by its volume average,

$$\bar{F} = \int_V d^3\mathbf{r} F[\delta(\mathbf{r})], \quad (1.17)$$

as long as the volume V provides a fair sample of the field δ . Fortunately, a homogeneous GRF possesses the ergodicity property if its power spectrum is continuous [57].

Point processes. To make the final connection between a continuous random field and the discrete galaxies one observes in the sky, a point process is needed. The basic model used in LSS is the *Poisson point process*, where the number of galaxies in any region follows the Poisson distribution and disjoint regions are mutually independent. For simplicity, one could consider an over-density field δ without intrinsic clustering and with mean number density \bar{n} . The number of galaxies N in any volume V has the Poisson PDF,

$$\mathbb{p}(N) = \frac{\bar{N}^N}{N!} e^{-\bar{N}} \quad \text{with} \quad \bar{N} = \bar{n}V. \quad (1.18)$$

Naïvely, one would expect the power spectrum for such a random field to be constant, given the absence of any preferred clustering scale. To prove this, one could note that the Fourier mode $\delta(\mathbf{k})$ for a discrete set of galaxies at $\{\mathbf{r}_i \in V\}_{i=1}^N$ is simply a sum of Dirac delta contributions,

$$\delta(\mathbf{k}) = \frac{1}{\bar{n}} \sum_{i=1}^N e^{-i\mathbf{k}\cdot\mathbf{r}_i} - V\delta_{\mathbf{k}\mathbf{0}}^{(K)}. \quad (1.19)$$

Since \mathbf{r}_i is uniformly randomly distributed for a homogeneous Poisson point process, it follows that for any wavevector $\mathbf{k} \neq \mathbf{0}$, the mode expectation is $\mathbb{E}[\delta(\mathbf{k})] = 0$ and its variance is $\text{Var}[\delta(\mathbf{k})] = N/\bar{n}^2$ from $\mathbb{E}[\sin^2(\mathbf{k}\cdot\mathbf{r}_i)] = \mathbb{E}[\cos^2(\mathbf{k}\cdot\mathbf{r}_i)] = 1/2$. The power spectrum is thus a constant,

$$P_{\text{shot}} = \frac{\text{var}[\delta(\mathbf{k})]}{V} = \frac{1}{\bar{n}} \quad (1.20)$$

⁹ Often in the literature, the discrete Fourier mode is defined as $\delta_{\mathbf{k}}$ which includes a volume normalisation factor so that the mode itself is dimensionless. In such a convention, $P(\mathbf{k}) = V\langle|\delta_{\mathbf{k}}|^2\rangle$, but P still has the physical dimension of volume.

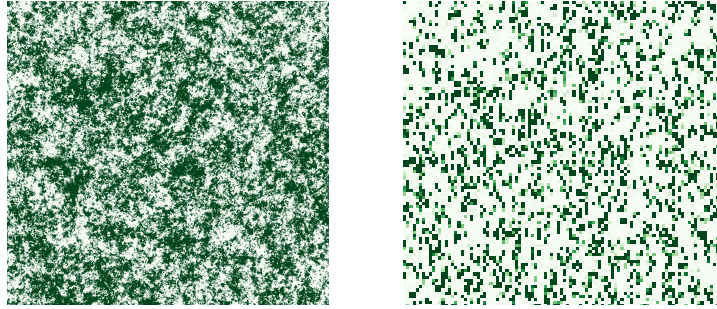


Figure 1.3. Poisson sampling of a continuous number density field (*left column*) produces a field of discrete tracers (*right column*). For homogeneous Poisson point processes, this adds a layer of constant shot noise, which is inversely proportional to the tracer number density, to the power spectrum.

— this is the Poisson *shot noise*. In general, for any point process that generates a discrete set of tracers from a continuous number density field, the shot noise power will add to the underlying clustering power spectrum. As a concrete example, suppose $n(\mathbf{r})$ is the number density of a tracer which has a constant bias b with respect to the underlying matter density field $\rho(\mathbf{r})$ (see appendix C for detailed discussion of tracer bias); that is to say, the tracer over-density field as defined by equation (1.1) is $\delta(\mathbf{r}) = b \delta_m(\mathbf{r}) \equiv b[\rho(\mathbf{r})/\bar{\rho} - 1]$, where δ_m is the matter density contrast and $\bar{\rho} = \langle \rho \rangle$ is the background matter density. In any small volume $\delta V \ni \mathbf{r}$, the tracer number count is $n(\mathbf{r}) \delta V$ and follows the Poisson distribution,

$$n(\mathbf{r}) \delta V \sim \text{Poisson}[\bar{n} \delta V (1 + \delta(\mathbf{r}))] \quad (1.21)$$

(see Fig. 1.3 for an illustration). The clustering power spectrum for this tracer is given by $P(\mathbf{k}) = b^2 P_m(\mathbf{k}) + \bar{n}^{-1}$, where P_m is the matter power spectrum (see also appendix B). It is interesting to point out that there is a different argument for the results above albeit in configuration space, where one imagines the volume V divided into a large number of microcells with a binary occupation number $N_{\text{cell}} = 0$ or 1. By exchanging discrete sums and continuous integrals and noting that $N_{\text{cell}}^2 = N_{\text{cell}}$, one arrives at

$$\langle n(\mathbf{r}_1)n(\mathbf{r}_2) \rangle = \bar{n}^2 [1 + \xi(\mathbf{r}_2 - \mathbf{r}_1)] + \bar{n} \delta^{(D)}(\mathbf{r}_1 - \mathbf{r}_2), \quad (1.22)$$

where the second term, the self-contribution from a single point, corresponds the shot noise power in Fourier space [10, 58]. The natural connection between the two arguments is that the Poisson distribution can be obtained as a large-number limit from the binomial distribution.



The rest of this thesis will rely on the basic assumption that the underlying large-scale distribution of tracers such as galaxies can be modelled as a discretely Poisson-sampled GRF satisfying homogeneity and isotropy, and the power spectrum will be the key observable used to probe the physics of galaxy clustering. Undoubtedly, there are numerous caveats to this assumption, some of which go beyond the scope of this thesis: the fact that $\delta(\mathbf{r}) \geq -1$ by definition, so the GRF model might be better replaced by, for instance, a log-normal model [e.g. 59]; non-Gaussianity from non-linear clustering [see e.g. 47, 48]; non-Poissonian shot noise due to halo exclusion and non-linear clustering [see e.g. 60–62]; and many other complications in both modelling and simulations.

1.2 Anisotropic galaxy clustering

In the idealised picture where the observer has access to the *exact* three-dimensional coordinates of galaxies in an instantaneous snapshot, clustering measurements could almost be directly compared with predictions of matter density fluctuations from cosmological models.¹⁰ In reality, the observer is passive and has to wait for distant photons to reach the telescope, and only then by measuring their redshift can a three-dimensional map of galaxies be constructed with their *inferred* coordinates in the observed volume. This map of detected galaxies does not depict a single snapshot but a past light cone, where matter flows towards the centres of gravitational wells amidst the expansion of the Universe; the measured large-scale distribution of galaxies, supposedly homogeneous and isotropic, would have been distorted by various physical effects that perturb the cosmological redshift of the photons as they traverse the cosmic web.

As a starting point, one could make the following connection between the galaxy overdensity field δ and the matter density contrast δ_m at the same redshift z ,

$$\delta(\mathbf{r}, z) = b(z) \delta_m(\mathbf{r}, z), \quad (1.23)$$

where the galaxy bias b is scale-independent on large, linear scales (see appendix B for the discussion of the evolution of matter density contrast δ_m). Since any detected galaxies lie on the past light cone from the observer, the comoving position \mathbf{r} and redshift z are not independent coordinates but are rather related through the comoving distance–redshift relation, $r = \chi(z)$ (see appendix A for discussions of the cosmological redshift and distances). Moreover, for incomplete galaxy survey observations due to instrumental sensitivity limit and sample selection effects, the homogeneous background number density $\bar{n}(z)$ is modulated by spatial variations, and the observable mean number density is

$$\bar{n}(\mathbf{r}, z) = W(\mathbf{r}) \bar{n}(z), \quad (1.24)$$

where W is the *survey window* function; for a full-sky survey with no angular variations, the window function is simply the *radial selection* function, $W(\mathbf{r}) = \phi(r)$. The observed number density field in the survey is thus

$$n(\mathbf{r}, z) = W(\mathbf{r}) \bar{n}(z) [1 + b(z) \delta_m(\mathbf{r}, z)] \quad (1.25)$$

following equation (1.1).

Apparent distortions in the observed galaxy distribution can arise when one converts a redshift measurement to a distance one: the cosmological redshift associated with the Hubble expansion may be altered by the peculiar motion of the source galaxy, an effect known as *redshift-space distortions*; the underlying cosmological model used to convert between the two, i.e. the fiducial distance–redshift relation, may not match the true cosmology, leading to the important *Alcock–Paczynski test*. These two effects, which are discussed below in detail, induce anisotropies around the line of sight in the measured galaxy distribution, since radial and angular distances are affected differently.

¹⁰ This is barring the issue of tracer bias—galaxies do not faithfully trace the underlying matter distribution, but instead have a tendency to form in local density peaks such as haloes and thus cluster differently.

1.2.1 Redshift-space distortions

For a galaxy that is moving relative to the background expansion of the Universe, its observed redshift z_{obs} differs from the cosmological redshift z owing to the Doppler effect,

$$1 + z_{\text{obs}} = (1 + z)(1 + \mathbf{v} \cdot \hat{\mathbf{r}}), \quad (1.26)$$

where \mathbf{v} is the peculiar velocity.¹¹ The observed *redshift-space* coordinates \mathbf{s} are mapped from the *real-space* coordinates \mathbf{r} along the radial direction by

$$s = \chi(z_{\text{obs}}) = r + u \quad \text{with} \quad u = \mathcal{H}^{-1}v \quad (1.27)$$

to linear order, where \mathcal{H} is the conformal Hubble parameter.

To obtain the relation between the redshift-space over-density and the real-space one, the local number conservation law $n_s(\mathbf{s}, z) d^3\mathbf{s} = n(\mathbf{r}, z) d^3\mathbf{r}$ is employed with equations (1.25) and (1.27):

$$\bar{n}_s(\mathbf{s}, z)[1 + \delta_s(\mathbf{s}, z)] = |\det \mathbf{J}_s(\mathbf{r})|^{-1} \bar{n}(\mathbf{r}, z)[1 + \delta(\mathbf{r}, z)], \quad (1.28)$$

where \cdot_s denotes a redshift-space quantity. Here \mathbf{J}_s is the Jacobian matrix for the real- to redshift-space coordinate mapping, the inverse absolute determinant of which can be expanded to linear order,

$$|\det \mathbf{J}_s(\mathbf{r})|^{-1} = \left(1 + \frac{u}{r}\right)^{-2} \left(1 + \frac{du}{dr}\right)^{-1} \simeq 1 - 2\frac{u}{r} - \frac{du}{dr}, \quad (1.29)$$

provided $|u| < r$, as $|du/dr| = O(\delta_m) < 1$ is already satisfied in linearised theory. Using the linearised continuity equation (B.21) and equation (1.23), one can then derive

$$\delta_s(\mathbf{s}, z) = [b(z) + f(z)\mathcal{D}_r] \delta_m(\mathbf{r}, z), \quad (1.30)$$

where f is the linear growth rate,

$$\mathcal{D}_r = \partial_r^2 \nabla^{-2} + \left[2 + \frac{\partial \ln \bar{n}(\mathbf{r}, z)}{\partial \ln r}\right] \frac{\partial_r}{r} \nabla^{-2} \quad (1.31)$$

is the RSD integro-differential operator, and $\delta_s(\mathbf{s}, z) = \delta_s(\mathbf{r}, z)$ in linearised theory [63].¹²

To calculate the Fourier modes of redshift-space galaxy clustering, one could simply Fourier transform equation (1.30) and use the inverse Fourier transform for the matter density contrast δ_m ,

$$\delta_s(\mathbf{k}, z) = b(z) \delta_m(\mathbf{k}, z) + f(z) \int \frac{d^3\mathbf{q}}{(2\pi)^3} R(\mathbf{q}, \mathbf{k}) \delta_m(\mathbf{q}, z), \quad (1.32)$$

where the Fourier kernel

$$R(\mathbf{q}, \mathbf{k}) = \int d^3\mathbf{r} e^{i(\mathbf{q}-\mathbf{k})\cdot\mathbf{r}} \left\{ (\hat{\mathbf{q}} \cdot \hat{\mathbf{r}})^2 - i \left[2 + \frac{\partial \ln \bar{n}(\mathbf{r}, z)}{\partial \ln r} \right] \frac{\hat{\mathbf{q}} \cdot \hat{\mathbf{r}}}{qr} \right\} \quad (1.33)$$

¹¹ Recall that the speed of light in vacuum is set to $c = 1$ in natural units, and here the Doppler effect is taken to be non-relativistic for $v \ll c$.

¹² On large scales, velocity bias between galaxies and dark matter is assumed to be negligible [63].

couples different clustering modes [64]. The complexity of this mode-coupling Fourier kernel poses computational challenges in modelling redshift-space clustering, but considerable simplification could be achieved if the *distant-observer limit*, $|u|/r \rightarrow 0$, holds—this condition translates to $v \ll \mathcal{H}r$, meaning the peculiar velocity is far slower than the Hubble flow, which is the case for distant galaxies.¹³ Consequently, the second part of the integro-differential operator \mathcal{D}_r proportional to $(\partial_r/r)\nabla^{-2}$, which corresponds to u/r , can be neglected, and

$$R(\mathbf{q}, \mathbf{k}) \rightarrow \int \frac{d^3\mathbf{r}}{(2\pi)^3} (\hat{\mathbf{q}} \cdot \hat{\mathbf{r}})^2 e^{i(\mathbf{q}-\mathbf{k})\cdot\mathbf{r}}. \quad (1.34)$$

Besides the distant-observer limit, the *global plane-parallel approximation* needs to be made to further simplify RSD modelling—if there exists a global line of sight $\hat{\mathbf{n}}$ such that $\mu_{\mathbf{k}} = \hat{\mathbf{k}} \cdot \hat{\mathbf{n}}$, the cosine of the angle between a mode wavevector and the line of sight, is constant, then

$$R(\mathbf{q}, \mathbf{k}) \rightarrow \mu_{\mathbf{k}}^2 \delta^{(D)}(\mathbf{q} - \mathbf{k}). \quad (1.35)$$

Therefore there is no longer a mode-coupling effect,

$$\delta_s(\mathbf{k}, z) = [b(z) + f(z)\mu_{\mathbf{k}}^2] \delta_m(\mathbf{k}, z), \quad (1.36)$$

and the redshift-space clustering power spectrum is simply

$$P_s(\mathbf{k}, z) = [b(z) + f(z)\mu_{\mathbf{k}}^2]^2 P_m(\mathbf{k}, z). \quad (1.37)$$

It is worth mentioning that a distinction is made here between the distant-observer and the global plane-parallel approximations, since they are two separate mathematical procedures. In some literature they are regarded as interchangeable, since for galaxy pairs at a given separation (in other words, number density fluctuations at a particular scale), the farther they are, the smaller the angular variations amongst the different lines of sight pointing towards them. However, this understanding could lead to confusion over the meaning of the so-called ‘wide-angle effect’, which is taken to mean the consequence of the plane-parallel approximation, rather than that of neglecting the mode-coupling RSD effect, when the galaxy survey has a large angular coverage and the line of sight could vary significantly. Although wide-angle effects will be discussed in more detail later, the reader is also referred to e.g. ref. [67] for clarification.

The results derived above in linearised theory are known as the Kaiser RSD model, first introduced in ref. [42];¹⁴ towards quasi-linear scales, this model no longer suffices, and in particular, there is a secondary RSD effect known as ‘the fingers of God’, which is not due to coherent peculiar motions under the influence of the local gravitational potential but rather the virialised random motions of galaxies. Figure 1.4 gives an illustration of both RSD effects. In Fourier space, one could introduce a phenomenological factor that accounts for the damping of the power spectrum due to the smearing of structure on smaller scales,

$$P_s(\mathbf{k}, z) = D_\sigma(k, \mu_{\mathbf{k}}) [b(z) + f(z)\mu_{\mathbf{k}}^2]^2 P_m(\mathbf{k}, z), \quad (1.38)$$

¹³ Assuming peculiar velocity $v \lesssim O(10^3) \text{ km s}^{-1}$ [65, 66], galaxies with redshift $z \gg 0.003$ are considered to be sufficiently distant.

¹⁴ Though the notion of redshift-space distortions and its use for probing the matter density parameter had been suggested even earlier by ref. [43].

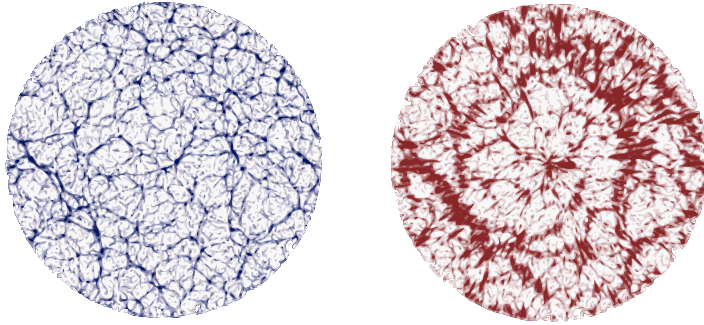


Figure 1.4. A simulated galaxy distribution in real space (*left column*) and in redshift space (*right column*) including the Kaiser effect and fingers of God. Filaments appear denser along directions transverse to the line of sight of the observer at the centre, and voids appear emptier due to the Kaiser effect. Fingers of God smear the galaxy distribution radially. Figure adapted from the original online source at [🔗 feldman.ku.edu](https://feldman.ku.edu) (credit: H Feldman).

where the damping factor D_σ depends on a velocity dispersion parameter σ and has several possibilities: for instance, the Gaussian (velocity distribution) model [68],

$$D_\sigma(k, \mu_k; \sigma_p) = \exp\left[-\frac{1}{2}(k\sigma_p\mu_k)^2\right]; \quad (1.39a)$$

the Lorentzian model (with exponential pairwise velocity distribution) [69–71],

$$D_\sigma(k, \mu_k; \sigma_p) = \left[1 + \frac{1}{2}(k\sigma_p\mu_k)^2\right]^{-1}; \quad (1.39b)$$

or the exponential (velocity distribution) model [72, 73],

$$D_\sigma(k, \mu_k; \sigma_v) = \left[1 + \frac{1}{2}(k\sigma_v\mu_k)^2\right]^{-2}. \quad (1.39c)$$

Here σ_p is the pairwise line-of-sight velocity dispersion and σ_v the one-dimensional velocity dispersion of a single galaxy,¹⁵ both of which have the dimension of inverse length after normalisation by the Hubble parameter and may have to be empirically determined in practice [47, 63]. These phenomenological forms are not actually physical, but could be formalised in streaming models where one considers the configuration-space over-density field convolved with a probability distribution for the random velocities of galaxies along the line-of-sight direction [see e.g. 74].

In recent years, many improved non-linear RSD models have been proposed (e.g. the Taruya–Nishimichi–Saito model [75, 76]), and the discussion above is far from exhaustive; for a recent detailed study, the reader may refer to e.g. ref. [77]. Regardless of the choice of models, it is clear that the presence of redshift-space distortions breaks any underlying homogeneity and isotropy of the galaxy distribution in observations: the former is lost as the observer is located at a special position; and the latter is reduced to only the rotational symmetry about the

¹⁵ There seems to be some confusion and misattribution in the literature between the Lorentzian and exponential models. Note also the difference in the overall power. If the velocity correlation between a galaxy pair is zero, then $\sigma_p = \sqrt{2}\sigma_v$ [70, 71].

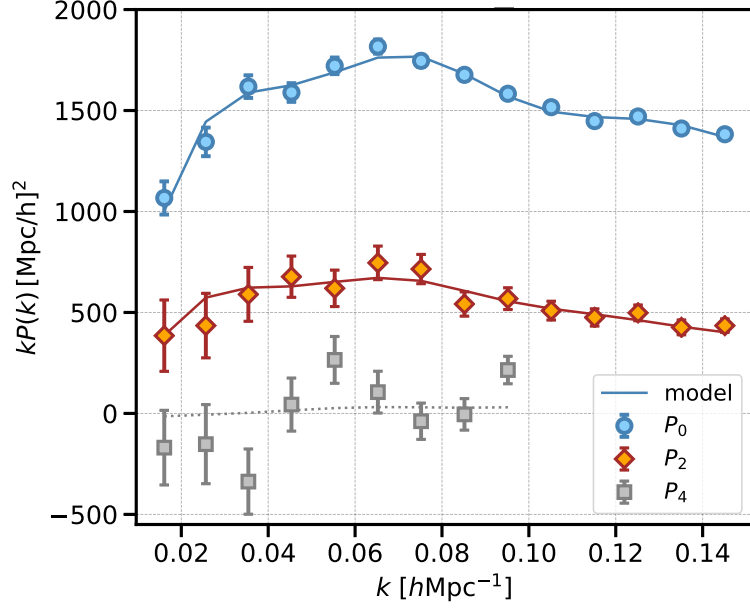


Figure 1.5. Measurements (data points with error bars) and the best-fitting model (solid and dotted lines) for the power spectrum multipoles $k P_\ell(k)$ of galaxies in the redshift range $0.4 < z < 0.6$ in BOSS Data Release 12 [79, 80]. Figure adapted from the original online source at ss.org/science/final-bao-and-rsd-measurements (credit: Jiamin Hou/BOSS Collaboration).

observer and azimuthal symmetry around the line of sight [78]. Generally, the redshift-space clustering power spectrum depends on two Fourier-space coordinates,

$$P_s(\mathbf{k}) \equiv P_s(k, \mu), \quad (1.40)$$

and thus may be decomposed into multipoles of the Legendre polynomials $\mathcal{L}_\ell(\mu)$,

$$P_\ell(k) = \frac{2\ell + 1}{2} \int_{-1}^1 d\mu \mathcal{L}_\ell(\mu) P_s(k, \mu), \quad (1.41)$$

where the redshift variable z and the subscript ‘ \mathbf{k} ’ for μ are omitted for brevity. The μ -dependence in the linear Kaiser RSD model means that only the first three even power spectrum multipoles—the monopole, quadrupole and hexadecapole—are non-zero,

$$P_0(k) = \left(b^2 + \frac{2}{3}bf + \frac{1}{5}f^2 \right) P_m(k), \quad (1.42a)$$

$$P_2(k) = \left(\frac{4}{3}bf + \frac{4}{7}f^2 \right) P_m(k), \quad (1.42b)$$

$$P_4(k) = \frac{8}{35}f^2 P_m(k). \quad (1.42c)$$

These multipoles have become the key summary statistics for anisotropic galaxy clustering analyses in recent surveys such as 2dFGRS, BOSS and eBOSS (see Fig. 1.5 for the BOSS Data Release 12 results as an example) [e.g. 24, 79, 81, 82].

1.2.2 The Alcock–Paczynski effect

The rescaling of radial and transverse distances by a fiducial distance–redshift relation, which does not necessarily match the true cosmology, is another cause for anisotropic distortions

in the observed galaxy distribution—this is the Alcock–Paczynski effect, which was originally proposed as an observational test for a non-zero cosmological constant in ref. [41]. From the comoving line element (eq. A.6a), it follows that the rescaling factors of line-of-sight and transverse comoving distances are given by

$$\alpha_{\parallel}(z)^{-1} = \frac{\check{\chi}'(z)}{\chi'(z)} \quad \text{and} \quad \alpha_{\perp}(z)^{-1} = \frac{\check{S}_k(z)}{S_k(z)} \quad (\text{I.43})$$

respectively, where a breve $\check{\cdot}$ denotes a quantity evaluated in a fiducial cosmological model and a thin prime \cdot' denotes the derivative of a function with respect to its argument. This is effectively a coordinate transformation,

$$\mathbf{r} \mapsto \check{\mathbf{r}} = \mathbf{A}\mathbf{r} \quad \text{and} \quad \mathbf{k} \mapsto \check{\mathbf{k}} = \mathbf{A}^{-1}\mathbf{k} \quad \text{with} \quad \mathbf{A} = \text{diag}(\alpha_{\parallel}^{-1}, \alpha_{\perp}^{-1}, \alpha_{\perp}^{-1}), \quad (\text{I.44})$$

where the coordinate vectors have been separated into parallel and transverse components with respect to the line of sight in the plane-parallel picture. The resulting three-dimensional power spectrum is

$$\check{P}(\check{k}_{\parallel}, \check{\mathbf{k}}_{\perp}) = |\mathbf{A}| P(\alpha_{\parallel}^{-1}\check{k}_{\parallel}, \alpha_{\perp}^{-1}\check{\mathbf{k}}_{\perp}), \quad (\text{I.45})$$

where $|\mathbf{A}|$ is the Jacobian determinant. Using equation (I.37) and

$$k_{\parallel} = k\mu, \quad k_{\perp} = k\sqrt{1 - \mu^2}, \quad \text{etc.}, \quad (\text{I.46})$$

one can derive the anisotropic power spectrum in the distorted coordinates,

$$\check{P}(\check{k}, \check{\mu}) = \alpha_{\parallel}^{-1} \alpha_{\perp}^{-2} P(k, \mu), \quad (\text{I.47})$$

where the true coordinates k and μ are related to the fiducial ones by

$$k = \check{k} \left[\alpha_{\perp}^{-2} + (\alpha_{\parallel}^{-2} - \alpha_{\perp}^{-2}) \check{\mu}^2 \right]^{1/2}, \quad (\text{I.48a})$$

$$\mu = \frac{\check{\mu}}{\alpha_{\parallel}} \left[\alpha_{\perp}^{-2} + (\alpha_{\parallel}^{-2} - \alpha_{\perp}^{-2}) \check{\mu}^2 \right]^{-1/2}. \quad (\text{I.48b})$$

The rescaling relations above were first derived by ref. [70], and nowadays it is common to rewrite the factors α_{\parallel} and α_{\perp} as

$$\alpha_{\parallel}(z) = \frac{\check{H}(z)}{H(z)} \quad \text{and} \quad \alpha_{\perp}(z) = \frac{D_{\text{M}}(z)}{\check{D}_{\text{M}}(z)}, \quad (\text{I.49})$$

since $d\chi/dz \propto H^{-1}$ and $S_k \propto D_{\text{M}}$ by equations (A.17a), (A.17b) and (A.18), where D_{M} is the transverse comoving distance. It is also equivalent to define α_{\parallel} and α_{\perp} using the Hubble distance D_{H} or the angular diameter distance D_{A} ; and for BAO analyses, these factors are sometimes anchored to the the sound horizon r_{d} of the photon–baryon fluid at the drag epoch $z = z_{\text{d}}$.¹⁶ This parametrisation anchored to the BAO scale set by r_{d} is helpful for breaking the degeneracy between the AP effect and RSD, since both give rise to anisotropies around the line of sight. In particular, the dilation and deformation parameters

$$\alpha_{\text{AP}} = (\alpha_{\parallel} \alpha_{\perp}^2)^{1/3} \quad \text{and} \quad \epsilon_{\text{AP}} = \left(\frac{\alpha_{\parallel}}{\alpha_{\perp}} \right)^{1/3} - 1 \quad (\text{I.50})$$

¹⁶ This refers to the Compton drag arising from the difference in the bulk velocities between photons and baryons. After the drag epoch, baryons are essentially released from photon pressure [83].

measure the stretching and warping of the BAO feature in the two-point correlators, and can in turn constrain cosmological distances [84].¹⁷



For the rest of this thesis, the linearised theory of RSD is assumed but not necessarily in the distant-observer limit with the plane-parallel approximation. Since the AP effect can be readily incorporated into anisotropic models of galaxy clustering via appropriate rescaling, it is neglected except where relevant.

1.3 Primordial non-Gaussianity signature

Although GRFs describe cosmic fluctuations in the observable Universe very well and form the basis of most cosmological analyses, a certain level of non-Gaussianity still exists, not only because of non-linear gravitational interactions but also due to primordial non-Gaussianity in the initial conditions of structure formation. In fact, the existence of PNG is predicted by inflationary theories including even the simplest single-field slow-roll model—this is due to the ‘gravitational floor’ imposed by minimal coupling of the inflaton field to gravity [85]—though the amplitude of PNG is generally expected to be small.

Whilst all GRFs are very similar, deviations from Gaussianity could take a multitude of forms.¹⁸ One way to categorise the different types of non-Gaussianity is by considering the triangular configuration of Fourier mode wavevectors, $(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)$ with $\sum_{i=1}^3 \mathbf{k}_i = \mathbf{0}$, for which the bispectrum $B(k_1, k_2, k_3)$,¹⁹ the leading-order non-Gaussian statistic, is enhanced. Common types of PNG include the squeezed, equilateral and orthogonal configurations, chiefly parametrised by f_{NL} (with appropriate labels for different types), which arise from different classes of inflationary models. For example, in multi-field models, a subdominant spectator field additional to the inflaton can generate curvature perturbations leading to $f_{\text{NL}}^{\text{local}} \gtrsim 1$, which corresponds to the squeezed configuration; in single-field models with non-standard kinetics, one might expect $f_{\text{NL}}^{\text{equil}} \gtrsim 1$ for the equilateral configuration; many models may have other distinctive features, e.g. a folded bispectrum or running f_{NL} [86, 87]. As such, PNG opens up a unique window onto the dynamics and interactions of fields in the very early Universe, and could help distinguish different classes of inflationary models. The most stringent constraints on some of the PNG parameters currently come from *Planck*, which has reported latest measurements of $f_{\text{NL}}^{\text{local}} = -0.9 \pm 5.1$, $f_{\text{NL}}^{\text{equil}} = -26 \pm 47$ and $f_{\text{NL}}^{\text{ortho}} = -38 \pm 24$ at the 68 % uncertainty level [88].

The local-type PNG $f_{\text{NL}}^{\text{local}}$, which is suppressed by the slow-roll parameters and thus almost negligible in the simplest single-field slow-roll models, is the focus of the discussion here, and thereafter the superscript ‘local’ shall be omitted for brevity. The motivation for f_{NL} comes from considering additional non-linear local terms of a Gaussian Bardeen potential Φ_{G} [89],

¹⁷ Sometimes the parameters ‘ D_V ’ and ‘ F_{AP} ’ are used, which are essentially the same as α_{AP} and ϵ_{AP} [see e.g. 81].

¹⁸ ☞ “All happy families are alike; each unhappy family is unhappy in its own way.” A (translated) quote from *Anna Karenina* by Leo Tolstoy.

¹⁹ The triangle can be completely specified by three side lengths, since the orientation of the triangle does not matter.

starting at the quadratic order,

$$\Phi(\mathbf{x}) = \Phi_G(\mathbf{x}) + f_{\text{NL}}(\Phi_G(\mathbf{x})^2 - \langle \Phi_G^2 \rangle) + O(\Phi_G^3). \quad (1.51)$$

The quadratic term becomes a convolution in Fourier space, and thus the three-point correlator for $\Phi(\mathbf{k})$, the bispectrum, is given by

$$\begin{aligned} & \langle \Phi(\mathbf{k}_1) \Phi(\mathbf{k}_2) \Phi(\mathbf{k}_3) \rangle \\ &= (2\pi)^3 f_{\text{NL}} \int \frac{d^3 \mathbf{q}_1}{(2\pi)^3} \frac{d^3 \mathbf{q}_2}{(2\pi)^3} \delta^{(D)}(\mathbf{q}_1 + \mathbf{q}_2 - \mathbf{k}_3) \left[\langle \Phi_G(\mathbf{k}_1) \Phi_G(\mathbf{k}_2) \Phi_G(\mathbf{q}_1) \Phi_G(\mathbf{q}_2) \rangle \right. \\ & \quad \left. - \langle \Phi_G(\mathbf{k}_1) \Phi_G(\mathbf{k}_2) \rangle \langle \Phi_G(\mathbf{q}_1) \Phi_G(\mathbf{q}_2) \rangle \right] + \dots \\ &= (2\pi)^3 f_{\text{NL}} \int \frac{d^3 \mathbf{q}_1}{(2\pi)^3} \frac{d^3 \mathbf{q}_2}{(2\pi)^3} \delta^{(D)}(\mathbf{q}_1 + \mathbf{q}_2 - \mathbf{k}_3) \left[\langle \Phi_G(\mathbf{k}_1) \Phi_G(\mathbf{q}_1) \rangle \langle \Phi_G(\mathbf{k}_2) \Phi_G(\mathbf{q}_2) \rangle \right. \\ & \quad \left. + \langle \Phi_G(\mathbf{k}_1) \Phi_G(\mathbf{q}_2) \rangle \langle \Phi_G(\mathbf{k}_2) \Phi_G(\mathbf{q}_1) \rangle \right] + \dots \\ &= (2\pi)^3 \delta^{(D)}(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) \\ & \quad \times 2f_{\text{NL}} [P_\Phi(k_1) P_\Phi(k_2) + P_\Phi(k_2) P_\Phi(k_3) + P_\Phi(k_3) P_\Phi(k_1)], \quad (1.52) \end{aligned}$$

where Isserlis' theorem has been applied and ellipses indicate contributions from the remaining cyclic permutations of \mathbf{k}_1 , \mathbf{k}_2 and \mathbf{k}_3 . Akin to the definition of kurtosis, one could consider the shape function of the bispectrum obtained after normalisation by the power spectrum,

$$S(k_1, k_2, k_3) = \frac{B(k_1, k_2, k_3)}{[P_\Phi(k_1) P_\Phi(k_2) P_\Phi(k_3)]^{2/3}}. \quad (1.53)$$

The basic single-field slow-roll model predicts an almost scale-invariant primordial scalar power spectrum, so that $P_\Phi(k) \propto k^{-3}$;²⁰ hence by equation (1.52),

$$S(k_1, k_2, k_3) \propto f_{\text{NL}} \left(\frac{k_1^2}{k_2 k_3} + \frac{k_2^2}{k_3 k_1} + \frac{k_3^2}{k_1 k_2} \right), \quad (1.54)$$

which has an enhanced amplitude in the squeezed limit, e.g. $k_3 \ll k_1, k_2$. However, $B(k_1, k_2, k_3)$ is still suppressed by slow-roll parameters which enter P_Φ [91]. In Figure 1.6, the shape function S for the bispectrum in the squeezed triangular configuration for local PNG is shown.

At first glance, and very naturally, it would seem that to uncover PNG, one would have to resort to high-order statistics solely and carefully disentangle from PNG any non-Gaussianity due to late-time non-linear gravitational evolution. However, in recent years there has been a breakthrough which has shown and verified with N -body simulations that PNG leaves a scale-dependent signature in the large-scale halo bias [93–95]. To see this, one could consider a peak–background split where background fluctuations are of long wavelengths and virialised haloes form from density peaks corresponding to small-scale perturbations [96, 97]. Taking

²⁰ The scalar spectral index has been set to $n_s = 1$ here assuming perfect scale invariance; otherwise, one can obtain a consistency relation for n_s and f_{NL} , which provides a test of single-field inflation [see e.g. 85, 90]. The primordial scalar power spectrum is evaluated at horizon exit for each Fourier mode, and on super-horizon scale, Φ is conserved. See also appendix B.

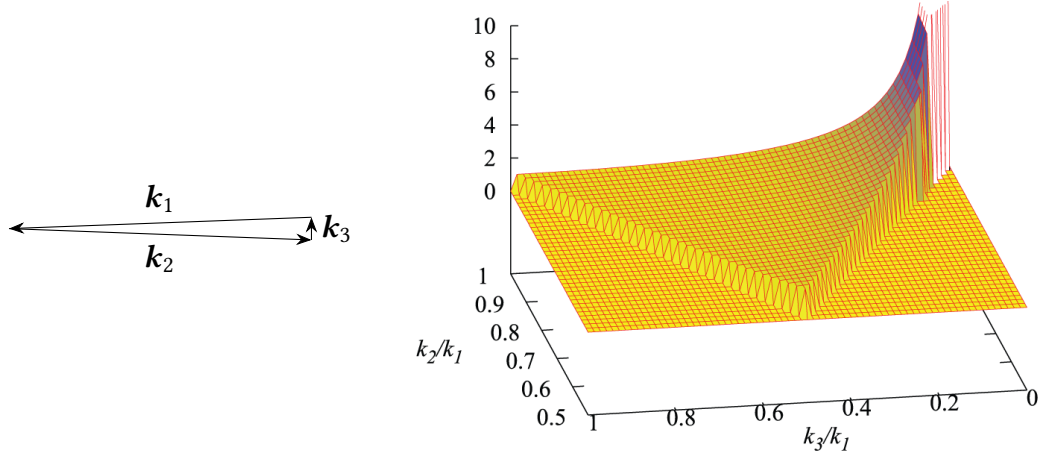


Figure 1.6. The squeezed triangular configuration $k_3 \ll k_1, k_2$ (left column) and the shape function $S(k_1, k_2, k_3)$ of the bispectrum (right column) for local PNG. Figure adapted from Fig. 1 in ref. [92] (credit: T Takahashi).

the non-Gaussian Bardeen potential field Φ_i described above at some initial epoch of structure formation (labelled with the subscript ‘i’), one could write the Gaussian part as $\Phi_G = \Phi_l + \Phi_s$, so that

$$\Phi_i = \Phi_l + f_{\text{NL}}^2 \Phi_l + (1 + 2f_{\text{NL}} \Phi_l) \Phi_s + f_{\text{NL}}^2 \Phi_s^2 + \dots, \quad (1.55)$$

where the subscripts ‘l’ and ‘s’ denote long- and short-wavelength contributions separately, and the ellipsis contains constant and higher-order terms. From the Friedmann equation (A.13a), the Poisson equation (B.10) and the discussions in appendix B about the linear growth factor D_1 and the transfer function T ,²¹ one can express the long-wavelength Fourier mode of matter density contrast as

$$\delta_{m,l}(k, z) = \mathcal{M}(k, z) \Phi_l(k) \quad \text{with} \quad \mathcal{M}(k, z) = \frac{2}{3} C(\Omega_{m,0})^{-1} \frac{D_1(z) T(k)}{\Omega_{m,0} H_0^2} k^2, \quad (1.56)$$

where H_0 and $\Omega_{m,0}$ are the Hubble and matter density parameters of today, and the factor C comes from the normalisation convention $D_1(0) = 1$ for today instead of $D_1(z) = a(z)$ in the matter-dominated era (see eq. B.19). Meanwhile, the short-wavelength Fourier mode is identified as

$$\delta_{m,s}(k, z) = \mathcal{M}(k, z) [1 + 2f_{\text{NL}} \Phi_l(k)] \Phi_s(k), \quad (1.57)$$

and it is evident that the effect of f_{NL} is a modulation of small-scale fluctuations by the long-wavelength background fluctuations. From the discussion of (proto-)halo bias in appendix C, it follows that this is equivalent to changing the local field variance,

$$\sigma_M^2 \mapsto \check{\sigma}_M^2 = (1 + 2f_{\text{NL}} \Phi_l)^2 \sigma_M^2, \quad (1.58)$$

so that it depends on long-wavelength modes; consequently, a bias modification is introduced to the scale-independent linear bias b_1 for haloes,

$$\Delta b(k, z) = \frac{2f_{\text{NL}}}{\mathcal{M}(k, z)} \frac{\partial \ln n_h}{\partial \ln \sigma_M} = C(\Omega_{m,0}) f_{\text{NL}} \frac{3\Omega_{m,0} H_0^2 \delta_c}{D_1(z) T(k)} \frac{b_1(z) - 1}{k^2}, \quad (1.59)$$

²¹ This transfer function links the Newtonian gravitational potential to the primordial Bardeen potential Φ ; $T(k) \simeq 1$ for $k \ll k_{\text{eq}}$, where k_{eq} corresponds to the comoving horizon scale at matter–radiation equality [98].

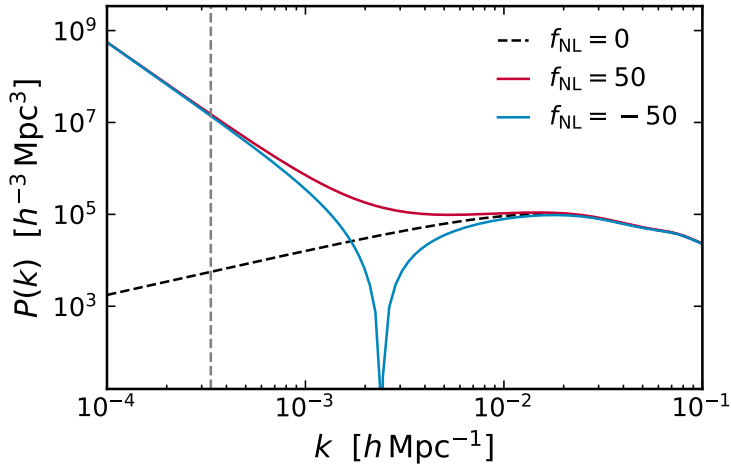


Figure 1.7. Scale-dependent signature of PNG in the clustering power spectrum at redshift $z = 0$ with $f_{\text{NL}} = 0, \pm 50$. The scale-independent linear bias is fixed at $b_1 = 2$ and *Planck* 2018 cosmology is adopted with $(h, \Omega_{\text{b},0}, \Omega_{\text{CDM},0}) = (0.674, 0.0493, 0.264)$ [101]. The vertical dashed line marks the horizon scale at $k = H_0$. The dip in the power spectrum for negative f_{NL} occurs when $(b_1 + \Delta b)$ changes sign.

where n_{h} is the halo mass function, and equations (1.56), (C.16) and (C.19) have been used. Since $T(k)$ tends to a constant as $k \rightarrow 0$, it follows that on large scales, Δb is enhanced by the k^{-2} factor—this is the scale-dependent PNG signature in halo bias.²² The scale-independent linear bias b_1 is sometimes termed the ‘Gaussian’ bias, and the modification Δb or the total bias $b = b_1 + \Delta b$ is known as the ‘non-Gaussian’ bias. In Figure 1.7, the effect of $f_{\text{NL}} = 0, \pm 50$ on the clustering power spectrum at redshift $z = 0$ is shown for a fixed linear bias $b_1 = 2$ and *Planck* 2018 cosmology with $(h, \Omega_{\text{b},0}, \Omega_{\text{CDM},0}) = (0.674, 0.0493, 0.264)$. For $f_{\text{NL}} < 0$, the bias modification Δb is negative and $(b_1 + \Delta b)$ changes sign at some intermediate k , thus leaving a dip in the power spectrum (as shown in the figure for $f_{\text{NL}} = -50$), which is a feature of the scale-dependent bias prediction.

Whilst this scale-dependent signature was demonstrated for haloes initially, tracers such as galaxies and quasars which reside in haloes are expected to exhibit similar behaviour in their bias in the presence of PNG. The precise calculation above for the amplitude of the bias modification is performed assuming a universal mass function \bar{n}_{h} ; more generally, it is described by

$$\Delta b(k, z) = \frac{2f_{\text{NL}}\delta_{\text{c}}}{\mathcal{M}(k, z)}(b_1 - p), \quad (1.60)$$

where p is a tracer-dependent parameter. For instance, for quasars that reside in haloes which have formed through mergers, $p = 1 + \delta_{\text{c}}^{-1} \approx 1.6$ is appropriate [95]; for a galaxy sample selected by stellar mass, $p \in [0.4, 0.7]$ might be expected [102].

Owing to the relative ease of measuring and computing the clustering power spectrum, the scale-dependent tracer bias offers a promising avenue for detecting PNG in large-scale structure. Indeed, its constraint on local PNG is currently tighter than those from the galaxy bis-

²² Besides the local-type PNG considered here, it has also been shown for other types of PNG that there is similarly a scale-dependent bias modification, though the dependence is weaker than k^{-2} for the local-type and thus harder to detect [e.g. 99, 100].

pectrum measurements, as the modelling of the latter is computationally challenging and suffers from the non-linearity of the late-time gravitational evolution of density perturbations. The most recent measurements by eBOSS using the scale-dependent quasar bias have reported $-81 < f_{\text{NL}} < 26$ at the 95 % uncertainty level [103]. Current and upcoming surveys, such as the Dark Energy Spectroscopic Instrument (DESI) and the *Euclid* mission, can reach uncertainty levels around $\sigma_{f_{\text{NL}}} = 5$ and bring competitive results in comparison with the CMB constraints, which may soon saturate the cosmic variance bound [104–106].



Local PNG appearing in large-scale tracer bias will serve as an important test parameter for validating novel approaches to the analysis of large-scale galaxy clustering proposed in this thesis. Although the modelling of its signature is relatively straightforward, survey systematics and other physical effects, such as relativistic corrections, may hamper its detection—these will be central to the discussions in chapters 3–5.

1.4 Cosmological likelihood inference

While LSS cosmology is described as an initial value problem earlier in this chapter, its probabilistic nature has also been made apparent by the preceding discussions. Therefore, LSS could also be framed as an *inverse problem of statistical inference*: given some clustering measurements, what could one learn about the underlying cosmological model, as parametrised by a set of parameters?

The philosophical interpretation of probability chiefly falls into two schools of thought, which take conceptually different approaches to statistical inference:²³

- in the frequentists’ framework, the underlying model parameters are unknown but fixed, and by constructing an appropriate estimator from the data, one would expect that with repeated sampling, the estimator should yield increasingly tighter confidence bounds enclosing the true value of model parameters;
- in the Bayesian framework, the underlying model parameters are random variables themselves, and any new information about them is encoded in their probability distributions conditional on the existing data.

Either of the two approaches has its own advantages and shortcomings, and in practice, a pragmatic cosmologist should take an agnostic position and adapt to the problem at hand with versatility.²⁴ However, in recent decades, the Bayesian framework has found wide application and success in the field of cosmology, not least because of the limited repeatability of many cosmological observations.²⁵ Using Bayes’ theorem, which is a basic yet profound result in probability theory, one can find the *posterior probability* \mathcal{P} of some cosmological model parameter(s) θ given the measurement data \mathbf{X} ,

$$\mathcal{P}(\theta | \mathbf{X}) = \frac{\pi(\theta)}{\mathbb{P}(\mathbf{X})} \mathcal{L}(\theta; \mathbf{X}). \quad (1.61)$$

²³ Many of the following discussions are based on textbook materials such as refs. [107–109] and reviews, e.g. refs. [110–112].

²⁴ ☞ “Black cat or white cat, if it can catch mice, it’s a good cat.” A (translated) quote by Deng Xiaoping.

²⁵ This alludes to the issue of cosmic variance in particular.

Here the *likelihood function* \mathcal{L} is given by the PDF of the data random variables conditional on the model parameters,

$$\mathcal{L}(\theta; \mathbf{X}) = \mathbb{p}(\mathbf{X} | \theta); \quad (1.62)$$

the PDF $\pi(\theta)$ specifies the *prior distribution* of the parameters; and the PDF of the data marginalised over all model parameters, $\mathbb{p}(\mathbf{X})$, is sometimes known as the evidence and acts to normalise the posterior. It is worth emphasising here that although the likelihood is formed from the PDF of the data, it should be viewed as a function of the model parameters θ with the data variables \mathbf{X} fixed at the observed values. In either the frequentists' or the Bayesian framework, the likelihood plays a central rôle in classical parameter inference.

In principle, both the prior distribution and the likelihood should ideally be motivated by the underlying theoretical model and physical processes that generate the measurement data, e.g. the Poisson distribution for the galaxy number count as described in section 1.1; however, this may not be always possible. In the absence of a theoretically motivated prior, an uninformative prior distribution may be used, e.g. the uniform prior for a location parameter such as the mean, or the Jeffreys prior for a scale parameter like the variance.²⁶ For the likelihood distribution, one could appeal to asymptotic normality when the data variable itself is the sum of N independent random variables as $N \rightarrow \infty$; one example is the power spectrum averaged in some wavenumber interval from many independent Fourier modes. Asymptotic normality as a universality law is warranted by the (generalised) central limit theorem, which states that the sum of a large number of independent random variables converges in distribution to a normal distribution regardless of their individual distributions, provided reasonable regularity conditions are met.

As such, it is common in LSS analyses to assume the *Gaussian likelihood*, which refers to the normal distribution of the data rather than the Gaussian shape of the likelihood in the model parameter space.²⁷ If one takes the estimated power spectrum at binned wavenumbers k_i as the data vector, $\mathbf{X} = \hat{\mathbf{P}} = (\hat{P}(k_i))$, then the Gaussian likelihood for some cosmological parameter(s) θ takes the following form:

$$\mathcal{L}(\theta; \hat{\mathbf{P}}) = \mathbb{p}(\hat{\mathbf{P}} | \theta) = \frac{1}{\sqrt{|2\pi \boldsymbol{\Sigma}|}} \exp\left\{-\frac{1}{2} [\hat{\mathbf{P}} - \mathbf{P}(\theta)]^\top \boldsymbol{\Sigma}^{-1} [\hat{\mathbf{P}} - \mathbf{P}(\theta)]\right\}. \quad (1.63)$$

Here a hat $\hat{\cdot}$ distinguishes an estimate from the underlying quantity itself, and $\mathbf{P}(\theta)$ is the mean vector given by model predictions. All that is required for likelihood evaluation is then the covariance matrix of the power spectrum,

$$\boldsymbol{\Sigma} = \text{cov}[\hat{\mathbf{P}}] = \langle \hat{\mathbf{P}} \hat{\mathbf{P}}^\top \rangle - \langle \hat{\mathbf{P}} \rangle \langle \hat{\mathbf{P}}^\top \rangle. \quad (1.64)$$

Since $\boldsymbol{\Sigma}$ is a four-point correlator of the over-density field, analytic expressions for it are not always available; even in the linear perturbative regime where the field is Gaussian and $\boldsymbol{\Sigma}$ can be decomposed into two-point correlators, the correlation between Fourier modes introduced by

²⁶ More generally, one could subscribe to the principle of maximum entropy [see e.g. 109, 113].

²⁷ The term 'Gaussian likelihood', commonly used in astronomy, is regarded as a possible misnomer in this thesis; it may have come from the misconception of equality as identity between the likelihood and the underlying data distribution. See the discussion under equation (1.62) and also footnote 6 on page 6.

the survey window can still pose significant challenges to its modelling.²⁸ Therefore in galaxy survey analyses, it is usually more reliable to estimate the covariance matrix from $N_{\text{mock}} (\gg 1)$ mock catalogues,

$$\hat{\Sigma} = \frac{1}{N_{\text{mock}} - 1} \sum_{j=1}^{N_{\text{mock}}} (\hat{P}_j - \bar{P})(\hat{P}_j - \bar{P})^\top, \quad (1.65)$$

where the averaged power spectrum of the mock catalogues is

$$\bar{P} = \frac{1}{N_{\text{mock}}} \sum_{j=1}^{N_{\text{mock}}} \hat{P}_j. \quad (1.66)$$

The number of mocks, N_{mock} , should satisfy $N_{\text{mock}} > d+2$ where d is the dimension of the data vector $\mathbf{X} = \hat{P}$ [116]. For typical galaxy survey analyses, $N_{\text{mock}} = O(10^3)$ (e.g. $N_{\text{mock}} = 1000$ in recent eBOSS analyses [81, 82]).

With the replacement of the unknown underlying covariance matrix Σ by its estimate $\hat{\Sigma}$, two key issues arise:

- 1) Σ appears in equation (1.63) as its inverse, the *precision matrix* $\Psi = \Sigma^{-1}$, but $\mathbb{E}[\hat{\Sigma}^{-1}] \neq \mathbb{E}[\hat{\Sigma}]^{-1}$, i.e. $\hat{\Sigma}^{-1}$ is not an unbiased estimate of Ψ ;
- 2) $\hat{\Sigma}$ is intrinsically noisy, which could introduce additional statistical uncertainty into any estimates of θ .

With regard to the first point, ref. [117] noted that the conditional distribution of $\hat{\Sigma}^{-1} | \Sigma^{-1}$ is inverse Wishart [118], and a multiplicative factor is needed to obtain an unbiased estimate of Ψ ,

$$\hat{\Psi} = \frac{m - d - 1}{m} \hat{\Sigma}^{-1}, \quad (1.67)$$

where $m = N_{\text{mock}} - 1$. As for the second point, refs. [116, 119, 120] have shown by means of Taylor expansions in the covariance matrix and parameter errors that, to account for the uncertainty in covariance estimation, the square uncertainties of the parameter estimates need to be corrected by a multiplicative factor of

$$\frac{\sigma_\theta^2(\text{with } \hat{\Sigma})}{\sigma_\theta^2(\text{with } \Sigma)} = \frac{1 + B(d - p)}{1 + A + B(p + 1)}, \quad (1.68a)$$

with the factors A and B given by

$$A = \frac{2}{(m - d)(m - d - 3)} \quad \text{and} \quad B = \frac{m - d - 1}{(m - d)(m - d - 3)}, \quad (1.68b)$$

where p is the number of parameters being estimated.

Recently, Sellentin & Heavens have employed a Bayesian approach to propagating covariance estimation uncertainties to the estimated parameters, where the unknown true covariance matrix Σ is treated as a random variable to be marginalised over [121, hereafter SH].²⁹ First,

²⁸ Nonetheless, significant progress has been made recently in deriving analytic covariance matrices; see e.g. refs. [114, 115].

²⁹ This approach has been considered before [see e.g. 120], but SH has found a prior which makes the calculation feasible.

one recalls that the conditional distribution of $\hat{\Sigma} | \Sigma$ is Wishart with the PDF [118]

$$\mathbb{P}(\hat{\Sigma} | \Sigma) = \frac{2^{-md/2}}{\Gamma_d(m/2)} \frac{|\hat{\Sigma}|^{(m-d-1)/2}}{|\Sigma/m|^{m/2}} \exp\left[-\frac{m}{2} \text{tr}(\Sigma^{-1} \hat{\Sigma})\right], \quad (1.69)$$

where the multivariate gamma function is given by

$$\Gamma_d(x) = \pi^{d(d-1)/4} \prod_{i=1}^d \Gamma\left(x - \frac{i-1}{2}\right). \quad (1.70)$$

If one specifies the Jeffreys prior for Σ ,

$$\pi(\Sigma) \propto |\Sigma|^{-(d+1)/2}, \quad (1.71)$$

then by equation (1.61), the posterior for Σ is given by the PDF of the inverse Wishart distribution,

$$\mathcal{P}(\Sigma | \hat{\Sigma}) = \frac{2^{-md/2}}{\Gamma_d(m/2)} \frac{|m\hat{\Sigma}|^{m/2}}{|\Sigma|^{(m+d+1)/2}} \exp\left[-\frac{m}{2} \text{tr}(\Sigma^{-1} \hat{\Sigma})\right]. \quad (1.72)$$

By marginalising equation (1.63) over this posterior distribution, one then obtains the modified Student's t -distribution likelihood,

$$\begin{aligned} \mathcal{L}(\theta; \hat{P}, \hat{\Sigma}) &= \int d\Sigma \mathbb{P}(\hat{P} | \theta, \Sigma) \mathcal{P}(\Sigma | \hat{\Sigma}) \\ &= (m\pi)^{-d/2} \frac{\Gamma(\frac{m+1}{2})}{\Gamma(\frac{m-d+1}{2})} |\hat{\Sigma}|^{-1/2} \left[1 + (\hat{P} - P)^\top \frac{\hat{\Sigma}^{-1}}{N-1} (\hat{P} - P) \right]^{-(m+1)/2}. \end{aligned} \quad (1.73)$$

In comparison with the original Gaussian likelihood (eq. 1.63), the evaluation of the modified- t likelihood is no more computationally expensive. To gauge the amount of uncertainty introduced by covariance estimation, Sellentin & Heavens use the Fisher information marginalised over the estimated covariance matrix $\hat{\Sigma}$ as a barometer, and considers the corresponding systematic loss of information through the volume change in the credibility regions of the posterior distribution \mathcal{P} as well as the figure of merit for the model parameters θ [122].



Despite the wide applicability of the Gaussian likelihood, there are inevitably situations where it should not be used: when one considers the binned power spectrum on very large scales, the number of contributing Fourier modes is small, and thus the central limit theorem does not apply. In those circumstances, it is imperative to consider non-Gaussian likelihoods, i.e. non-normal underlying distributions of the data. Moreover, the discussion so far has neglected the possibility that the covariance matrix itself may depend on the model parameters being estimated, i.e. $\Sigma \equiv \Sigma(\theta)$ in equations (1.63) and (1.64). These are the focal issues explored in chapter 3; however, in situations where non-Gaussian likelihoods and parameter dependence of the covariance matrix are not of concern, the results presented in this section remain applicable.

1.5 Motivation and outline

As primary CMB probes gradually saturate the cosmic variance bound, LSS probes such as BAO and RSD will become ever more indispensable. Their statistical power lies within the large three-dimensional data sets to be gathered by upcoming galaxy surveys such as DESI and *Euclid*, which will map out unprecedented cosmic volumes and tighten cosmological constraints sensitive to clustering measurements on large scales; these include signatures of PNG and relativistic effects, which are respectively important for inflationary model selection and testing the nature of gravity on ultra-large scales [93–95, 123–130].

Although the study of galaxy clustering is now an established discipline, techniques for analysing large galaxy surveys are still being constantly improved and updated. In order to further motivate the research underlying this thesis, chapter 2 will review the current framework of galaxy clustering analyses, where one constructs estimators of the power spectrum multipoles, which are compatible with fast Fourier transform (FFT) algorithms, out of a weighted over-density field formed from the galaxy catalogue and a high-density, high-fidelity mock catalogue. Chapters 3–5 will explore the different cross-sections of the confrontation between theoretical modelling and data analysis of LSS observations *on large, linear scales*—

- 1) The issues of non-Gaussian likelihoods and parameter-dependent covariance matrices mentioned in section 1.4 are addressed in chapter 3 (based on Paper I). Two new techniques are proposed to resolve them: Gaussianisation of the data variable, and the rescaling of covariance matrices for different cosmological models through the variance–correlation decomposition;
- 2) In chapter 4 (based on Paper II), a forward-modelling approach is employed to propagate measurements of the tracer luminosity function to constraints of the relativistic bias parameters, which crucially determine the amplitude of relativistic corrections to the over-density field, and thus their scale-dependent modifications of clustering statistics, on scales close to the Hubble horizon;
- 3) A hybrid-basis Fourier analysis framework is introduced in chapter 5 (based on Paper III), where the computational edge of the conventional power spectrum analysis is combined with advantages of a spherical Fourier analysis that is more suited to realistic survey geometry. The spherical Fourier analysis can circumvent the distant-observer and plane-parallel approximations discussed in section 1.2, and by constructing its likelihood function directly from the over-density field, it does not require covariance estimation.

For reasons elaborated in section 1.3, local PNG f_{NL} will be the primary parameter considered for cosmological inference from large-scale galaxy clustering, as it serves as a natural test parameter for the methodologies presented in this thesis. It is also a key scientific objective for next-generation large galaxy surveys such as DESI and *Euclid*. Finally, chapter 6 summarises the results from these novel approaches to galaxy clustering analysis, points out new directions for future investigations, and concludes with an optimistic outlook for what may come to fruition in the golden age of cosmology.

Galaxy Clustering Analysis

The common framework of galaxy clustering analysis has evolved over the last two decades from the *Feldman–Kaiser–Peacock* (FKP) formalism originally outlined in ref. [58, hereafter FKP]: the galaxy over-density field is first calculated from the survey galaxy catalogue and a synthetic random catalogue, the latter essentially Monte Carlo sampling the background galaxy number density as modulated by survey geometry and selection effects; at the same time, a general weighting scheme can be applied to individual galaxies, which seeks to minimise the variance of the summary statistics (most commonly the power spectrum multipoles binned in wavenumbers). Based on this weighted galaxy over-density field, different estimators of the summary statistics can be constructed, and the choice appropriate for a particular survey analysis depends on the required accuracy and precision as well as the computational cost associated with the estimator. Finally, in a typical likelihood analysis where models are fitted to the data, one must also compute model predictions for the chosen summary statistics, accounting for the survey geometry, selection effects and any other data transformations involved in processing the raw catalogue data into the summary statistics.

Throughout this chapter, the Fourier analysis of galaxy clustering statistics is performed in redshift space, so the subscript ‘*s*’ denoting redshift-space quantities will be omitted for brevity; in any case, the ambiguity between real- and redshift-space quantities can be eliminated by the context or the coordinates explicitly used.

2.1 The Feldman–Kaiser–Peacock formalism

Given a survey catalogue of N_g galaxies, one needs to synthesise a random catalogue with the same geometry and selection effects as those of the survey catalogue, which should contain a large number of galaxies, N_s . The ratio of the number counts (or the mean number densities), $\alpha = N_g/N_s \ll 1$, is chosen to be small to reduce the shot noise contribution from the synthetic random catalogue. The FKP weighted over-density field is then given by

$$F(\mathbf{s}) = \frac{w(\mathbf{s})}{I^{1/2}} [n(\mathbf{s}) - \alpha n_s(\mathbf{s})], \quad (2.1)$$

where n and n_s are the number densities of the survey and synthetic catalogues, w represents some weighting scheme,

$$I = \int d^3\mathbf{s} w(\mathbf{s})^2 \bar{n}(\mathbf{s})^2 \quad (2.2)$$

is a normalisation factor, and $\bar{n}(\mathbf{s})$ is the modulated background number density (recall from eq. 1.2.4).

The number density fields can be evaluated as sums of Dirac delta contributions from each galaxy in the catalogue, so the Fourier modes of the weighted over-density field are calculated to be

$$F(\mathbf{k}) = \frac{1}{I^{1/2}} \sum_{i=1}^{N_g} w(\mathbf{s}_i) e^{-i\mathbf{k}\cdot\mathbf{s}_i} - \frac{\alpha}{I^{1/2}} \sum_{i=1}^{N_s} w(\mathbf{s}_i) e^{-i\mathbf{k}\cdot\mathbf{s}_i}, \quad (2.3)$$

where the normalisation factor can be expressed as

$$I = \alpha \sum_{i=1}^{N_s} w(\mathbf{s}_i)^2 \bar{n}(\mathbf{s}_i), \quad (2.4)$$

and to ensure that $\langle F \rangle = 0$, the number count ratio is recalculated to be

$$\alpha = \frac{\sum_{i=1}^{N_g} w(\mathbf{s}_i)}{\sum_{i=1}^{N_s} w(\mathbf{s}_i)}. \quad (2.5)$$

For later convenience, one could define a series of integrals given by

$$G_{a,b}(\mathbf{k}) = I^{-a/2} \int d^3\mathbf{s} e^{-i\mathbf{k}\cdot\mathbf{s}} w(\mathbf{s})^a \bar{n}(\mathbf{s})^b, \quad a, b \in \mathbb{N}, \quad (2.6)$$

which is discretely evaluated as

$$G_{a,b}(\mathbf{k}) = \alpha \sum_{i=1}^{N_s} w(\mathbf{s}_i)^a \bar{n}(\mathbf{s}_i)^{b-1} e^{-i\mathbf{k}\cdot\mathbf{s}_i}. \quad (2.7)$$

In practice, evaluation of each Fourier mode by direct summation can be prohibitively slow for a large data set; to utilise FFT algorithms, galaxies are instead interpolated on a regular cubic grid, and then Fourier modes of the over-density field can be computed with appropriate multiplicative corrections to compensate for the amplitude change due to the interpolation window [131]. The use of discrete Fourier transform also means that there are aliasing effects in the sampled Fourier modes, which are non-trivial to remove; however, this is only significant for Fourier wavenumbers close to the Nyquist wavenumber, $k_N = N_{\text{grid}} k_f / 2$, where N_{grid} is the grid number in each dimension and k_f is the fundamental wavenumber (recall from §1.1).¹

In Fourier space, it follows from equation (2.1) that

$$\langle |F(\mathbf{k})|^2 \rangle = \int \frac{d^3\mathbf{q}}{(2\pi)^3} |G_{1,1}(\mathbf{k} - \mathbf{q})|^2 P(\mathbf{q}) + G_{2,1}(\mathbf{0}), \quad (2.8)$$

where the second term is the shot noise contribution. In the original FKP analysis of the *Infrared Astronomical Satellite (IRAS)* survey, the window function $G_{1,1}(\mathbf{k})$ is very localised, so given the choice of the normalisation factor I in equation (2.2), it follows that

$$\langle |F(\mathbf{k})|^2 \rangle \approx P(\mathbf{k}) + P_{\text{shot}}. \quad (2.9)$$

¹ See also ref. [132] which introduces the interlacing method for eliminating the odd-parity aliasing images.

Therefore an estimator for the power spectrum without shot noise is given by

$$\widehat{P}(\mathbf{k}) = |F(\mathbf{k})|^2 - P_{\text{shot}}, \quad (2.10)$$

where the shot noise power is calculated discretely by

$$P_{\text{shot}} = \frac{1}{I} \sum_{i=1}^{N_g} w(\mathbf{s}_i)^2 + \frac{\alpha^2}{I} \sum_{i=1}^{N_s} w(\mathbf{s}_i)^2. \quad (2.11)$$

If one is only concerned with the information in some isotropic band power spectrum obtained by averaging in spherical shells of thickness Δk ,

$$\bar{P}(k) = \frac{1}{V_k} \int_{q \in S_k} d^3 \mathbf{q} P(\mathbf{q}), \quad (2.12)$$

where the shell $S_k = [k - \Delta k/2, k + \Delta k/2)$ centred at wavenumber k has volume

$$V_k = 4\pi k^2 \Delta k \left[1 + \frac{1}{12} \left(\frac{\Delta k}{k} \right)^2 \right], \quad (2.13)$$

then the estimator $\widehat{P}(\mathbf{k})$ can be binned,

$$\widehat{P}(k) = \frac{1}{N_k} \sum_{q \in S_k} \widehat{P}(\mathbf{q}), \quad (2.14)$$

where N_k is the number of wavevectors \mathbf{k} sampled within the shell.

To derive the FKP weighting scheme, one starts by considering the covariance between Fourier modes,

$$\begin{aligned} \langle F(\mathbf{k}_1) F^*(\mathbf{k}_2) \rangle &= \int \frac{d^3 \mathbf{q}}{(2\pi)^3} G_{1,1}(\mathbf{k}_1 - \mathbf{q}) G_{1,1}^*(\mathbf{k}_2 - \mathbf{q}) P(\mathbf{q}) \\ &\quad + (1 + \alpha) G_{2,1}(\mathbf{k}_1 - \mathbf{k}_2). \end{aligned} \quad (2.15)$$

If the window function $G_{1,1}(\mathbf{k})$ is assumed to be localised (i.e. with sufficiently compact support), the covariance is non-vanishing only for small $\Delta \mathbf{k} = \mathbf{k}_2 - \mathbf{k}_1$,

$$\langle F(\mathbf{k}_1) F^*(\mathbf{k}_2) \rangle \approx G_{2,2}(\Delta \mathbf{k}) P(\mathbf{k}_1) + (1 + \alpha) G_{2,1}(\Delta \mathbf{k}). \quad (2.16)$$

With F being a GRF, the covariance of the power spectrum estimator is then

$$\begin{aligned} \text{cov} \left[\widehat{P}(\mathbf{k}_1), \widehat{P}(\mathbf{k}_2) \right] &\approx |\langle F(\mathbf{k}_1) F^*(\mathbf{k}_2) \rangle|^2 \\ &\approx \left| G_{2,2}(\Delta \mathbf{k}) P(\mathbf{k}_1) + (1 + \alpha) G_{2,1}(\Delta \mathbf{k}) \right|^2. \end{aligned} \quad (2.17)$$

For sufficiently large width Δk of the aforementioned spherical shells, the variance of the isotropic band power spectrum estimator $\widehat{P}(k)$ can be expressed as

$$\sigma_{\widehat{P}}^2(k) \approx \frac{1}{V_k} \int d^3 \mathbf{q} \left| G_{2,2}(\mathbf{q}) \bar{P}(k) + (1 + \alpha) G_{2,1}(\mathbf{q}) \right|^2, \quad (2.18)$$

and the fractional variance is thus

$$\sigma_{\bar{P}}^2(k) / \bar{P}(k)^2 \approx \frac{(2\pi)^3}{V_k I^2} \int d^3s w(\mathbf{s})^4 \bar{n}(\mathbf{s})^4 \left[1 + \frac{1}{\bar{n}(\mathbf{s}) \bar{P}(k)} \right]. \quad (2.19)$$

By considering its functional variation with respect to the weight w and the stationary point satisfying

$$\frac{\delta}{\delta w} (\sigma_{\bar{P}}^2 / \bar{P}^2) = 0, \quad (2.20)$$

one finally obtains the *FKP weight* (taking $\alpha \approx 0$),

$$w(\mathbf{s}; k) = \frac{1}{1 + \bar{n}(\mathbf{s}) \bar{P}(k)}. \quad (2.21)$$

This weight is explicitly dependent on the wavenumber k of the band power spectrum one attempts to estimate; in practice, it suffices to set it to a constant value typical of the power spectrum in the range of scales of interest, e.g. $\bar{P} = 10000 h^3 \text{Mpc}^{-3}$ set in the BOSS Data Release 12 analysis by ref. [79], or $\bar{P} = 30000 h^3 \text{Mpc}^{-3}$ for the PNG analysis with the eBOSS QSO sample in ref. [103].

The arguments presented above follow from the original work of FKP; subsequently, ref. [133] extended the FKP weighting scheme to account for luminosity dependence in galaxy clustering, and ref. [134] has formulated a more general and formal derivation of FKP-style weights. Additional weights accounting for observational systematics such as redshift failure and fibre collision can also be incorporated [e.g. 135, 136]. However, the redshift dependence of power spectrum estimators has so far only been treated rather simplistically. Since galaxies of different redshifts and weights may contribute to the same power spectrum measurement, it is only sensible to define an effective redshift for a given catalogue sample [e.g. 79],

$$z_{\text{eff}} = \frac{\sum_{i=1}^{N_g} w(\mathbf{s}_i) z_i}{\sum_{i=1}^{N_g} w(\mathbf{s}_i)}. \quad (2.22)$$

For two-point function measurements, ref. [137] argues for a different definition through a Taylor expansion around the stationary redshift point,

$$z_{\text{eff}} = \frac{\sum_{i=1}^{N_g} w(\mathbf{s}_i)^2 z_i}{\sum_{i=1}^{N_g} w(\mathbf{s}_i)^2}; \quad (2.23)$$

there are alternative definitions quadratic in weights, but their differences are negligible given the current level of statistical uncertainties [138]. This ambiguity in the effective redshift is a motivation for non-tomographic analyses, e.g. the spherical Fourier analysis discussed in chapter 5. It is also worth noting that from the FKP derivation, an approximate diagonal covariance matrix for the band power spectrum estimator emerges,

$$\text{cov}[\hat{P}(k_j), \hat{P}(k_m)] \approx \frac{V_c}{V_{k_j}} P(k_j) P(k_m) \delta_{jm}^{(K)}, \quad (2.24)$$

where the coherence volume factor is given by

$$V_c(k) = \frac{(2\pi)^3}{I^2} \int d^3s w(\mathbf{s})^4 \bar{n}(\mathbf{s})^4 \left[1 + \frac{1}{\bar{n}(\mathbf{s}) \bar{P}(k)} \right]. \quad (2.25)$$

However, this covariance matrix does not account for the correlation between wavenumber bins as introduced by the survey window, selection and weighting scheme.

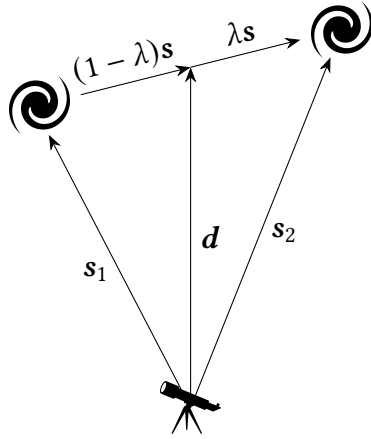


Figure 2.1. Geometric configuration of the local plane-parallel picture. For a galaxy pair at positions \mathbf{s}_1 and \mathbf{s}_2 separated by $\mathbf{s} = \mathbf{s}_2 - \mathbf{s}_1$, any lines of sight \mathbf{d} in between (i.e. $\lambda \in [0, 1]$) are considered equivalent.

2.2 Local plane-parallel estimators

Whilst the FKP formalism and weighting scheme enjoy wide applicability, the power spectrum estimator (eq. 2.10) presented earlier, which assumes a localised window function for narrow, pencil beam-like surveys, is not suitable for wide-area surveys. Moreover, as one recalls from section 1.2, galaxy clustering in redshift space is anisotropic with the key observables prescribed in the global plane-parallel limit. However, the assumption of a global line of sight is an invalid one for large surveys, since the line of sight to each galaxy varies in direction significantly. This poses the question as to whether the global plane-parallel power spectrum $P(k, \mu)$, or equivalently its multipoles $P_\ell(k)$, can actually be recovered from clustering measurements.

One attempt to address this problem, introduced by Yamamoto et al. [139], is to adopt the *local*, or *pairwise*, *plane-parallel approximation*, whereby the power spectrum multipole is estimated by

$$\hat{P}_\ell(k) = (2\ell + 1) \int \frac{d^2 \hat{\mathbf{k}}}{4\pi} \int d^3 \mathbf{s}_1 e^{i\mathbf{k} \cdot \mathbf{s}_1} \int d^3 \mathbf{s}_2 e^{-i\mathbf{k} \cdot \mathbf{s}_2} \mathcal{L}_\ell(\hat{\mathbf{k}} \cdot \hat{\mathbf{d}}) F(\mathbf{s}_1) F(\mathbf{s}_2) \quad (2.26)$$

with a pairwise line of sight,

$$\mathbf{d} = \lambda \mathbf{s}_1 + (1 - \lambda) \mathbf{s}_2 \quad (0 \leq \lambda \leq 1), \quad (2.27)$$

for two galaxies located at positions \mathbf{s}_1 and \mathbf{s}_2 that contribute to the clustering signal (see Fig. 2.1 for a diagram) [78]. For any value of $\lambda \in [0, 1]$, \mathbf{d} lies between \mathbf{s}_1 and \mathbf{s}_2 and is considered equivalent, i.e. locally parallel. The advantage of this estimator, hereafter referred to as the *Yamamoto estimator*, stems from the exploitation of the remaining rotational symmetry around the observer, despite anisotropies around the line of sight induced by RSDs and the AP effect. It works effectively when the clustering scale of interest is small in comparison with the typical distance to a pair of galaxies. However, it is evident that the Yamamoto estimator cannot match the global plane-parallel limit perfectly; to see this, one could note that the choice of $\lambda = 0$ or 1 above breaks the exchange symmetry in a galaxy pair, and the resulting estimate will be different from that with, for instance, $\lambda = 1/2$. The discrepancy between the expectation $\langle \hat{P}_\ell \rangle$ (with any choice of $\lambda \in [0, 1]$) and the global plane-parallel P_ℓ is the so-called

‘*wide-angle effect*’, which motivates the spherical (or hybrid-basis) Fourier analysis discussed in chapter 5.

To unravel the relationship between the Yamamoto estimator and the global plane-parallel prediction, it is instructive to consider the so-called ‘local power spectrum’,

$$P(\mathbf{k}, \mathbf{d}; \lambda) = \int d^3 \mathbf{s} e^{-i\mathbf{k}\cdot\mathbf{s}} \langle \delta(\mathbf{d} - (1 - \lambda)\mathbf{s}) \delta(\mathbf{d} + \lambda\mathbf{s}) \rangle, \quad (2.28)$$

where $\mathbf{s} = \mathbf{s}_2 - \mathbf{s}_1$ is the separation vector [140]. By expressing δ as an inverse Fourier transform and changing variables, one arrives at

$$\begin{aligned} P(\mathbf{k}, \mathbf{d}; \lambda) &= \iint \frac{d^3 \mathbf{q}}{(2\pi)^3} \frac{d^3 \mathbf{p}}{(2\pi)^3} e^{i(\mathbf{q}+\mathbf{p})\cdot\mathbf{d}} \langle \delta(\mathbf{q}) \delta(\mathbf{p}) \rangle \int d^3 \mathbf{s} e^{-i\mathbf{k}\cdot\mathbf{s}} e^{-i(1-\lambda)\mathbf{q}\cdot\mathbf{s}} e^{i\lambda\mathbf{p}\cdot\mathbf{s}} \\ &= \int \frac{d^3 \mathbf{q}}{(2\pi)^3} e^{i(\mathbf{k}+\mathbf{q})\cdot\mathbf{d}/\lambda} \langle \delta(\mathbf{q}) \delta\left(-\mathbf{q} + \frac{\mathbf{k}+\mathbf{q}}{\lambda}\right) \rangle \\ &= \int \frac{d^3 \mathbf{q}}{(2\pi)^3} e^{i\mathbf{q}\cdot\mathbf{d}} \langle \delta(\mathbf{k} + (1 - \lambda)\mathbf{q}) \delta(-\mathbf{k} + \lambda\mathbf{q}) \rangle, \end{aligned} \quad (2.29)$$

where the Fourier representation and the scaling property of the Dirac delta distribution have been used in the second and third lines respectively. In the last line, the Fourier-space two-point correlator does not reduce to the usual power spectrum, since homogeneity and translational invariance are broken by anisotropy. However, in the global plane-parallel limit, the two-point correlator collapses to become diagonal, $\langle \delta(\mathbf{k} + (1 - \lambda)\mathbf{q}) \delta(-\mathbf{k} + \lambda\mathbf{q}) \rangle \rightarrow (2\pi)^3 \delta^{(D)}(\mathbf{q}) P(\mathbf{k})$, so that one recovers $P(\mathbf{k}, \mathbf{d}; \lambda) \rightarrow P(\mathbf{k})$. If one instead considers a finite survey volume and averages over all possible lines of sight, then

$$\frac{1}{V_s} \int d^3 \mathbf{d} P(\mathbf{k}, \mathbf{d}; \lambda) = \delta_{q_0}^{(K)} \langle \delta(\mathbf{k} + (1 - \lambda)\mathbf{q}) \delta(-\mathbf{k} + \lambda\mathbf{q}) \rangle = \langle \delta(\mathbf{k}) \delta^*(\mathbf{k}) \rangle, \quad (2.30)$$

where equation (1.12) has been used. This suggests that the volume average of the local power spectrum is equivalent to the two-point auto-correlator. With the inclusion of the Legendre polynomials, an analogous result holds for the Yamamoto estimator (with any choice of $\lambda \in [0, 1]$; label omitted hereafter) [78],

$$\langle \widehat{P}_\ell(k) \rangle = \frac{2\ell + 1}{V_s} \int d^3 \mathbf{d} \int \frac{d^2 \hat{\mathbf{k}}}{4\pi} \mathcal{L}_\ell(\hat{\mathbf{k}} \cdot \hat{\mathbf{d}}) P(\mathbf{k}, \mathbf{d}) = \frac{1}{V_s} \int d^3 \mathbf{d} P_\ell(k, d), \quad (2.31)$$

where the following Legendre expansion holds,

$$P(\mathbf{k}, \mathbf{d}) = \sum_\ell \mathcal{L}_\ell(\hat{\mathbf{k}} \cdot \hat{\mathbf{d}}) P_\ell(k, d). \quad (2.32)$$

If one considers the local correlation function analogously defined in configuration space,

$$\xi(\mathbf{s}_1, \mathbf{s}_2) \equiv \xi(\mathbf{s}, \mathbf{d}) = \sum_\ell \xi_\ell(s, d) \mathcal{L}_\ell(\hat{\mathbf{s}} \cdot \hat{\mathbf{d}}), \quad (2.33)$$

where each multipole $\xi_\ell(s, d)$ admits a *wide-angle expansion* of the form

$$\xi_\ell(s, d) = \sum_i \left(\frac{s}{d}\right)^i \xi_\ell^{(i)}(s), \quad (2.34)$$

then the multipole $P_\ell(k, d)$ can be obtained by a Hankel transform,

$$P_\ell(k, d) = 4\pi i^{-\ell} \int s^2 ds j_\ell(ks) \xi_\ell(s, d), \quad (2.35)$$

where j_ℓ is the spherical Bessel function of the first kind. The coefficients $\xi_\ell^{(i)}$ of the wide-angle expansion (eq. 2.34) depend on the choice of the line of sight, and they are formed from the correlation function multipoles $\xi_\ell(s)$ derived in the global plane-parallel limit. For detailed derivation of the wide-angle expansion and these coefficients, the reader may refer to refs. [78, 141, 142].² The key point here is that, as it is evident now, the Yamamoto estimator as the volume average of the local power spectrum $P_\ell(k, d)$ contains wide-angle corrections at all orders to the global plane-parallel limit [142].

When the Yamamoto estimator (eq. 2.26) was first introduced with $\lambda = 1/2$, it was considered computationally expensive to evaluate because of the nested doubled integral. However, with the endpoint line-of-sight choice ($\lambda = 0$ or 1), refs. [140, 145] recognised that the double integral becomes separable and is compatible with FFT algorithms. If one considers the transformation

$$F_n(\mathbf{k}) = \int d^3\mathbf{s} e^{-i\mathbf{k}\cdot\mathbf{s}} (\hat{\mathbf{k}} \cdot \hat{\mathbf{s}})^n F(\mathbf{s}), \quad (2.36)$$

then by writing each Legendre polynomial as a sum of monomials,

$$\mathcal{L}_\ell(\mu) = 2^{-\ell} \sum_{n=0}^{\lfloor \ell/2 \rfloor} (-1)^n \binom{\ell}{n} \binom{2\ell-2n}{\ell} \mu^{\ell-2n}, \quad (2.37)$$

one can obtain the Yamamoto estimator as

$$\hat{P}_\ell(k) = \frac{2\ell+1}{2^\ell} \sum_{n=0}^{\lfloor \ell/2 \rfloor} (-1)^n \binom{\ell}{n} \binom{2\ell-2n}{\ell} \int \frac{d^2\hat{\mathbf{k}}}{4\pi} F_0^*(\mathbf{k}) F_{\ell-2n}(\mathbf{k}). \quad (2.38)$$

Here $F_n(\mathbf{k})$ can be calculated using $(n+1)(n+2)/2 = O(n^2)$ FFTs of the form

$$\int d^3\mathbf{s} e^{-i\mathbf{k}\cdot\mathbf{s}} \hat{s}_{i_1} \cdots \hat{s}_{i_n} F(\mathbf{s}), \quad (2.39)$$

where \hat{s}_{i_j} is the i_j th Cartesian component ($i_j \in \{1, 2, 3\}$) of $\hat{\mathbf{s}}$;³ for a given power spectrum multipole of degree ℓ , the total number of FFTs is thus $O(\ell^3)$; for all ℓ s up to some maximum degree ℓ_{\max} , the total remains $O(\ell_{\max}^3)$ as F_n does not need to be recomputed for each ℓ . Ref. [140] has suggested a method to reduce the number of FFTs by breaking down higher-degree Legendre polynomials, for instance \mathcal{L}_4 , into powers of \mathcal{L}_2 which may be redistributed between a pair of δ s in equation (2.26); however, this method was found to increase the variance of the estimator. Instead, ref. [146] has improved this algorithm by noting that it is more natural to decompose Legendre polynomials into orthogonal spherical harmonics $Y_{\ell m}$,

$$\mathcal{L}_\ell(\hat{\mathbf{k}} \cdot \hat{\mathbf{s}}) \equiv \frac{4\pi}{2\ell+1} \sum_{m=-\ell}^{\ell} Y_{\ell m}(\hat{\mathbf{k}}) Y_{\ell m}^*(\hat{\mathbf{s}}), \quad (2.40)$$

² Note that the use of the notation $\xi_\ell^{(i)}$ agrees with that in ref. [142] but is reserved for a different yet related quantity in ref. [78]. Many of the formulæ contained in these references have been originally derived for various wide-angle estimators for the correlation function; see e.g. refs. [143, 144].

³ The number of FFTs here is equal to the number of unique terms in a trinomial expansion, since $(\hat{\mathbf{k}} \cdot \hat{\mathbf{s}})^n$ is a power of trinomials in three dimensions.

so one can instead compute

$$F_\ell(\mathbf{k}) = \frac{4\pi}{2\ell+1} \sum_{m=-\ell}^{\ell} Y_{\ell m}(\hat{\mathbf{k}}) \int d^3\mathbf{s} e^{-i\mathbf{k}\cdot\mathbf{s}} Y_{\ell m}^*(\hat{\mathbf{s}}) F(\mathbf{s}) \quad (2.41)$$

by $2\ell+1 = O(\ell)$ FFTs, and thus obtain the Yamamoto estimator by

$$\hat{P}_\ell(k) = (2\ell+1) \int \frac{d^2\hat{\mathbf{k}}}{4\pi} F_0^*(\mathbf{k}) F_\ell(\mathbf{k}). \quad (2.42)$$

The total number of FFTs for all ℓ s up to ℓ_{\max} is therefore only $O(\ell_{\max}^2)$.

The FFT-based Yamamoto estimator is currently the standard summary statistics for galaxy clustering analysis; it was adopted in the most recent eBOSS analyses such as refs. [81, 103, 147].⁴

2.3 Window convolution of models

It was shown earlier in equation (2.8) or (2.15) that in general, the two-point correlator of the weighted over-density field F is related to the underlying power spectrum P but convolved with the function $|G_{1,1}|^2$ and layered with shot noise $G_{2,1}$. Whilst the shot noise is additive and can be either subtracted from the Yamamoto estimator or included in model predictions (possibly as a nuisance parameter to be marginalised over), the convolution with a window function changes both the amplitude and shape of the power spectrum and is thus more difficult to be undone. Ref. [149] argues that deconvolution, an inversion process, generally amplifies statistical noise, and hence a forward-modelling approach is more desirable. However, the convolution of a model power spectrum naïvely evaluated as a Riemannian integral by FFTs is not only slow but may also suffer from discrete sampling effects such as aliasing. Although these issues also affect power spectrum measurements on smaller scales, FFTs are only performed once; in contrast, convolution needs to be repeated whenever the base model changes, and the mixing of wavenumbers exacerbates any discrete sampling effects. Furthermore, Cartesian grids used for FFTs do not provide a regular sampling of the angular variable μ of the anisotropic power spectrum.

In light of these challenges, ref. [149] has devised a technique for the window convolution of power spectrum models albeit performed in configuration space. To avoid cluttered notation, here the FKP galaxy over-density field is simply written as

$$F(\mathbf{s}) = W(\mathbf{s}) \delta(\mathbf{s}), \quad (2.43)$$

where the generalised window function W encompasses the survey window (eq. 1.24), the weighting scheme w and any normalisation factors. In configuration space, the windowed correlation function is

$$\tilde{\xi}(\mathbf{s}) = \langle F(\mathbf{x}) F(\mathbf{x} + \mathbf{s}) \rangle = Q(\mathbf{s}) \xi(\mathbf{s}), \quad (2.44)$$

⁴ Some previous analyses have also considered power spectrum ‘wedges’ [148], i.e. the anisotropic power spectrum $P(k, \mu)$ binned in μ , which allow certain anisotropic modes affected by survey systematics to be removed.

where the tilde $\tilde{\cdot}$ denotes a windowed quantity, and the function Q is given as a volume average,

$$Q(\mathbf{s}) = \int d^3\mathbf{x} W(\mathbf{x}) W(\mathbf{x} + \mathbf{s}) = \int \frac{d^3\mathbf{k}}{(2\pi)^3} e^{i\mathbf{k}\cdot\mathbf{s}} |W(\mathbf{k})|^2, \quad (2.45)$$

by invoking ergodicity (see §1.1).^{5,6} Decomposing $\tilde{\xi}$ and Q into Legendre multipoles defined for the cosine of the angle between the pairwise line of sight and the separation vector, $\mu_{\mathbf{s}} = \hat{\mathbf{d}} \cdot \hat{\mathbf{s}}$, one finds

$$\tilde{\xi}_{\ell}(s) = \frac{2\ell + 1}{2} \iint d\mu_{\mathbf{s}} \frac{d\varphi_{\mathbf{s}}}{2\pi} \mathcal{L}_{\ell}(\mu_{\mathbf{s}}) \sum_{\ell_1} Q_{\ell_1}(s) \mathcal{L}_{\ell_1}(\mu_{\mathbf{s}}) \sum_L \tilde{\xi}_L(s) \mathcal{L}_L(\mu_{\mathbf{s}}), \quad (2.46)$$

where $\varphi_{\mathbf{s}}$ is the azimuthal angle of the separation vector \mathbf{s} around the line of sight \mathbf{d} . By further decomposing products of Legendre polynomials as

$$\mathcal{L}_{\ell}(\mu_{\mathbf{s}}) \mathcal{L}_L(\mu_{\mathbf{s}}) = \sum_{\ell_2} A_{\ell L}^{\ell_2} \mathcal{L}_{\ell_2}(\mu_{\mathbf{s}}), \quad (2.47)$$

one obtains

$$\tilde{\xi}_{\ell}(s) = (2\ell + 1) \sum_L \left[\sum_{\ell_1} \frac{1}{2\ell_1 + 1} A_{\ell L}^{\ell_1} Q_{\ell_1}(s) \right] \tilde{\xi}_L(s). \quad (2.48)$$

The coefficients $A_{\ell L}^{\ell_1}$ can be found from the identity

$$\mathcal{L}_{\ell} \mathcal{L}_L \equiv \sum_{\ell'=0}^{\min\{\ell, L\}} \frac{2\ell + 2L - 4\ell' + 1}{2\ell + 2L - 2\ell' + 1} \frac{C_{\ell-\ell'} C_{\ell'} C_{L-\ell'}}{C_{\ell+L-\ell'}} \mathcal{L}_{\ell+L-2\ell'}, \quad (2.49)$$

where $C_{\ell} = (2\ell - 1)!!/p! \equiv (2\ell - 1)(2\ell - 3) \cdots 1/p!$ [150]. For the monopole, quadrupole and hexadecapole with $\ell = 0, 2, 4$ in the Kaiser RSD model (see §1.2), the results can be written as a matrix multiplication [79, 149],

$$\begin{pmatrix} \tilde{\xi}_0 \\ \tilde{\xi}_2 \\ \tilde{\xi}_4 \end{pmatrix} = \sum_{i=0}^4 Q_{2i} \mathbf{M}_{2i} \begin{pmatrix} \xi_0 \\ \xi_2 \\ \xi_4 \end{pmatrix}, \quad (2.50)$$

where the matrices \mathbf{M}_{2i} are $\mathbf{M}_0 = \mathbf{I}_3$ (the identity matrix) and

$$\mathbf{M}_2 = \begin{pmatrix} 0 & 1/5 & 0 \\ 1 & 2/7 & 2/7 \\ 0 & 18/35 & 20/77 \end{pmatrix}, \quad \mathbf{M}_4 = \begin{pmatrix} 0 & 0 & 1/9 \\ 0 & 2/7 & 100/693 \\ 1 & 20/77 & 162/1001 \end{pmatrix}, \quad (2.51)$$

$$\mathbf{M}_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 25/143 \\ 0 & 45/143 & 20/143 \end{pmatrix}, \quad \mathbf{M}_8 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 490/2431 \end{pmatrix}.$$

To link the windowed correlation function to the expectation of the Yamamoto estimator (eq. 2.26), one may use the plane wave expansion

$$e^{-i\mathbf{k}\cdot\mathbf{s}} \equiv \sum_{\ell} (2\ell + 1) i^{-\ell} j_{\ell}(ks) \mathcal{L}_{\ell}(\mathbf{k} \cdot \mathbf{s}) \quad (2.52)$$

⁵ Note that W itself is non-stochastic, and $Q = W^2$ is analogous to the function $G_{1,1}$ in equation (2.8).

⁶ No volume normalisation factor is needed here if the FKP normalisation factor I (eq. 2.2) is included, in which case the physical dimension of the generalised window W is $[W] = L^{-3/2}$, and thus $[Q] = 1$.

to write

$$\begin{aligned}
 \langle \widehat{P}_\ell(k) \rangle &= \frac{2\ell+1}{2} \sum_{\ell_1} (2\ell_1+1) i^{-\ell_1} \iint d\mu_k \frac{d\varphi_k}{2\pi} \int d^3\mathbf{x} \int d^3\mathbf{s} j_{\ell_1}(ks) \\
 &\quad \times \mathcal{L}_{\ell_1}(\hat{\mathbf{k}} \cdot \hat{\mathbf{s}}) \mathcal{L}_\ell(\hat{\mathbf{k}} \cdot \hat{\mathbf{d}}) W(\mathbf{x}) W(\mathbf{x} + \mathbf{s}) \sum_L \xi_L(s) \mathcal{L}_L(\hat{\mathbf{s}} \cdot \hat{\mathbf{d}}) \\
 &= (2\ell+1) i^{-\ell} \int d^3\mathbf{s} j_\ell(ks) \sum_L \xi_L(s) \sum_{\ell_2} A_{\ell L}^{\ell_2} \int d^3\mathbf{x} W(\mathbf{x}) W(\mathbf{x} + \mathbf{s}) \mathcal{L}_{\ell_2}(\hat{\mathbf{s}} \cdot \hat{\mathbf{d}}) \\
 &= 4\pi i^{-\ell} \int s^2 ds j_\ell(ks) (2\ell+1) \sum_L \xi_L(s) \left[\sum_{\ell_2} \frac{1}{2\ell_2+1} A_{\ell L}^{\ell_2} Q_{\ell_2}(s) \right] \\
 &= 4\pi i^{-\ell} \int s^2 ds j_\ell(ks) \widetilde{\xi}_\ell(s), \tag{2.53}
 \end{aligned}$$

where to derive the second line, the orthogonality condition for Legendre polynomials and equation (2.47) are used; to derive the third line, the definition of the multipole Q_ℓ is used; and to derive the last line, equation (2.48) is used.

The derivation above suggests the following procedure for convolving any model of the power spectrum multipoles P_ℓ with the window function $Q = W^2$:

- 1) perform the inverse spherical Bessel transform (commonly referred to as the Hankel transform) of P_ℓ into the correlation function multipoles,⁷

$$\xi_\ell(s) = 4\pi i^\ell \int k^2 dk j_\ell(ks) P_\ell(k); \tag{2.54}$$

- 2) couple the multipoles ξ_ℓ with the window function using equation (2.48), and thus obtain the windowed correlation function multipoles $\widetilde{\xi}_\ell$;
- 3) perform the spherical Bessel transform of $\widetilde{\xi}_\ell$ into the windowed model of power spectrum multipoles,

$$\widetilde{P}_\ell(k) = 4\pi i^{-\ell} \int s^2 ds j_\ell(ks) \widetilde{\xi}_\ell(s). \tag{2.55}$$

There are two technicalities in the procedure above. The first is the computation of the window function multipoles $Q_\ell(s)$. These can be determined by a pair counting approach used in correlation function measurements [151], as suggested by ref. [149]. Alternatively, one could still employ the FFT-based Yamamoto estimator to determine $Q_\ell(k)$ by measuring the power spectrum of a synthetic random catalogue, which possesses the same geometry, sample selection and weighting as the survey catalogue, and is embedded in a large and finely spaced Cartesian grid for better determination at both low and high wavenumbers (compared to the fundamental and Nyquist wavenumbers respectively). This FFT procedure only needs to be performed once, and by a spherical Bessel transform, $Q_\ell(s)$ can then be readily computed [142]. The second technicality concerns the numerical implementation of the spherical Bessel, or Hankel, transform, which may suffer from poor accuracy due to the oscillatory behaviour of spherical Bessel functions j_ℓ . However, a FFT algorithm based on logarithmically-spaced coordinate grid addresses this problem [152], which is numerically implemented by codes such as FFTLOG [153] and MCFIT.

⁷ The actual Hankel transform, also known as the Fourier–Bessel transform, is defined in terms of the Bessel function of the first kind, J_ℓ , with a slightly different measure.

Before one could substitute the power spectrum multipoles \widehat{P}_ℓ measured by the Yamamoto estimator and the window-convolved model \widetilde{P}_ℓ into a likelihood function, there is one final issue which needs to be considered in conjunction with the window function: an integral constraint due to the estimation of the background galaxy number density [154]. Since the background galaxy number density is unknown, the mean number density of the survey catalogue is used as a proxy, and by construction, the measured weighted over-density field \check{F} is shifted from the actual value F ,

$$\check{F}(\mathbf{s}) = F(\mathbf{s}) - \frac{W(\mathbf{s})}{\int d^3\mathbf{x} W(\mathbf{x})} \int d^3\mathbf{x} F(\mathbf{x}), \quad (2.56)$$

so that $\int d^3\mathbf{s} \check{F}(\mathbf{s}) = 0$. In Fourier space, this becomes

$$\check{F}(\mathbf{k}) = F(\mathbf{k}) - \frac{W(\mathbf{k})}{W(\mathbf{k} = \mathbf{0})} F(\mathbf{k} = \mathbf{0}). \quad (2.57)$$

It is evident that the integral constraint forces the zero-wavenumber Fourier mode to vanish, $\check{F}(\mathbf{k} = \mathbf{0}) = 0$, and with the window function, this mode attenuation propagates to Fourier modes at other wavenumbers, in particular at large scales. Therefore one needs to correct the window-convolved model of power spectrum multipoles by [79, 142, 149, 154, 155]

$$\widetilde{P}_\ell(k) \mapsto \check{P}_\ell(k) - \frac{Q_\ell(k)}{Q_\ell(k = 0)} \check{P}_0(0). \quad (2.58)$$

This is actually an approximate prescription merely to ensure that $\lim_{k \rightarrow 0} \check{P}_\ell(k) = 0$, and a more general scheme for integral constraint corrections valid in the local plane-parallel approximation is provided by ref. [156], which also accounts for the fact that the radial selection function may be determined from the survey catalogue as well. Finally, it is worth commenting here that the FKP normalisation factor I (eq. 2.2) ensures that $Q_0(s = 0) = 1$; therefore, if one neglects the factor I when measuring the weighted over-density field, a simple method to ensure that any convolved power spectrum multipoles have the correct amplitude is by normalising the window function multipoles to the value $Q_0(s = 0) = 1$. Similarly, for the integral constraint correction, one could also normalise the window function multipoles to the value $Q_0(k = 0) = 1$.

Power Spectrum Gaussianisation and Covariance Decomposition

At the interface of the confrontation between theoretical models and observational measurements is the probabilistic interpretation connecting the two. Efforts to expand data sets and refine analytic models would thus be in vain if the underpinning probability distribution assumed in a cosmological likelihood analysis is inadequate or outright wrong. In section 1.4, it is discussed how, in the absence of any theoretically motivated probability distributions, one might appeal to asymptotic normality or the principle of indifference to prescribe a normal distribution for cosmological data sets, which are often compressed or ‘binned’ from many independent data points.

The clustering of galaxies as biased tracers of the underlying matter distribution, as discussed in chapter 1, provides important probes of the large-scale structure of the Universe. On large scales where cosmic density fluctuations are described by GRFs, the power spectrum contains all of the information in the galaxy distribution.¹ In practice, the galaxy clustering power spectrum can only be sampled at discrete wavenumbers and measured in wavenumber bands. Except at scales close to the size of the galaxy survey, the band power spectrum is the average square amplitude of many independent clustering modes and hence, by the central limit theorem, can be modelled by the multivariate normal distribution. This has served as the common basis for most galaxy survey analyses in the past [e.g. 79–82, 147]. As discussed in section 1.4, given a theoretical model for the data, the remaining ingredient of a multivariate normal distribution is its covariance matrix, and in the past, many efforts have been devoted to the accurate estimation of covariance matrices subject to limited computational resources. Unbiased covariance matrix estimates are often made from a set of mock catalogues synthesised using algorithms ranging from fast but approximate perturbation theory computations to slow yet detailed N -body simulations, or combinations of the two [157–159] (see also refs. [160–162] for a comprehensive comparison of these methods). One could further reduce the computa-

¹ Of course, a non-zero value of PNG means this is not exactly true, though the level of PNG is expected to be small and its main manifestation is in the clustering amplitude change; see the discussion in section 1.3.

tional burden and enhance the precision of covariance matrix estimates through techniques such as the shrinkage method, which essentially interpolates between empirical estimates and theoretical models [163], or the covariance tapering method [164, 165].

However, as already discussed in section 1.4, there are numerous caveats to using a Gaussian likelihood with an estimated covariance matrix; they are not just limited to the possible need to de-bias the precision matrix and the requirement to account for covariance estimation uncertainties in the inferred parameter constraints. In a frequentists' framework, a Gaussian approximation for the underlying probability distribution may be sufficient when one is primarily interested in the maximum likelihood point with the associated uncertainties given by the curvature of the likelihood in its neighbourhood. However, when asymptotic normality does not apply and one needs to adopt a Bayesian perspective for parameter inference, the overall shape of the probability distribution matters. In this case, even an arbitrarily precise covariance matrix estimate does not guarantee accurate parameter inference owing to non-negligible higher moments in the non-normal distribution [166]. Moreover, as mock catalogues are computationally expensive, they are usually produced at a fixed fiducial cosmological model, whereas in reality, the covariance matrix can depend on the underlying cosmological parameters and therefore must vary with the model being tested; this has adverse implications for likelihood analysis, as demonstrated by ref. [167] in the context of cosmic shear. In summary, failure to account for either the full shape of the likelihood distribution (and thus the posterior distribution) or the parameter dependence of the covariance matrix may diminish the potential of future surveys, which are promised to deliver more accurate and precise cosmological constraints.

There have been recent efforts to go beyond the Gaussian likelihood approximation in various contexts: refs. [166, 168] have considered the likelihood for non-Gaussian fields;² motivated by constraints imposed by non-negativity of the power spectrum, refs. [169–171] have found a transformation that improves the data normality for the bounded configuration-space correlation function which has a non-normal distribution; ref. [172] has considered the gamma distribution for the power spectrum multipoles and the log-normal plus Gaussian approximation, and assessed their effects on parameter estimation in comparison with the Gaussian approximations; for the three-dimensional power spectrum, ref. [173] has rederived the probability distribution of single clustering modes for GRFs, and compared the resulting posterior distribution with approximations inspired by similar studies of the CMB [e.g. 174–177]. However, these earlier works are either limited to univariate or bivariate distributions, or have neglected the window function due to survey geometry and selection effects which can correlate independent clustering modes.

In this chapter, we focus on the power spectrum monopole as the summary statistic of a galaxy clustering analysis in the FKP framework (see §2), though our methodology trivially generalises to other power spectrum multipoles, or indeed any linear combinations of quadratic amplitudes of Gaussian clustering modes.³ The underlying non-normal probability dis-

² These considerations concern the quasi-linear scales which are beyond the scope of this thesis, and are mentioned here as part of a broader context.

³ Although the Gaussianisation technique proposed in this chapter also works with other types of data such as the bispectrum, the precise form of the Gaussianising transformation required is non-trivial to derive and different from the case of two-point statistics considered here.

tribution is rederived for the windowed galaxy clustering power spectrum of a Poisson sampled GRF, though it is analytically too complex for practical evaluations. Therefore the multivariate normal distribution is reinstated through a Gaussianising transformation, or *Gaussianisation*, which improves the data normality. We also propose the variance–correlation decomposition to capture the cosmological dependence of the covariance matrix in the linear perturbative regime using a rescaling method. It is worth emphasising that the data Gaussianisation procedure is a non-linear transformation, so the expectation value of the resulting data vector is not equal to the same transformation applied to the original model for the untransformed power spectrum measurements; instead, the procedure is simply designed to give a likelihood possessing a Gaussian form with respect to the data vector. To test the new methodology and demonstrate its advantages over the standard likelihood treatments, we consider the problem of inferring the local PNG parameter f_{NL} and compare the different posterior distributions with the true one for simulated data sets.

3.1 Distribution of power spectrum measurements

As elaborated in both chapters 1 and 2, clustering measurements are usually made for the power spectrum multipoles in a galaxy survey analysis. As seen from equation (2.26), the Yamamoto estimator \hat{P}_ℓ for power spectrum multipoles has a functional form that is quadratic in the clustering mode δ . Barring the issue of mode coupling due to RSD, i.e. assuming the distant-observer limit in redshift space, $\langle \delta\delta^* \rangle$ is diagonal; $\hat{P}_\ell(k)$ is then essentially a linear combination of the quadratic amplitudes $|\delta(\mathbf{q})|^2$ of clustering modes. For simplicity without loss of generality, we shall consider measurements of the power spectrum monopole or equivalently the spherically averaged band power, $\tilde{Y} = (\hat{P}_0(k_a))_{a=1}^d$, which include the effect of the survey window (and any weighting schemes etc.) and are thus denoted with a tilde. The data vector \tilde{Y} in effect comes from a linear transformation,

$$\tilde{Y} = \mathbf{B}Y, \quad (3.1)$$

where $\mathbf{B} \in \mathbb{R}^{d \times d'}$ is the mixing matrix, and the random vector $Y = (|\delta(\mathbf{q}_i)|^2)_{i=1}^{d'}$ consists of quadratic amplitudes of independent clustering modes that can be sampled on a FFT grid, i.e. measurements of the underlying mode power in the absence of window effects. The mixing matrix is in nature a representation of the window function: by considering the discrete form of equation (2.8), one can see that the elements of the mixing matrix are

$$(\mathbf{B})_{ai} = \frac{1}{VN_{k_a}} \sum_{\mathbf{k} \in S_{k_a}} |G_{1,1}(\mathbf{k} - \mathbf{q}_i)|^2, \quad (3.2)$$

where V is the volume of a rectangular domain on which the periodic boundary condition is applied (see §1.1), and N_{k_a} is the number of clustering modes sampled in a spherical shell S_{k_a} at the a th band wavenumber k_a (see also §2.1). To account for the Poisson shot noise power (see §1.1), one simply adds a random vector $\boldsymbol{\zeta}$ so that

$$\tilde{Y} = \mathbf{B}(Y + \boldsymbol{\zeta}). \quad (3.3)$$

For a single Fourier mode of a GRF, its real and imaginary parts independently follow the same normal distribution, $\text{Re } \delta(\mathbf{q}), \text{Im } \delta(\mathbf{q}) \sim \text{N}(0, VP(\mathbf{q})/2)$. By considering the relation $\delta(\mathbf{q}) \equiv \text{Re } \delta(\mathbf{q}) + i \text{Im } \delta(\mathbf{q}) \equiv |\delta(\mathbf{q})| \exp[-i \arg \delta(\mathbf{q})]$ as a transformation of variables,

one can readily derive the PDF for the mode amplitude $|\delta(\mathbf{q})|$ which follows the Rayleigh distribution,

$$\mathbb{p}(|\delta(\mathbf{q})|) = \frac{2|\delta(\mathbf{q})|}{VP(\mathbf{q})} \exp\left[-\frac{|\delta(\mathbf{q})|^2}{VP(\mathbf{q})}\right]. \quad (3.4)$$

The same derivation also shows that the phase of the Fourier mode, $\arg \delta$, follows a uniform random distribution. The power spectrum estimated from the amplitude of a single clustering mode, $\hat{P} = V^{-1}|\delta|^2$, is therefore described by the exponential distribution with the PDF

$$\mathbb{p}(\hat{P}) = P^{-1} \exp(-\hat{P}/P). \quad (3.5)$$

Before the probability distribution of the power spectrum data vector \tilde{Y} can be derived, the distribution of the shot noise power ζ also needs to be determined. Following the argument of ref. [154], one could consider N galaxies randomly located at positions $\{\mathbf{x}_j\}_{j=1}^N$ in some volume. The over-density field in Fourier space is then (see also eq. 1.19)

$$\delta(\mathbf{q}) = \frac{1}{N} \sum_{j=1}^N e^{i\mathbf{q}\cdot\mathbf{x}_j} \quad (\mathbf{q} \neq 0), \quad (3.6)$$

which, by the central limit theorem, converges in distribution to a normal random variable in the large number limit $N \rightarrow \infty$ regardless of the distribution of individual summands $\exp(i\mathbf{q}\cdot\mathbf{x}_j)$ where $\{\mathbf{x}_j\}$ are independent. Therefore shot noise also behaves like a GRF, and each component of ζ follows an exponential distribution with mean value P_{shot} by the argument above.

Since both the clustering mode power and the shot noise power are exponentially distributed, components of the band power spectrum data vector \tilde{Y} are linear combinations of exponential random variables by equation (3.3); to determine the probability distribution of these components, it is then instructive to consider generic independent exponential random variables $\{X_j \sim \text{Exp}(\lambda_j)\}_{j=1}^N$ with the PDF

$$\mathbb{p}_j(X_j) = \lambda_j \exp(-\lambda_j X_j), \quad (3.7)$$

and derive the distribution of the sum $X = \sum_j X_j$. Here $\lambda_j^{-1} = \mathbb{E}[X_j]$ is the mean, and λ_j is also the inverse of the scale parameter β_j .⁴ By considering the characteristic function of probability distributions, it follows that the PDF of the sum X is the convolution of individual exponential PDFs,

$$\mathbb{p}(X) = (\mathbb{p}_1 * \cdots * \mathbb{p}_N)(X) = \sum_{j=1}^N \mathbb{p}_j(X) \prod_{\substack{m \neq j \\ m=1}}^N \frac{\lambda_m}{\lambda_m - \lambda_j}, \quad (3.8)$$

where the second equality can be established using proof by induction. This is an example of the *hypo-exponential distribution*, sometimes also referred to as the generalised Erlang distribution [178]. One should note the particular case where $\lambda_m = \lambda_j$ for some $m \neq j$, when

⁴ A family of probability distributions can sometimes be parametrised by a scale parameter, which determines the dispersion or spread of the probability distribution. If β is a scale parameter, then the cumulative distribution function $F(x; \beta)$ satisfies $F(x; \beta) \equiv F(x/\beta; 1)$. For exponential distributions, the mean is the scale parameter; the gamma distribution, which is discussed later, is also commonly parametrised with a scale parameter.

two random variables X_m and X_j are identically distributed. By the result above, in the limit $\Delta\lambda = \lambda_m - \lambda_j \rightarrow 0$, the PDF of $X_{mj} = X_m + X_j$ is

$$\mathbb{P}(X_{mj}) = \lim_{\Delta\lambda \rightarrow 0} \frac{\lambda_m(\lambda_m - \Delta\lambda)}{\Delta\lambda} e^{-\lambda_m X_{mj}} (1 - e^{X_{mj}\Delta\lambda}) = \lambda_m^2 X_{mj} e^{-\lambda_m X_{mj}}, \quad (3.9)$$

which one recognises as the PDF of a gamma distribution $\Gamma(2, \lambda_m^{-1})$ in the shape–scale parametrisation. By this general derivation, it is now clear that components of the data vector \tilde{Y} , which are exponential mixtures, follow the hypo-exponential distribution with the PDF

$$\mathbb{P}(\tilde{Y}_a; \{\beta_I\}) = \sum_I \left(\prod_{I' \neq I} \frac{1}{1 - \beta_{I'}/\beta_I} \right) \beta_I^{-1} \exp(-\beta_I^{-1} \tilde{Y}_a) \quad (3.10)$$

parametrised by the positive scale parameters

$$\beta_I \in \{B_{ai} P(\mathbf{q}_i), B_{ai} P_{\text{shot}} : i = 1, \dots, d'\}, \quad (3.11)$$

which are the windowed contributions of the underlying clustering mode power and shot noise power to the a th band. It is understood that a well-defined limit is taken above in the case $\beta_I = \beta_{I'}$. By the central limit theorem, when the number of modes $\delta(\mathbf{q}_i)$ contributing to the a th band power (with $B_{ai} \neq 0$) is large, the hypo-exponential variable \tilde{Y}_a converges in distribution to a normal variable—this is the basis for the normality assumption often used in power spectrum analyses. However, on the largest scales in a survey where the number of clustering modes is the fewest, the underlying hypo-exponential distribution deviates from the normal distribution, and hence it would be erroneous to use the Gaussian likelihood for cosmological parameter inference.

Although the hypo-exponential distribution is the correct description of power spectrum measurements, it is difficult to manipulate as it involves many scale parameters β_I instead of just a few summary parameters. A robust approximation is the exponentially modified gamma distribution which has one shape parameter R and two scale parameters (η_1, η_2) [179], but determining these parameters involves solving a cubic algebraic equation, which is still numerically cumbersome. We suggest an even simpler approximation using the gamma distribution with the PDF

$$\mathbb{P}_\Gamma(\tilde{Y}_a; R, \eta) = \frac{\eta^{-R}}{\Gamma(R)} \tilde{Y}_a^{R-1} \exp(-\tilde{Y}_a/\eta), \quad (3.12)$$

where by matching the mean and variance of the hypo-exponential distribution, the effective shape and scale parameters,

$$R = \frac{\mathbb{E}[\tilde{Y}_a]^2}{\text{var}[\tilde{Y}_a]} = \frac{(\sum_I \beta_I)^2}{\sum_I \beta_I^2} \quad \text{and} \quad \eta = \frac{\text{var}[\tilde{Y}_a]}{\mathbb{E}[\tilde{Y}_a]} = \frac{\sum_I \beta_I^2}{\sum_I \beta_I}, \quad (3.13)$$

can be determined. In place of the mean and variance parameters of a normal distribution $\mathcal{N}(R\eta, R\eta^2)$ with the PDF

$$\mathbb{P}_\mathcal{N}(\tilde{Y}_a; R\eta, R\eta^2) = \frac{1}{\sqrt{2\pi R\eta^2}} \exp\left[-\frac{(\tilde{Y}_a - R\eta)^2}{2R\eta^2}\right], \quad (3.14)$$

the shape–scale parameters (R, η) determine the non-normal distribution of band power spectrum measurements. They have a natural interpretation in this context: the shape parameter R is the effective number of independent clustering modes contributing to the wavenumber band, and the scale parameter η is the effective windowed power in that band including shot noise. In the limiting case that there is only a single clustering mode in a wavenumber band (e.g. pure shot noise), the gamma distribution coincides with the hypo-exponential distribution, both of which reduce to the exponential distribution. Henceforth, we shall assume components of the power spectrum data vector \tilde{Y} follow the gamma distribution with shape–scale parameters (R_a, η_a) for the a th component. In principle, the shape–scale parameters (R, η) can be calculated for each wavenumber band using analytical models of the band power expectation $\mathbb{E}[\tilde{Y}_a]$ and variance $\text{var}[\tilde{Y}_a]$ with knowledge of the mixing matrix \mathbf{B} and FFT-sampled wavevectors \mathbf{q}_i , though these quantities are not always available. In practice, window convolution of power spectrum models is often performed in configuration space and then Hankel transformed back into Fourier space (see §2.3), and the covariance matrix for power spectrum measurements is estimated from mock catalogues.

Even with the simplifying approximation for the univariate distribution of band power measurements as a gamma distribution, the full multivariate distribution for the data vector \tilde{Y} remains intractable, since its components are not independent as a result of the mixing of clustering modes. One might consider $\mathbf{Y} \xrightarrow{\mathbf{B}} \tilde{\mathbf{Y}}$ as a transformation of random vectors with the Jacobian determinant $\det(\mathbf{B}\mathbf{B}^\top)^{1/2}$, where the distribution function of \mathbf{Y} is just the product of independent exponential components. However, the mixing matrix \mathbf{B} is generally non-square ($d < d'$), and expressing the transformed probability distribution explicitly in terms of the band power data vector $\tilde{\mathbf{Y}}$ requires the inversion of \mathbf{B} . One could perhaps introduce $(d' - d)$ ‘helper’ components in $\tilde{\mathbf{Y}}$ and pad \mathbf{B} into a square matrix before eventually marginalising out these additional helper components. Unfortunately, the linear transformation induced by the padded square matrix would map the domain of the random vector \mathbf{Y} from $\mathbb{R}_+^{d'}$ to a different domain in \mathbb{R}^d : whereas the random vector \mathbf{Y} spans the positive orthant (a higher-dimensional quadrant) as its components being the mode power are non-negative, the transformed random vector generally spans an irregularly-shaped domain, which makes the marginalisation over the helper components difficult and susceptible to ‘the curse of dimensionality’, especially when $d \gg d'$ for massive data compression.

3.2 Beyond standard Gaussian likelihoods

3.2.1 Data Gaussianisation

Instead of attempting to determine the exact full multivariate distribution for the windowed power spectrum measurements $\tilde{\mathbf{Y}}$, we subscribe to a component-wise Gaussianisation strategy and reinstate the multivariate normality assumption for the Gaussianised data vector $\mathbf{Z} \leftarrow \tilde{\mathbf{Y}}$. The reasoning behind this is two-fold. First, each component of the random vector \mathbf{Z} is now certainly univariate normal, as should be the case for a bona fide multivariate normal distribution. Secondly, if the covariance matrix of the data vector has a narrow-band structure and the cross-correlation between components is weak, then univariate Gaussianisation is approximately equivalent to multivariate Gaussianisation; this can be achieved with a suitable bin-

ning strategy where the wavenumber band widths are chosen to be greater than the correlation length of the Fourier-space window. Although sophisticated full Gaussianisation schemes exist [e.g. 180], they are often based on the empirical data at hand and iterative in nature, and thus may not be appropriate for forward modelling of theoretical predictions; therefore we leave more advanced multivariate Gaussianisation methods for future considerations.

In the case of univariate Gaussianisation, it is possible to find the precise transformation $\tilde{Y} \mapsto Z$ where $\tilde{Y} \sim \Gamma(R, \eta)$ is a gamma random variable and $Z \sim \mathcal{N}(0, 1)$ is the standard normal random variable (vector indices are irrelevant and thus suppressed for notational brevity). This requires the matching of their cumulative distribution functions,

$$\int_0^{\tilde{Y}} dy \mathbb{P}_\Gamma(y; R, \eta) = \int_{-\infty}^Z dz \mathbb{P}_\mathcal{N}(z; 0, 1), \quad (3.15)$$

which leads to the error function transformation,

$$Z = -\sqrt{2} \operatorname{erfc}^{-1} \left[2 \frac{\gamma(R, \tilde{Y}/\eta)}{\Gamma(R)} \right], \quad (3.16)$$

where erfc^{-1} is the inverse complementary error function and γ is the lower incomplete gamma function. However, similar to the cause of the AP effect, the true cosmology underlying the band power measurement \tilde{Y} is unknown and thus R and η determined in a fiducial cosmological model may not result in perfect Gaussianisation. This then necessitates the recalculation of the mean and variance of the supposedly standard normal random variable Z for each cosmological model to be fitted to the data, which involves computationally intensive numerical integrations.

We instead propose a simple Gaussianisation scheme which has been applied to cosmological analyses before in e.g. refs. [181, 182] though implemented differently in different contexts: we define the *Box–Cox transformation* [183] to be

$$Z = \tilde{Y}^\nu, \quad (3.17)$$

where the transformation parameter ν is positive to ensure regularity.⁵ This transformation achieves *approximate* Gaussianisation by essentially suppressing higher-order moments of the data; in our case, the transformation involves the shape–scale parameters $(\check{R}, \check{\eta})$ of the gamma distribution, which are denoted with a breve as their values are determined by the fixed power spectrum model \check{P} , including Poisson shot noise, in a fiducial cosmological model with parameters $\theta = \check{\theta}$. The PDF of the Gaussianised random variable Z is

$$\mathbb{P}(Z; \check{R}, \check{\eta}) = |\mathbf{J}_Z(\tilde{Y})| \frac{\check{\eta}^{-\check{R}}}{\Gamma(\check{R})} \tilde{Y}^{\check{R}-1} \exp(-\tilde{Y}/\check{\eta}), \quad (3.18)$$

where $\mathbf{J}_Z(\tilde{Y}) = \nu \tilde{Y}^{\nu-1}$ is the Jacobian determinant of the transformation. The K th moment of Z is given by

$$\mathbb{E}[Z^K] = \frac{\Gamma(\check{R} + K\nu)}{\Gamma(\check{R})} \check{\eta}^{K\nu}, \quad (3.19)$$

⁵ Note that the letters R, ν here and the letters K, μ below denote different physical quantities outside the scope of this chapter.

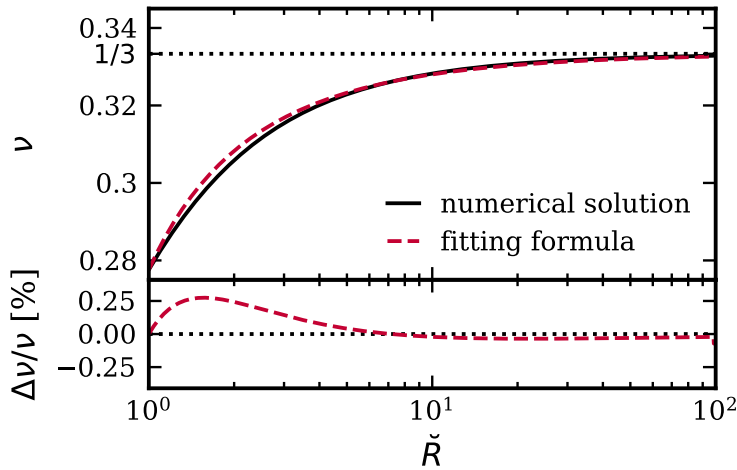


Figure 3.1. Solution of the Box–Cox transformation parameter ν as a function of the fiducial gamma distribution shape parameter \check{R} . The numerical solution to equation (3.21) (solid line) is compared with the fitting formula (eq. 3.24) (dash-dotted line) in the top panel. The relative error of the fitting formula (dashed line) is shown in the bottom panel. Figure adapted from Fig. 1 in Paper I.

and thus its mean and variance are

$$\mu_Z(\check{R}, \check{\eta}) = \frac{\Gamma(\check{R} + \nu)}{\Gamma(\check{R})} \check{\eta}^\nu \quad (3.20a)$$

$$\text{and } \sigma_Z^2(\check{R}, \check{\eta}) = \frac{\Gamma(\check{R} + 2\nu) \Gamma(\check{R}) - \Gamma(\check{R} + \nu)^2}{\Gamma(\check{R})^2} \check{\eta}^{2\nu}. \quad (3.20b)$$

To determine the transformation parameter ν , one demands the third central moment of Z to vanish,

$$\begin{aligned} 0 &= \mathbb{E}[Z^3] - 3 \mathbb{E}[Z^2] \mathbb{E}[Z] + 2 \mathbb{E}[Z]^3 \\ &= \left\{ \frac{\Gamma(\check{R} + 3\nu)}{\Gamma(\check{R})} - 3 \frac{\Gamma(\check{R} + 2\nu)}{\Gamma(\check{R})} \frac{\Gamma(\check{R} + \nu)}{\Gamma(\check{R})} + 2 \left[\frac{\Gamma(\check{R} + \nu)}{\Gamma(\check{R})} \right]^3 \right\} \check{\eta}^{3\nu}. \end{aligned} \quad (3.21)$$

The numerical solution of ν as a function of \check{R} is shown in Figure 3.1. The observed asymptotic behaviour, namely $\nu \rightarrow 1/3$ as $\check{R} \rightarrow \infty$, can be understood by considering the expansion of the gamma function ratio [184]

$$\frac{\Gamma(A+B)}{\Gamma(A)} \sim A^B \left[1 + \frac{(B-1)B}{2} A^{-1} + \frac{(3B-1)(B-2)(B-1)B}{24} A^{-2} + \dots \right], \quad (3.22)$$

so as $\check{R} \rightarrow \infty$ with $\check{R}\check{\eta} < \infty$ (i.e. finite mean), equation (3.21) becomes

$$\check{R}^{-2} (3\nu - 1) \nu^3 = 0. \quad (3.23)$$

The non-trivial solution is indeed $\nu = 1/3$. For the purpose of Gaussianisation, however, the precise value of ν matters less as \check{R} becomes large owing to the suppression factor \check{R}^{-2} in the

third central moment—this is essentially a manifestation of asymptotic normality. It is useful to present an empirical fitting formula for the numerical solution to equation (3.21):

$$\nu \approx \frac{1}{3} + 0.042 \left[1 - \exp(0.85 / \check{R}^{0.85}) \right]. \quad (3.24)$$

Comparison of this fitting formula to the numerical solution in Figure 3.1 shows that it performs well, being accurate to sub-percent levels. In fact, adopting the fixed value $\nu = 1/3$ works well in realistic situations, which indicates that our Gaussianisation scheme is robust to variation in ν and thus \check{R} , so that it possesses the desirable property of being insensitive to the choice of fiducial cosmology.

It is worth noting that our Gaussianisation scheme differs from the approaches taken in some CMB studies in that they usually seek an approximation to the probability distribution or likelihood function at the maximum point: for instance, ref. [185] makes a Gaussian approximation to the likelihood function by employing the Fisher information matrix in place of the curvature matrix around the peak; ref. [174] combines a log-normal likelihood with a Gaussian one by adding them logarithmically; ref. [177] writes the log-likelihood as a quadratic form in a variable that is a non-linear transformation of the data; interestingly, ref. [176] proposes a power-law transformation of the estimated angular power spectrum with the exponent being $1/3$ —this should be seen as an analogue to the limiting case of our Gaussianisation scheme with $\nu = 1/3$, where the likelihood is only weakly non-Gaussian when the shape parameter R is large. In contrast to the motivation behind all these approaches, our Gaussianising transformation targets the overall shape of the distribution by suppressing higher-order moments instead of seeking a local approximation.

3.2.2 Covariance treatment

The second issue to be addressed in standard likelihood analyses is the neglect of cosmological parameter dependence of the covariance matrix. Because of the reliance on mock catalogues for covariance estimation, it is computationally expensive to produce covariance matrix estimates that vary with the cosmological model being fitted. However, if one were able to separate and analytically predict the cosmologically varying part of the covariance matrix, then its estimate could be rescaled accordingly with the cosmological model.

A generic covariance matrix Σ can be decomposed into a variance part and a correlation part,

$$\Sigma = \Lambda \mathbf{C} \Lambda, \quad (3.25)$$

where the diagonal matrix Λ consists of standard deviations, $(\Lambda)_{aa} = (\Sigma)_{aa}^{1/2}$, and \mathbf{C} is the correlation matrix. This decomposition offers the following insight: for power spectrum measurements on large scales, whether Gaussianised or not, the off-diagonal correlation in \mathbf{C} is solely induced by window effects as manifested by the mixing matrix \mathbf{B} , which should not depend on the cosmological model being tested through the power spectrum.⁶ This insight makes the variance–correlation decomposition particularly useful, for the decomposition into Λ and \mathbf{C} is precisely the separation of cosmological dependence from the cosmology-independent part in

⁶ A minor caveat is that in redshift space, the window function may be slightly dependent on the fiducial cosmology overall because, for instance, the survey boundary is defined by a maximum redshift.

the covariance matrix Σ . Therefore one may obtain a covariance matrix estimate $\hat{\Sigma}_f = \check{\Lambda} \hat{C} \check{\Lambda}$ from mock catalogues produced at fiducial cosmological parameters $\theta = \check{\theta}$ with the fiducial power spectrum model \check{P} , and rescale this estimate using the diagonal matrix Λ to allow for varying cosmological models,

$$\hat{\Sigma}(\theta) = \Lambda(\theta) \check{\Lambda}^{-1} \hat{\Sigma}_f \check{\Lambda}^{-1} \Lambda(\theta). \quad (3.26)$$

For instance, for the Gaussianised power spectrum measurements Z at cosmological parameters θ , the entries of the diagonal matrix Λ are

$$(\Lambda(\theta))_{aa} = \text{var}[Z_a] = \sigma_Z^2(R_a, \eta_a) \quad (3.27)$$

as given by equation (3.20), where the gamma distribution shape–scale parameters (R_a, η_a) for each component depend on cosmological parameters θ through the power spectrum model (see eqs. 3.11 and 3.13).

As one recalls from section 1.4, covariance matrices estimated from mock catalogues are inherently random. Instead of directly substituting the estimated covariance in a likelihood function, a more principled Bayesian approach is to marginalise out the unknown underlying covariance matrix [121, SH], which changes the Gaussian likelihood based on the normal distribution to a likelihood based on the modified Student’s t -distribution given by equation (1.73). Fortuitously but unsurprisingly, our covariance decomposition and rescaling method to account for parameter dependence does not alter the derivation of SH’s modified t -distribution likelihood. This is a consequence of the invariance property of Wishart and inverse Wishart distributions W_d and W_d^{-1} (of data dimension d): given a non-singular matrix $\mathbf{A} \in \mathbb{R}^{d \times d}$, if $\mathbf{W} \sim W_d(\mathbf{V}, m)$, then $\mathbf{A}^T \mathbf{W} \mathbf{A} \sim W_d(\mathbf{A}^T \mathbf{V} \mathbf{A}, m)$; if $\mathbf{W} \sim W_d^{-1}(\mathbf{V}^{-1}, m)$, then $\mathbf{A}^T \mathbf{W} \mathbf{A} \sim W_d^{-1}(\mathbf{A}^T \mathbf{V}^{-1} \mathbf{A}, m)$ [118]. Here the random matrix \mathbf{W} could be either the covariance matrix estimate (the former case) or the precision matrix estimate (the latter case); the corresponding scale matrix \mathbf{V} is then either the true covariance matrix or the true precision matrix; the number of degrees of freedom, $m = N_{\text{mock}} - 1$, is given by the number of mock catalogues, N_{mock} , used in the estimation. In addition to the stochasticity introduced by covariance matrix estimation for the Gaussianised data Z , if one estimates the gamma distribution shape–scale parameters (R_a, η_a) (see eq. 3.13) from the fiducial covariance matrix estimate of the un-Gaussianised data \tilde{Y} , then these parameters are also random and should be treated jointly with the covariance matrix estimate in SH’s approach. However, as Monte Carlo tests in the next section will show, this is not an issue in practice since the scatter in the estimated covariance matrix dominates over the uncertainties of the estimated shape and scale parameters.

3.2.3 A novel likelihood function

With data Gaussianisation, covariance estimate rescaling via the variance–correlation decomposition, and SH’s covariance matrix marginalisation approach, we suggest the following likelihood function for galaxy clustering power spectrum measurements:

$$\mathcal{L}(\theta; \mathbf{Z}, \boldsymbol{\mu}_Z, \hat{\Sigma}_f) = c_{d,m} \left| \hat{\Sigma}(\theta) \right|^{-1/2} \left[1 + m^{-1} \chi^2(\mathbf{Z}; \boldsymbol{\mu}_Z(\theta), \hat{\Sigma}(\theta)) \right]^{-(m+1)/2}, \quad (3.28)$$

where the normalisation constant is

$$c_{d,m} = (m\pi)^{-d/2} \Gamma\left(\frac{m+1}{2}\right) / \Gamma\left(\frac{m-d+1}{2}\right), \quad (3.29)$$

the usual chi-square is given by

$$\chi^2(\mathbf{Z}; \boldsymbol{\mu}_z(\theta), \widehat{\boldsymbol{\Sigma}}(\theta)) = [\mathbf{Z} - \boldsymbol{\mu}_z(\theta)]^\top \widehat{\boldsymbol{\Sigma}}(\theta)^{-1} [\mathbf{Z} - \boldsymbol{\mu}_z(\theta)], \quad (3.30)$$

the components of the mean vector $\boldsymbol{\mu}_z$ are given by equation (3.20), and the cosmology-varying covariance matrix estimate $\widehat{\boldsymbol{\Sigma}}(\theta)$ is related to the fiducial estimate $\widehat{\boldsymbol{\Sigma}}_f$ by equation (3.26).

This likelihood function is the key result of this chapter. In the next section, we shall test our likelihood treatments using Monte Carlo simulations matched to the typical specifications of future survey data.

3.3 Testing with Monte Carlo simulations

In order to test the proposed likelihood function, we simulate power spectrum measurements with Monte Carlo sampling. Exponential random variables representing clustering mode power and shot noise are generated and then convolved with a survey window before forming the power spectrum data vectors. Given the survey volume and the largest scales accessible for upcoming missions such as DESI and *Euclid*, we consider wavenumbers $k \in [1.58 \times 10^{-3}, 1.58 \times 10^{-2}] h \text{ Mpc}^{-1}$ covering an order of magnitude,⁷ and sample clustering mode wavevectors \mathbf{q} in this range using an inverse-volume distribution $\mathbb{p}(\mathbf{q}) \propto |\mathbf{q}|^3$. The Fourier-space window $|G_{1,1}|^2$ is chosen to be a centred, normalised Gaussian function with full width at half maximum being $1.88 \times 10^{-3} h \text{ Mpc}^{-1}$, and we leave it unnormalised when computing the mixing matrix \mathbf{B} (eq. 3.2) for simplicity. The wavenumbers k_a are divided into $d = 9$ bands (or bins) so that the cross-band correlation is weak under this window function. Given an input cosmological model, the band power spectrum is predicted as follows:

- 1) The underlying galaxy clustering power spectrum is calculated using the Eisenstein–Hu matter transfer function [186], with the large-scale linear bias fixed at $b_1 = 1.87$ and other cosmological parameters set to *Planck* 2018 values [101];
- 2) The Poisson shot noise power $P_{\text{shot}} = 1/\bar{n}$ is calculated using a pessimistic number density $\bar{n} = 5 \times 10^{-4} h^3 \text{ Mpc}^{-3}$, though the number densities for DESI and *Euclid* are predicted to be likely higher [e.g. 187];
- 3) The convolution of the underlying clustering mode power and shot noise is performed using equation (3.3), where elements of the mixing matrix are given by equation (3.2).

3.3.1 Testing multivariate normality and covariance estimate rescaling

In Figure 3.2, the sample distribution of 40000 simulated power spectrum measurements for the wavenumber band centred at $k \approx 0.0024 h \text{ Mpc}^{-1}$ is compared with the gamma distribution approximation (eq. 3.12) and the normal distribution assumption (eq. 3.14), which have

⁷ The lowest wavenumber roughly corresponds to a cubic box of side length $4 h^{-1} \text{ Mpc}$, or comoving radius at redshift $z \approx 2.5$.

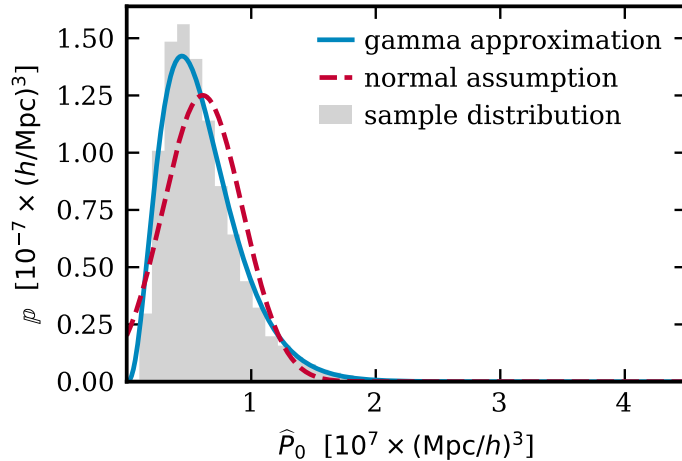


Figure 3.2. Sample distribution of 40000 simulated power spectrum measurements in the lowest wavenumber band centred at $k \approx 0.0024 h \text{ Mpc}^{-1}$ (grey histogram) compared with the gamma distribution approximation (solid blue line) and the normal distribution assumption (dashed red line) with the same mean and variance. Figure originally appearing as Fig. 2 in Paper I and reproduced here with re-simulated data sets.

the same mean and variance. The gamma distribution used has a shape parameter $R \approx 4$, as calculated from the underlying power spectrum model and the mixing matrix using equations (3.11) and (3.12). This suggests that the wavenumber bin has very few effectively independent clustering modes for which the central limit theorem certainly does not hold. Indeed, the peak of the normal distribution fitted is shifted from the underlying hypo-exponential distribution that describes the sample; in contrast, the gamma distribution approximation matches both the peak and the tails well.

One might wish to quantify the improvement in multivariate normality that our component-wise Gaussianising transformation can bring to power spectrum measurements. A key defining property of a multivariate normal variable \mathbf{X} is that any projection $\mathbf{X} \mapsto \mathbf{t}^\top \mathbf{X} \in \mathbb{R}$ with a vector \mathbf{t} should result in a univariate normal variable. Hence as a simple multivariate normality test, given a sample $\{\mathbf{x}_i\}$ for the random vector \mathbf{X} , one could randomly choose some direction \mathbf{t} and test the univariate normality of the projected sample $\{\mathbf{t}^\top \mathbf{x}_i\}$ [188]. Here we perform the D’Agostino–Pearson normality test [189] on 10000 random projections of 40000 samples of the band power data vector $\tilde{\mathbf{Y}}$, which returns the p -value that characterises how unlikely the sample realisations are under the null hypothesis that the underlying distribution is indeed normal. The p -value itself is not a meaningful indicator of normality, as it varies depending on the sample size; rather it is the comparison of the p -values with the same sample that signifies the relative departure from normality. We find $p = 0.01$ for $\tilde{\mathbf{Y}}$ without Gaussianisation; after Gaussianisation $\tilde{\mathbf{Y}} \mapsto \mathbf{Z}$ with the Box–Cox transformation parameter ν_a computed using the fitting formula (eq. 3.24) for each data vector component, the p -value is increased to $p = 0.08$ for the same sample, suggesting our Gaussianisation scheme does improve the multivariate normality of power spectrum measurements.

To test the covariance estimate rescaling method using the variance–correlation decomposition proposed in section 3.2.2, we generate one set of 40000 band power data realisations at the ‘true’ cosmological model with the Hubble parameter set to $H_0 = 67.4 \text{ km s}^{-1} \text{ Mpc}^{-1}$

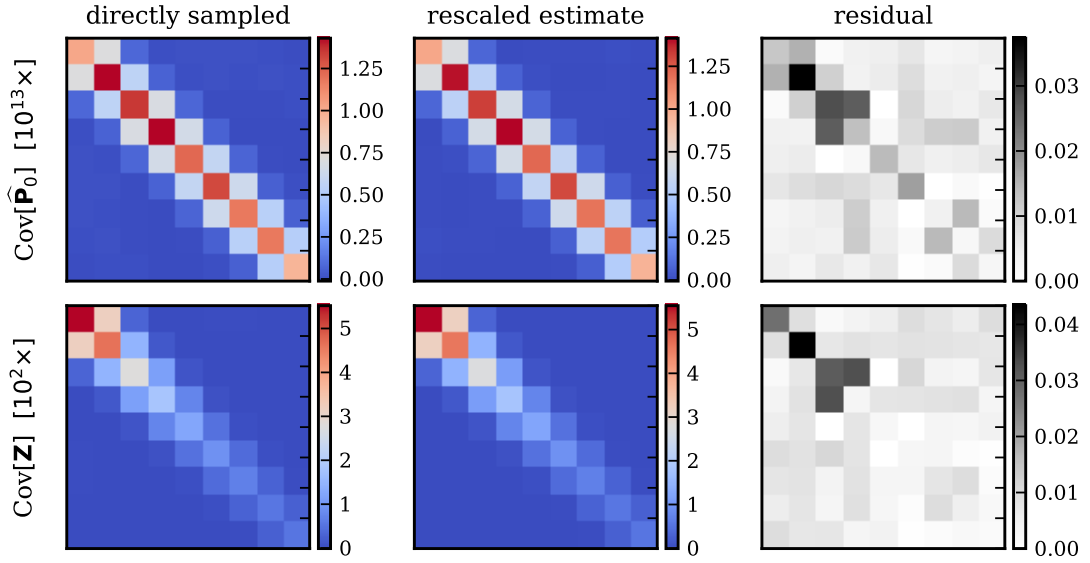


Figure 3.3. Comparison between directly sampled covariance matrices and rescaled fiducial estimates from 40000 simulations of the band power data vector either without Gaussianisation (top panel) or Gaussianised (bottom panel). Covariance matrices sampled at the ‘true’ cosmological model with $H_0 = 67.4 \text{ km s}^{-1} \text{ Mpc}^{-1}$ are shown in the *left column*, and the *middle column* shows the covariance matrix estimates at the ‘fiducial’ value $H_0 = 73.2 \text{ km s}^{-1} \text{ Mpc}^{-1}$, which have been rescaled to match the ‘true’ cosmology. The residuals between the directly sampled covariance matrices and the rescaled fiducial estimates are shown in the *right column*. Figure originally appearing as Fig. 3 in Paper I and reproduced here with re-simulated data sets.

(*Planck* 2018 cosmology), and an additional set of 40000 ‘fiducial’ realisations generated with $H_0 = 73.2 \text{ km s}^{-1} \text{ Mpc}^{-1}$. The former set gives a directly sampled covariance matrix estimate $\hat{\Sigma}$ whereas the latter gives a fiducial estimate $\hat{\Sigma}_f$ which is then rescaled using equation (3.26) to match the true cosmology. For both the band power \tilde{Y} without Gaussianisation and the Gaussianised band power Z , Figure 3.3 shows that the differences between the directly sampled true covariance matrices and the rescaled covariance matrix estimates are small; this validates the variance–correlation decomposition as an accurate method to include cosmological dependence of the covariance matrix.

3.3.2 Testing likelihoods for parameter inference

To summarise, we have made the following proposals in obtaining the improved new likelihood function (eq. 3.28), which can be tweaked in our Monte Carlo simulation tests to isolate their effects in a likelihood analysis—

- 1) Gaussianisation: the data vector can remain as \tilde{Y} without Gaussianisation, Gaussianised as $Z_{1/3}$ with a fixed Box–Cox parameter $\nu = 1/3$, or Gaussianised as Z_ν with a variable parameter ν given by the fitting formula (eq. 3.24);
- 2) Covariance estimate rescaling: the covariance matrix estimate can either remain fixed at the fiducial value $\hat{\Sigma}_f$, or be rescaled to $\hat{\Sigma}(\theta)$ using equation (3.26) to account for parameter dependence;
- 3) Covariance matrix marginalisation: one could either substitute the de-biased precision

Table 3.1. Different forms of the likelihood function for power spectrum measurements with or without data Gaussianisation and/or covariance matrix marginalisation and estimate rescaling. For each likelihood function, the covariance matrix estimate $\hat{\Sigma}$ is understood to be defined for the corresponding data vector, so any additional labels (such as a subscript ‘ Z_ν ’) are omitted for notational brevity.

Data variable	Distribution	Fiducial estimate	Rescaled estimate
No Gauss.	Modified t	$\mathcal{L}_t(\theta; \tilde{Y}, \hat{\Sigma}_f)$	$\mathcal{L}_t(\theta; \tilde{Y}, \hat{\Sigma}(\theta))$
With Gauss.	$\nu = 1/3$	Gaussian	$\mathcal{L}_G(\theta; Z_{1/3}, \hat{\Sigma}(\theta))$
		Modified t	$\mathcal{L}_t(\theta; Z_{1/3}, \hat{\Sigma}_f)$
	Variable ν	Modified t	$\mathcal{L}_t(\theta; Z_\nu, \hat{\Sigma}_f)$
			$\mathcal{L}_t(\theta; Z_\nu, \hat{\Sigma}(\theta))$

matrix estimate directly into the Gaussian likelihood (see eq. 1.63 in §1.4), or adopt the SH marginalisation procedure and use the modified t -distribution likelihood (eq. 1.73). Different combinations of these choices give the likelihood function forms tabulated in Table 3.1 which are to be tested. For notational brevity, we do not label covariance matrix estimates $\hat{\Sigma}$ by the data vector with respect to which they are defined (e.g. labelled as $\hat{\Sigma}_{\tilde{Y}}$), since they invariably appear in conjunction with the corresponding data vector in a likelihood function. The key test of our novel likelihood treatments for power spectrum measurements is the inference of cosmological parameters, and all these likelihood functions can be compared with the exact true likelihood, which is inaccessible in a realistic survey but can be constructed simply as a product from the exponentially distributed clustering power and shot noise generated in our Monte Carlo simulations in the absence of window effects,

$$\mathcal{L}_{\text{true}}(\theta; Y) = \prod_{i=1}^{d'} \frac{e^{-Y_i/P_i} - e^{-Y_i/P_{\text{shot}}}}{P_i - P_{\text{shot}}}. \quad (3.31)$$

Since our methodology mainly concerns power spectrum measurements on the largest survey scales, where asymptotic normality does not apply to the few clustering modes available and theoretical models are in the linear perturbative regime, the local PNG parameter f_{NL} is a well-motivated test parameter [172, 173]. PNG enters the galaxy clustering power spectrum by modifying the large-scale tracer bias on large scales as described by equation (1.60) (recall from section 1.3), where the cosmological parameters are set to *Planck* 2018 values as above and the tracer-dependent parameter is set to $p = 1$. Hereafter the generic cosmological parameter θ is identified with f_{NL} and the two notations are used interchangeably. To leading order in f_{NL} , which is small as constrained by the most recent *Planck* CMB measurements [88], we continue to treat galaxy over-density as a GRF.

To properly examine the ensemble behaviour of the different likelihood functions listed in Table 3.1, we have produced 250000 Monte Carlo realisations of the band power data vector at some ‘true’ input cosmology, and a fixed set of $N_{\text{mock}} = 1000$ mock catalogues simulated at the fiducial cosmology with $f_{\text{NL}} = 0$ and other parameters being the same as in the true input model. The mock catalogues provide covariance matrix estimates for both the un-Gaussianised band power \tilde{Y} and the Gaussianised data vector Z . The prior is set to be uniform

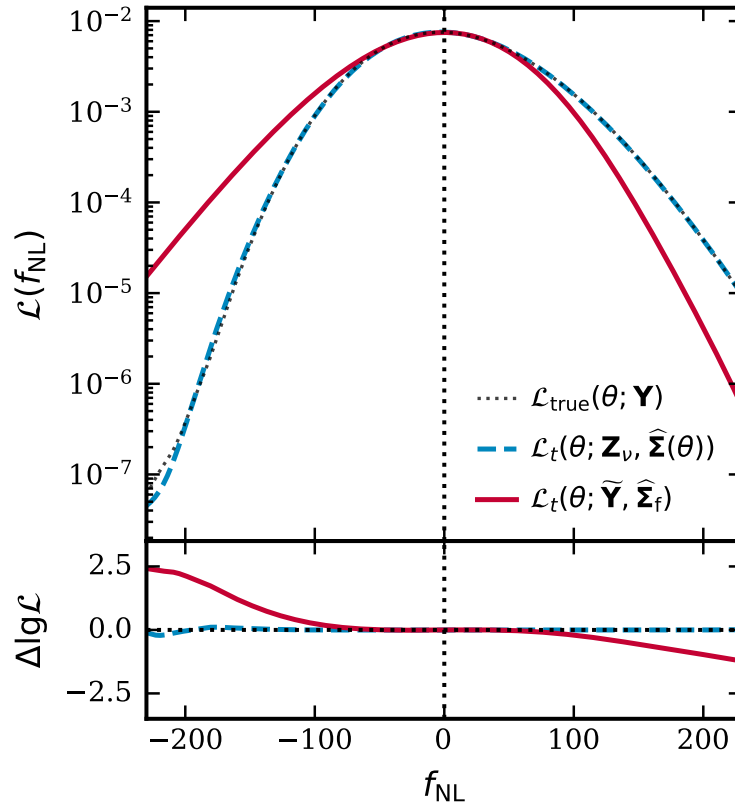


Figure 3.4. Comparison of the exact true likelihood $\mathcal{L}_{\text{true}}(\theta; Y)$ (dotted black line), the modified- t likelihood $\mathcal{L}_t(\theta; \mathbf{Z}_\nu, \hat{\Sigma}(\theta))$ with data Gaussianisation and covariance estimate rescaling (dashed blue line), and the modified- t likelihood $\mathcal{L}_t(\theta; \tilde{Y}, \hat{\Sigma}_f)$ without Gaussianisation and rescaling (solid red line) for $\theta \equiv f_{\text{NL}}$, all marginalised over 250000 data realisations produced at $f_{\text{NL}} = 0$ (marked by vertical dotted lines) (top panel). The bottom panel shows the deviations of the likelihoods from the true one $\mathcal{L}_{\text{true}}$. Note the logarithmic scale of the vertical axes and that $\text{lg} \equiv \log_{10}$.

with $f_{\text{NL}} \in [-250, 250]$, and the posterior PDFs are sampled by scanning through this range with a resolution of $\Delta f_{\text{NL}} = 0.05$. This prior range is chosen to be wide in order to fully sample the tails of the posterior distribution, since a key improvement in our likelihood treatments comes from the use of the full shape of non-normal distributions.

Direct likelihood form comparison

As an intuitive comparison, we show in Figure 3.4 the logarithms of the exact true likelihood $\mathcal{L}_{\text{true}}(\theta; Y)$, the modified- t likelihood $\mathcal{L}_t(\theta; \tilde{Y}, \hat{\Sigma}_f)$ without data Gaussianisation and covariance estimate rescaling, and the novel likelihood $\mathcal{L}_t(\theta; \mathbf{Z}_\nu, \hat{\Sigma}(\theta))$ derived using our full methodology with a fitted Box–Cox transformation parameter ν , all marginalised over data realisations produced at $f_{\text{NL}} = 0$. In addition, we also show the comparison for some individual data realisations in Figure 3.5, for each of which the new likelihood is a closer match in shape to the true likelihood. Although throughout this chapter we focus on the inference of f_{NL} as a well-motivated test parameter, the novel likelihood also performs well when tested on other cosmological parameters; for instance, in Figure 3.6 we also infer the H_0 parameter

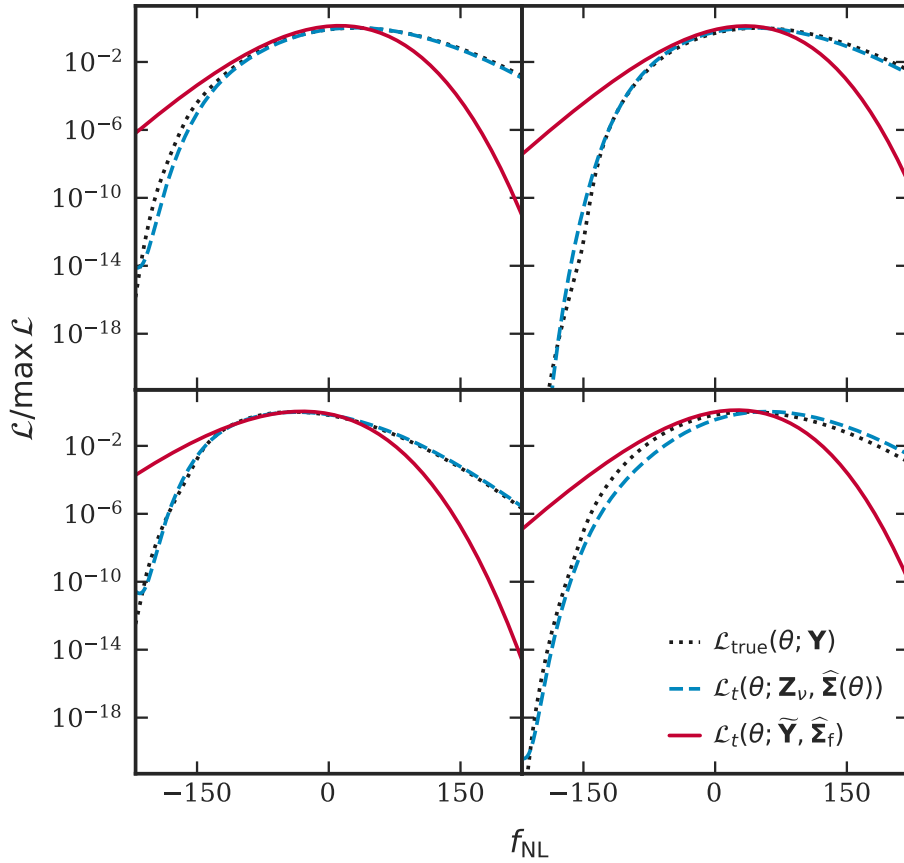


Figure 3.5. Comparison of the exact true likelihood $\mathcal{L}_{\text{true}}(\theta; \mathbf{Y})$ (dotted black line), the modified- t likelihood $\mathcal{L}_t(\theta; \mathbf{Z}_\nu, \hat{\Sigma}(\theta))$ with data Gaussianisation and covariance estimate rescaling (dashed blue line), and the modified- t likelihood $\mathcal{L}_t(\theta; \tilde{\mathbf{Y}}, \hat{\Sigma}_f)$ without Gaussianisation and rescaling (solid red line) for $\theta \equiv f_{\text{NL}}$ for four individual data realisations produced at $f_{\text{NL}} = 0$.

from the same ensemble of data realisations. All these comparisons demonstrably show that our novel likelihood is a far superior approximation to the exact true likelihood when compared to the standard likelihood with a fixed fiducial covariance estimate which also fails to account for the non-normal distribution of power spectrum measurements.

Point estimation comparison

For different ‘true’ f_{NL} input values and the same fiducial cosmological model at $f_{\text{NL}} = 0$, we compare the performance of both frequentists’ and Bayes estimators calculated from the likelihoods $\mathcal{L}_{\text{true}}(\theta; \mathbf{Y})$, $\mathcal{L}_t(\theta; \tilde{\mathbf{Y}}, \hat{\Sigma}_f)$, $\mathcal{L}_t(\theta; \tilde{\mathbf{Y}}, \hat{\Sigma}(\theta))$, $\mathcal{L}_t(\theta; \mathbf{Z}_\nu, \hat{\Sigma}_f)$ and $\mathcal{L}_t(\theta; \mathbf{Z}_\nu, \hat{\Sigma}(\theta))$ (see Table 3.1 for reference), with reported results having been marginalised over all data realisations.

The *maximum likelihood estimator* is a common frequentists’ estimator given by

$$\hat{\theta} = \arg \max \mathcal{L}(\theta). \quad (3.32)$$

The estimates are compared in Table 3.2 with uncertainties given by the standard deviations estimated from the ensemble of data realisations.

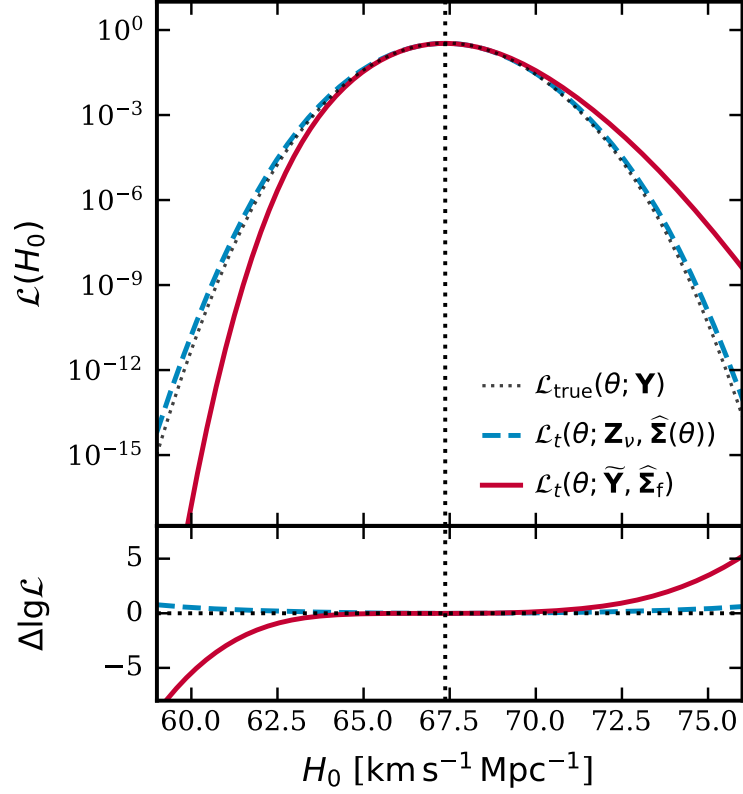


Figure 3.6. Comparison of the exact true likelihood $\mathcal{L}_{\text{true}}(\theta; Y)$ (dotted black line), the modified- t likelihood $\mathcal{L}_t(\theta; Z_\nu, \hat{\Sigma}(\theta))$ with data Gaussianisation and covariance estimate rescaling (dashed blue line), and the modified- t likelihood $\mathcal{L}_t(\theta; \tilde{Y}, \hat{\Sigma}_f)$ without Gaussianisation and rescaling (solid red line) for $\theta \equiv H_0$ marginalised over $f_{\text{NL}} = 0$ data realisations produced at $H_0 = 67.4 \text{ km s}^{-1} \text{ Mpc}^{-1}$ (marked by vertical dotted lines) (top panel). The bottom panel shows the deviations of the likelihoods from the true one $\mathcal{L}_{\text{true}}$. Note the logarithmic scale of the vertical axes and that $\text{lg} \equiv \log_{10}$.

Table 3.2. Maximum likelihood estimates of f_{NL} from likelihood functions $\mathcal{L}_{\text{true}}(\theta; Y)$, $\mathcal{L}_t(\theta; \tilde{Y}, \hat{\Sigma}_f)$, $\mathcal{L}_t(\theta; Z_\nu, \hat{\Sigma}_f)$, $\mathcal{L}_t(\theta; \tilde{Y}, \hat{\Sigma}(\theta))$ and $\mathcal{L}_t(\theta; Z_\nu, \hat{\Sigma}(\theta))$ (see Table 3.1 for reference), marginalised over 250000 data realisations with different true f_{NL} inputs and the same fiducial cosmological at $f_{\text{NL}} = 0$. The reported uncertainties are the estimated standard deviations from the ensemble of data realisations.

Input f_{NL}	Maximum likelihood estimates				
	$\mathcal{L}_{\text{true}}(\theta; Y)$	$\mathcal{L}_t(\theta; \tilde{Y}, \hat{\Sigma}_f)$	$\mathcal{L}_t(\theta; Z_\nu, \hat{\Sigma}_f)$	$\mathcal{L}_t(\theta; \tilde{Y}, \hat{\Sigma}(\theta))$	$\mathcal{L}_t(\theta; Z_\nu, \hat{\Sigma}(\theta))$
50	49 ± 42	49 ± 44	51 ± 42	42 ± 44	48 ± 42
10	9 ± 39	9 ± 39	11 ± 39	2 ± 41	8 ± 40
0	-1 ± 39	-2 ± 38	1 ± 38	-8 ± 41	-2 ± 39
-10	-12 ± 38	-12 ± 38	-10 ± 37	-18 ± 40	-12 ± 38
-50	-52 ± 35	-52 ± 37	-50 ± 34	-58 ± 36	-52 ± 35

Table 3.3. Posterior median estimates of f_{NL} from likelihood functions $\mathcal{L}_{\text{true}}(\theta; Y)$, $\mathcal{L}_t(\theta; \tilde{Y}, \hat{\Sigma}_f)$, $\mathcal{L}_t(\theta; Z_\nu, \hat{\Sigma}_f)$, $\mathcal{L}_t(\theta; \tilde{Y}, \hat{\Sigma}(\theta))$ and $\mathcal{L}_t(\theta; Z_\nu, \hat{\Sigma}(\theta))$ (see Table 3.1 for reference), marginalised over 25 0000 data realisations with different true f_{NL} inputs and the same fiducial cosmological at $f_{\text{NL}} = 0$. The reported uncertainties correspond to the equal-tailed 68.3 % credible interval.

Input f_{NL}	Posterior median estimates				
	$\mathcal{L}_{\text{true}}(\theta; Y)$	$\mathcal{L}_t(\theta; \tilde{Y}, \hat{\Sigma}_f)$	$\mathcal{L}_t(\theta; Z_\nu, \hat{\Sigma}_f)$	$\mathcal{L}_t(\theta; \tilde{Y}, \hat{\Sigma}(\theta))$	$\mathcal{L}_t(\theta; Z_\nu, \hat{\Sigma}(\theta))$
50	53^{+43}_{-39}	46^{+33}_{-35}	51^{+39}_{-38}	48^{+40}_{-35}	53^{+43}_{-38}
10	13^{+41}_{-36}	6^{+36}_{-39}	11^{+38}_{-38}	8^{+38}_{-32}	13^{+40}_{-35}
0	3^{+40}_{-36}	-5^{+37}_{-40}	1^{+38}_{-37}	-2^{+38}_{-32}	3^{+40}_{-35}
-10	-7^{+39}_{-35}	-15^{+37}_{-41}	-9^{+38}_{-37}	-12^{+37}_{-31}	-7^{+39}_{-34}
-50	-46^{+36}_{-31}	-56^{+41}_{-44}	-49^{+37}_{-37}	-52^{+35}_{-28}	-46^{+36}_{-30}

With uniform priors, the posterior $\mathcal{P}(\theta | X)$ of a cosmological parameter θ given any data X is simply the likelihood $\mathcal{L}(\theta; X)$ suitably normalised. Common Markov chain Monte Carlo (MCMC) analyses usually return the posterior median or mean as the Bayes estimate [110, 190], and here we choose the absolute loss function

$$\text{loss}(\lambda; \theta) = |\lambda - \theta| \quad (3.33)$$

and minimise its expectation over the posterior of θ to obtain the *posterior median estimator* [191]

$$\hat{\theta} = \arg \min_{\lambda} \mathbb{E}_{\theta | X} [\text{loss}(\lambda; \theta)]. \quad (3.34)$$

As the name suggests, this estimator is simply the median point of the posterior distribution. The corresponding uncertainties can be quoted as the equal-tailed 68.3 % credible interval. For the same likelihood functions considered above, we show the results in Table 3.3.

It is evident that overall the novel likelihood $\mathcal{L}_t(\theta; Z_\nu, \hat{\Sigma}(\theta))$ performs the best in producing parameter estimates as well as uncertainty bounds close to those from the exact true likelihood, whether a frequentists' or a Bayes estimator is used. We have found that Gaussianisation with a fixed transformation parameter $\nu = 1/3$ gives similar results to Gaussianisation with a variable ν given by the fitting formula (eq. 3.24); likewise, assuming a wrong fiducial cosmological model for Gaussianisation, even when it deviates significantly from the true cosmology, has negligible impact on recovered parameters. This demonstrates that our Gaussianisation scheme is robust to variations of the fiducial cosmological model.

To get a sense of how the f_{NL} posteriors vary with each data realisation, in Figure 3.7 we show the normalised histograms of the posterior median estimates and 68.3 % credible bounds produced from the exact true likelihood $\mathcal{L}_{\text{true}}(\theta; Y)$, the standard likelihood $\mathcal{L}_t(\theta; \tilde{Y}, \hat{\Sigma}_f)$ without data Gaussianisation and covariance estimate rescaling, and the novel likelihood $\mathcal{L}_t(\theta; Z_\nu, \hat{\Sigma}(\theta))$ over the ensemble of $f_{\text{NL}} = 0$ data realisations. Because of the relatively large uncertainty in f_{NL} , the posterior median estimates given by these likelihoods are statistically consistent. Nevertheless, it is apparent that the standard likelihood without Gaussianisation or covariance estimate rescaling tends to negatively shift the credible bounds on f_{NL} ,

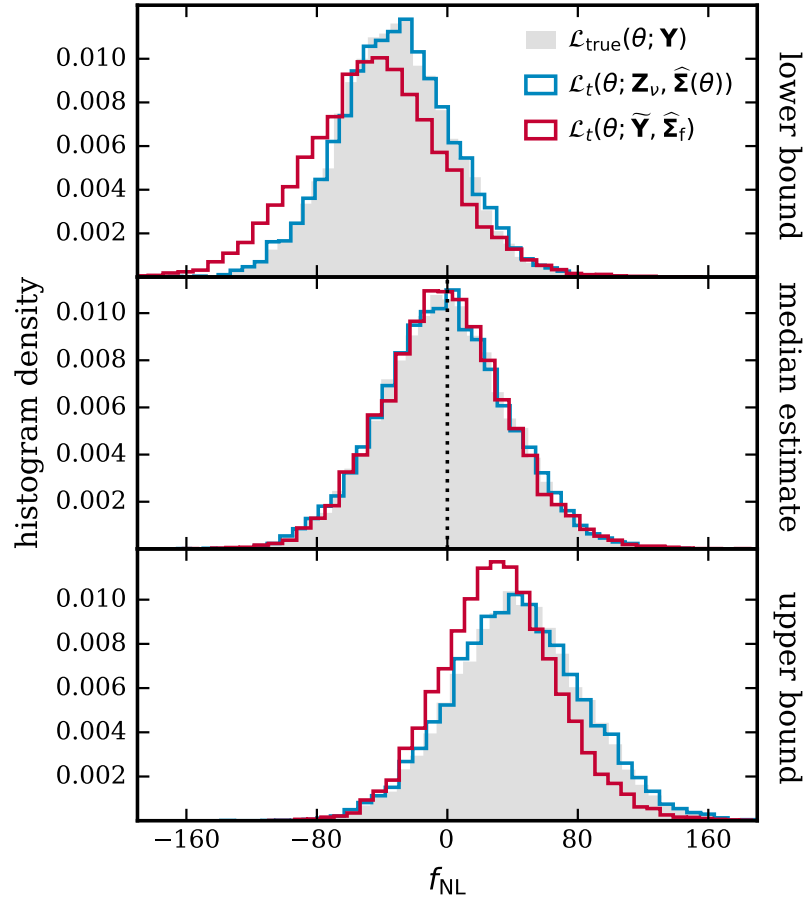


Figure 3.7. Normalised histograms of the posterior median estimates (middle panel) and 68.3 % credible bounds (top and bottom panels) produced from the exact true likelihood $\mathcal{L}_{\text{true}}(\theta; \mathbf{Y})$ (shaded), the standard likelihood $\mathcal{L}_t(\theta; \tilde{\mathbf{Y}}, \hat{\Sigma}_f)$ without data Gaussianisation and covariance estimate rescaling (in red), and the novel likelihood $\mathcal{L}_t(\theta; \mathbf{Z}_\nu, \hat{\Sigma}(\theta))$ (in blue) over the ensemble of $f_{\text{NL}} = 0$ data realisations.

whereas the novel likelihood is in excellent agreement with the exact true likelihood. Again, this reflects the fact that away from the peak of the posterior distribution, our methodology produces a superior likelihood approximation to the true one in terms of the overall shape of the distribution and in particular its tails.

Posterior shape comparison

A graphical comparison of probability distributions is the *quantile–quantile (Q–Q) probability plot* [192]. Figure 3.8 shows the f_{NL} percentiles inferred from likelihood functions listed in Table 3.1 except $\mathcal{L}_G(\theta; \mathbf{Z}_{1/3}, \hat{\Sigma}(\theta))$ against f_{NL} percentiles of the exact true likelihood $\mathcal{L}_{\text{true}}(\theta; \mathbf{Y})$, where we contrast no Gaussianisation against Gaussianisation, and fixed covariance estimates against rescaled covariance estimates. There are two trends that match our expectation: 1) posterior distributions of the Gaussianised data \mathbf{Z} match the shape of the true posterior better than those of the un-Gaussianised data $\tilde{\mathbf{Y}}$, especially away from the peak and near the tails of the distribution; 2) not accounting for parameter dependence of the covariance matrix by the rescaling method also noticeably distorts the shape of the posterior dis-

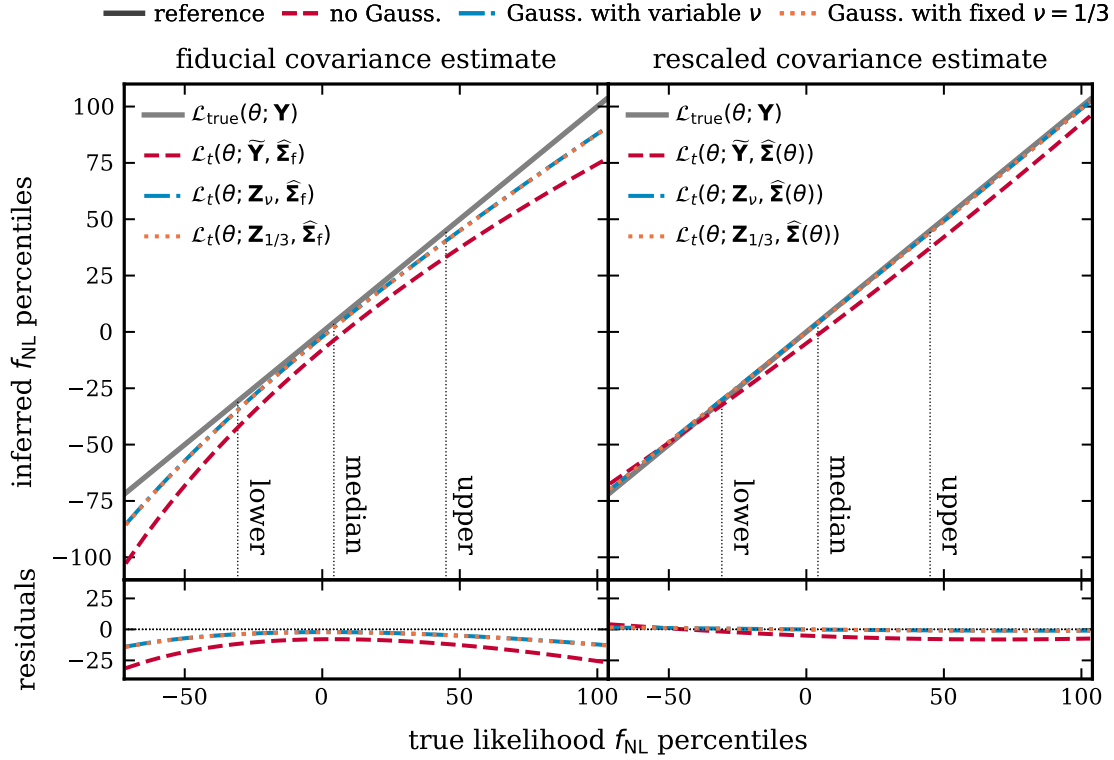


Figure 3.8. Q–Q plots comparing f_{NL} posterior distributions obtained from the different likelihood functions without or with data Gaussianisation and/or covariance estimate rescaling (see Table 3.1 for reference). The top panel shows the inferred f_{NL} percentiles from the different likelihood functions against the true f_{NL} percentiles, all marginalised over 250000 data realisations at $f_{\text{NL}} = 0$, and the bottom panel shows residuals of the f_{NL} percentiles. The *right column* shows the cases with covariance estimate rescaling and the *left column* the cases without. The unit-slope lines are reference lines showing no deviation from the exact true likelihood, and the dotted vertical lines mark the f_{NL} posterior median estimate and the equal-tailed 68.3% credible interval from the exact true likelihood. Figure adapted from Fig. 5 in Paper I.

tribution. Again we have found that Gaussianisation with a fixed Box–Cox transformation parameter $\nu = 1/3$ produces nearly indistinguishable results from Gaussianisation with a variable ν .

Another quantitative measure of the ‘statistical distance’ between a probability density function $f(\theta)$ and an approximate one $g(\theta)$ is the *Kullback–Leibler* (KL) *divergence* [193]

$$D_{\text{KL}}(f \parallel g) = \int d\theta f(\theta) \ln \frac{f(\theta)}{g(\theta)}. \quad (3.35)$$

For instance, if one takes f to be the posterior of the exact true likelihood against which the posterior g of the various likelihood functions is compared, then the KL divergence marginalised over the entire ensemble of data realisations could quantify the ‘information loss’ due to the replacement of the exact true likelihood with the approximate ones. Table 3.4 shows the posterior KL divergence values with respect to the exact true likelihood for all likelihood functions in Table 3.1 except $\mathcal{L}_{\text{G}}(\theta; \mathbf{Z}_{1/3}, \hat{\Sigma}(\theta))$, with the results having been averaged over all data realisations generated at each true f_{NL} input. As before, the novel likelihood $\mathcal{L}_{\text{t}}(\theta; \mathbf{Z}_{\nu}, \hat{\Sigma}(\theta))$

Table 3.4. KL divergence values with respect to the true f_{NL} posterior for posteriors of various likelihood functions (see Table 3.1 for reference). Results have been marginalised over 250000 data realisations with different true f_{NL} inputs and the same fiducial cosmology at $f_{\text{NL}} = 0$.

Input f_{NL}	D_{KL} from the true posterior				
	50	10	0	-10	-50
$\mathcal{L}_t(\theta; \tilde{Y}, \hat{\Sigma}_f)$	0.35	0.18	0.17	0.17	0.23
$\mathcal{L}_t(\theta; Z_{1/3}, \hat{\Sigma}_f)$	0.05	0.04	0.05	0.05	0.10
$\mathcal{L}_t(\theta; Z_\nu, \hat{\Sigma}_f)$	0.05	0.04	0.05	0.05	0.09
$\mathcal{L}_t(\theta; \tilde{Y}, \hat{\Sigma}(\theta))$	0.12	0.11	0.10	0.10	0.11
$\mathcal{L}_t(\theta; Z_{1/3}, \hat{\Sigma}(\theta))$	0.02	0.02	0.02	0.02	0.03
$\mathcal{L}_t(\theta; Z_\nu, \hat{\Sigma}(\theta))$	0.02	0.02	0.02	0.02	0.03

with both data Gaussianisation and covariance estimate rescaling is the closest to the true likelihood, and Gaussianisation with $\nu = 1/3$ fixed produces very similar results.

Sources of error in parameter inference

Although two major sources of error in parameter inference, namely the Gaussian likelihood assumption and the neglect of parameter dependence of the covariance matrix, have been identified and mitigated by Gaussianisation and variance–correlation rescaling of the covariance matrix respectively, there are other sources of error which we now consider.

The first concern is how the correlation between band power measurements may adversely affect the univariate Gaussianisation scheme we have adopted for simplicity. Univariate Gaussianising transformations are expected to work the best when the off-diagonal elements of the correlation matrix are comparatively small, and hence for a given window function, one would ideally choose broader wavenumber bins for band power measurements. For future surveys with both wider coverage and greater depth, this will be easier as the Fourier-space window function will be narrower. However, broadening the wavenumber bins also means that relatively more clustering modes contribute to each band power, thus making the data variables more Gaussian. The limit to how broad wavenumber bins should be will depend on whether there are sufficiently many bins to retain the cosmological information in the survey data.

The second concern is that when one cannot directly calculate the mixing matrix, the gamma distribution shape–scale parameters may have to be estimated from mock catalogues as part of covariance matrix estimation; this can bring additional statistical scatter to likelihood inference. Ideally the shape–scale parameters need to be marginalised out together with the unknown true covariance matrix in the SH procedure, but this unfortunately makes the likelihood function analytically intractable after Gaussianisation (see also the discussion in §3.2.2). We investigate this issue with Monte Carlo simulations as before and compare the impact on the posterior shape resulting from covariance matrix estimation and the evaluation of gamma distribution shape–scale parameters (R, η) using the estimated band power variance. Both estimations are made with the same ensemble of mock catalogues; for 25000 data realisations,

we have generated one ensemble of $N_{\text{mock}} = 1000$ mock catalogues for *each* data realisation. In a second Q–Q probability plot (Fig. 3.9) for the posteriors of likelihood functions of the Gaussianised data \mathbf{Z} with a fixed Box–Cox parameter $\nu = 1/3$ and rescaled covariance matrix estimates, we consider three scenarios:

- 1) The de-biased and rescaled precision matrix estimate is directly substituted into the Gaussian likelihood $\mathcal{L}_G(\theta; \mathbf{Z}_{1/3}, \hat{\Sigma}(\theta))$;
- 2) The SH marginalisation procedure is performed and thus the modified- t likelihood $\mathcal{L}_t(\theta; \mathbf{Z}_{1/3}, \hat{\Sigma}(\theta))$ is used;
- 3) A high-precision covariance matrix estimate is used as a proxy for the exact covariance matrix in the Gaussian likelihood $\mathcal{L}_G(\theta; \mathbf{Z}_{1/3}, \Sigma(\theta))$.

For each of these scenarios, we further compare the case where the parameters (R, η) are analytically calculated (using the mixing matrix and the fiducial power spectrum including shot noise) against the case where they are obtained from estimated band power variance. In addition, we also vary the size of the mock catalogue ensemble to differentiate between these multiple scenarios.

The results in Figure 3.9 indicate that the SH marginalisation procedure (i.e. the modified- t likelihood) indeed accounts for the statistical scatter due to covariance matrix estimation. The impact of covariance matrix estimation appears to be subdominant to that of band power variance estimation used for evaluating (R, η) if the ensemble size of mock catalogues is large, e.g. with $N_{\text{mock}} = 1000$ in the left column of the figure, where the residual lines fall into two distinctive groups depending on whether (R, η) are estimated or exact. In contrast, when the ensemble size of mock catalogues is small, e.g. $N_{\text{mock}} = 50$ in the bottom right panel, the impact of covariance estimation seems to be the greater of the two. However, by comparing with Figure 3.8, it is evident that the impact from these estimations is far less significant than the naïve Gaussian likelihood assumption and the neglect of the cosmological dependence of the covariance matrix, which are the focal problems addressed in this work. In particular, the errors due to these estimations with $N_{\text{mock}} = 1000$, which is a typical number of mock catalogues required for upcoming surveys such as DESI and *Euclid*, are approaching the errors inherent in our gamma distribution approximation and univariate Gaussianisation in any case. In light of these results, we recommend using the SH marginalisation procedure for the covariance matrix estimate, i.e. adopting the modified- t likelihood, even when the gamma distribution parameters are estimated and cannot be marginalised out in a similar fashion.

3.4 Likelihood construction for survey application

For the benefit of the reader who wishes to apply the methodology proposed in this work to cosmological inference with future survey data, we now present a straightforward pipeline for constructing the likelihood function.

Gamma distribution approximation. Measurements of the band power spectrum are modelled as gamma random variables. Given band power measurements \hat{P} at binned wavenumbers k_a , the gamma distribution shape and scale parameters are determined from the power

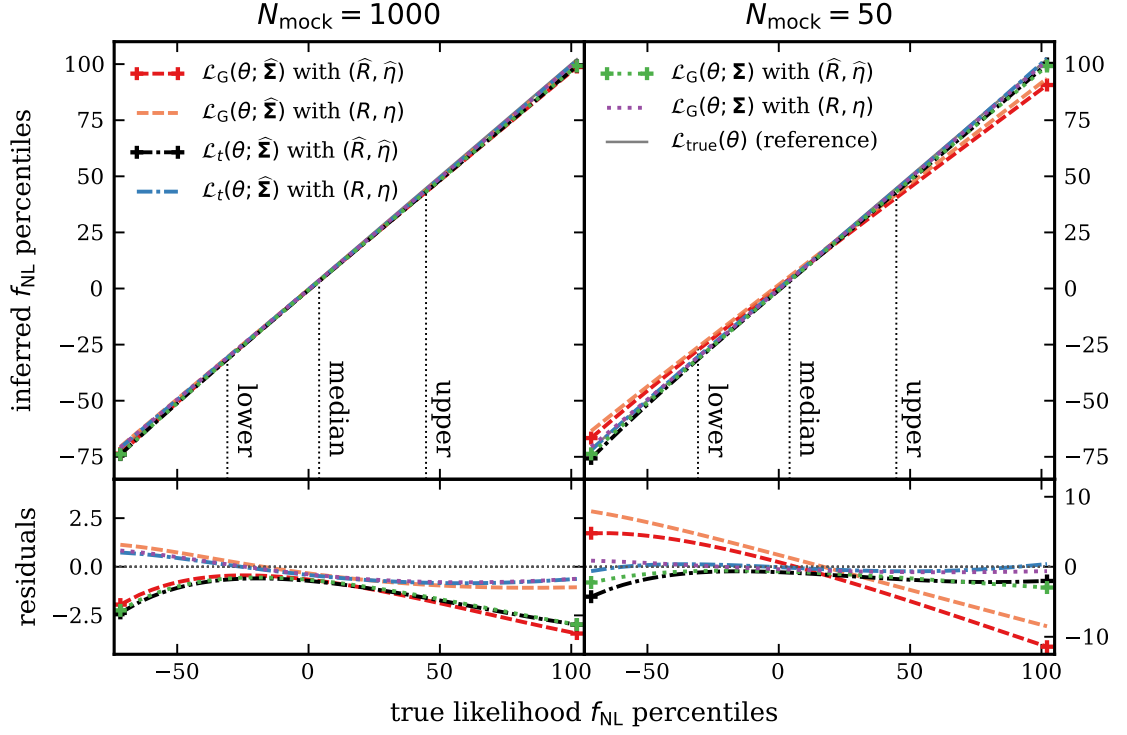


Figure 3.9. Q–Q plots of f_{NL} posterior distributions for comparing the relative impact of covariance matrix estimation and the evaluation of gamma distribution parameters (R, η) using estimated band power variance. All likelihood functions compared are of the Gaussianised data $Z_{1/3}$ with the covariance matrix estimate rescaled for varying cosmological models. The covariance matrix is either estimated without applying the SH marginalisation procedure which corresponds to the Gaussian likelihoods $\mathcal{L}_G(\theta; \hat{\Sigma}(\theta))$ (dashed lines), or has been marginalised using the SH procedure which corresponds to the modified- t likelihoods $\mathcal{L}_t(\theta; \hat{\Sigma}(\theta))$ (dash-dotted lines), or an ‘exact’ covariance matrix is used which corresponds to the Gaussian likelihoods $\mathcal{L}_G(\theta; \Sigma(\theta))$ (dotted lines). The gamma distribution parameters are either estimated (lines with ‘+’ markers) or exact (lines without markers). Similar to Figure 3.8, the top panel shows the inferred f_{NL} posterior percentiles for the different likelihood functions against posterior percentiles of the exact true likelihood $\mathcal{L}_{\text{true}}(\theta; Y)$, with results having been averaged over 25000 data realisations produced at the fiducial cosmological model with $f_{\text{NL}} = 0$. The bottom panel shows the residuals of f_{NL} posterior percentiles with respect to the exact true likelihood. The *left column* shows the case where $N_{\text{mock}} = 1000$ mock catalogues per data realisation have been used for estimation, whereas the *right column* shows the case with $N_{\text{mock}} = 50$. The dotted vertical lines mark the f_{NL} posterior median estimate and the equal-tailed 68.3% credible interval from the exact true likelihood. Note that the vertical axes in the bottom left and right panels have different scales. Figure adapted from Fig. 6 in Paper I.

spectrum and its variance at a fiducial cosmological model with parameters $\check{\theta}$,

$$\check{R}_a = \frac{\mathbb{E}[\widehat{P}_a]^2}{\text{var}[\widehat{P}_a]} \quad \text{and} \quad \check{\eta}_a = \frac{\text{var}[\widehat{P}_a]}{\mathbb{E}[\widehat{P}_a]}. \quad (3.36)$$

If the power spectrum variance could not be predicted analytically, it should be replaced by a fiducial estimate $\widehat{\text{var}}[\widehat{P}_a]$ calculated from mock catalogues.

Data Gaussianisation. The Box–Cox transformation is applied to Gaussianise the band power data vector \widehat{P} univariately,

$$\widehat{P}_a \mapsto Z_a = \widehat{P}_a^{\nu_a}, \quad (3.37)$$

where the transformation parameter ν_a is either calculated using the fitting formula (eq. 3.24) with the fiducial shape–scale parameters $(\check{R}_a, \check{\eta}_a)$, or simply fixed at $\nu_a = 1/3$ which has been shown to perform equally well. After the transformation, the mean $\mu_a(\theta)$ and variance $\sigma_a^2(\theta)$ of the Gaussianised random variable Z_a at any cosmological parameters θ are given by equation (3.20).

Covariance estimate rescaling. For any cosmological model to be tested, the parameter-dependent covariance matrix estimate $\widehat{\Sigma}(\theta)$ can be rescaled from a fiducial estimate $\widehat{\Sigma}_f$ using equation (3.26), where entries of the diagonal matrix Λ are given by $\sigma_a(\theta)$.

Likelihood evaluation. The modified- t likelihood (eq. 3.28) is evaluated with the Gaussianised data vector Z , the mean vector $\boldsymbol{\mu}_Z$ with entries being the transformed mean $\mu_a(\theta)$, and the rescaled covariance matrix estimate $\widehat{\Sigma}(\theta)$. Standard Bayesian inference techniques such as MCMC can be readily employed to extract cosmological parameter estimates and associated uncertainties from the posterior distribution in a multidimensional parameter space.

3.5 Summary and discussion

In preparation for next-generation galaxy surveys such as DESI and *Euclid* measuring large-scale cosmic fluctuations never probed before, we have revisited the Gaussian likelihood assumption commonly found in galaxy clustering power spectrum analyses, which may adversely impact cosmological parameter inference from measurements limited by the number of clustering modes on the largest survey scales. Extending previous work by ref. [169–173], we have carefully derived the underlying distribution of windowed power spectrum measurements for GRFs. To deal with the non-normal probability distribution and the cosmological parameter dependence of the covariance matrix, we have

- 1) devised a univariate Gaussianisation scheme using the Box–Cox transformation to improve data normality;
- 2) proposed a variance–correlation decomposition of the covariance matrix to allow for varying cosmological models;
- 3) synthesised a simple pipeline for straightforward application of the new methodology.

In conclusion, parameter dependence of the covariance matrix can give rise to significant systematic effects if unaccounted for in likelihood inference, as it is an integral part of the shape of the posterior distribution. Our rescaling method for the covariance matrix is valid in the linear perturbative regime and correctly includes this parameter dependence. Recognising and mitigating the non-normality of the underlying distribution of power spectrum measurements is equally crucial; although asymptotic normality may be valid for measurements made from a large number of clustering modes below the largest survey scales, we still recommend the use of our Gaussianisation scheme because of its simplicity.

By testing with Monte Carlo simulations, we have shown that the novel likelihood (eq. 3.28) derived with our proposed methodology closely matches the shape of the exact true likelihood that is inaccessible in realistic surveys. When applied to the inference of the local PNG parameter f_{NL} whose constraint is sensitive to large-scale clustering measurements, the novel likelihood outperforms the standard Gaussian likelihood without data Gaussianisation or covariance estimate rescaling in both frequentists' and Bayes estimations. Although the numerical tests performed in this work have focused on the isotropic power spectrum in real space, as explained in section 3.1, the new methodology also applies to redshift-space power spectrum multipoles measured in the local plane-parallel approximation with the Yamamoto estimator. Yet an all-encompassing framework for galaxy clustering power spectrum analysis is still out of reach. Towards the quasi-linear regime where clustering modes of the galaxy overdensity field are no longer independent but coupled due to non-linear gravitational evolution, the GRF description is no longer valid and the power spectrum covariance structure is fundamentally more complex, though these considerations are beyond the scope of this work.

Impact of Relativistic Effects in Large-Scale Clustering

Despite the relativistic nature of gravitational theories governing structure formation, the Newtonian description of fluctuations in the observed distribution of galaxies is usually adequate as relativistic effects are suppressed below the Hubble horizon scale. In the past, the full relativistic modelling of galaxy clustering has been unnecessary to obtain cosmological parameter constraints, as cosmic variance dominates over any corrections. With the next generation of galaxy surveys probing far wider and deeper cosmic volumes, however, such approximate prescriptions might no longer be sufficient to attain unbiased constraints. The necessary relativistic corrections for galaxy clustering observations have been derived by refs. [123–125]. Many subsequent works have demonstrated their importance for constraining cosmological parameters, in particular local PNG f_{NL} , as its scale-dependent signature on large scales may be disguised as relativistic effects [126–130, 194–197]. At the same time, the investigation of relativistic corrections is itself a valuable exercise, as it offers tests of relativistic gravitational theories on cosmological scales [198, 199], including the equivalence principle [200]. Future galaxy surveys like DESI are forecast to deliver the first detections of these relativistic corrections [201].

One crucial aspect of relativistic corrections is that their total amplitude does not only depend on the cosmological and gravitational models, but also on the background number density of the tracer population being examined through its redshift evolution and sensitivity to the luminosity threshold of observations, as respectively captured by parameters known as evolution bias b_e and magnification bias s . Previous works have mostly only considered relativistic effects in Fisher forecasts for f_{NL} by assuming fiducial values of b_e and s , but the exact dependence of these parameters on redshift and the luminosity threshold, as well as how their uncertainties propagate to the observed power spectrum, remains much less clear. In this chapter, we concretise these considerations for quasars, which are an ideal tracer for detecting f_{NL} thanks to their high redshift range and bias. The next section first reviews linear-order general relativistic corrections to galaxy clustering, including contributions from evolution bias b_e and mag-

nification bias s which shall be formally introduced. This motivates the need for determining the tracer *luminosity function* (LF), and based on the previous work by Palanque-Delabrouille et al. [202, hereafter PD2016], we fit the quasar luminosity function with eBOSS QSO measurements before deriving constraints on b_e , s and thus relativistic corrections. We then compare scale-dependent modifications to the quasar clustering power spectrum due to relativistic corrections and due to f_{NL} at different redshifts for two different magnitude thresholds. The chapter concludes by emphasising the need to include luminosity function constraints in forward modelling of relativistic clustering statistics for future galaxy surveys.

4.1 Relativistic effects in galaxy clustering

Whilst the Newtonian description of galaxy clustering is appropriate for observations on sub-horizon scales, as the clustering scale k^{-1} approaches the horizon scale \mathcal{H}^{-1} , the observed galaxy over-density field δ receives relativistic corrections of $O(\mathcal{H}/k)$ or higher that are otherwise suppressed [123–125],

$$\begin{aligned} \delta(\mathbf{r}, z) = & b_1(z) \delta_{\text{m}}(\mathbf{r}, z) - \frac{\hat{\mathbf{r}} \cdot \partial_{\mathbf{r}} \mathbf{v}}{\mathcal{H}(z)} \\ & - g_1(z) \hat{\mathbf{r}} \cdot \mathbf{v} - [b_e(z) - 3] \mathcal{H}(z) \nabla^{-2} \nabla \cdot \mathbf{v} \\ & + \frac{\Phi'(\mathbf{r}, z)}{\mathcal{H}(z)} - [2 - 5s(z)] \Phi(\mathbf{r}, z) + \Psi(\mathbf{r}, z) + g_1(z) \Psi(\mathbf{r}, z) + \dots \end{aligned} \quad (4.1)$$

Here Φ and Ψ are the Bardeen potentials (recall from appendix B); a thick prime denotes a conformal-time derivative; b_1 is the scale-independent bias with respect to the matter density contrast δ_{m} in comoving synchronous gauge; \mathbf{v} is the peculiar velocity in the Newtonian gauge; and g_1 is a dimensionless quantity given by

$$g_1 = \frac{\mathcal{H}'}{\mathcal{H}^2} + \frac{2 - 5s}{\mathcal{H}\chi} + 5s - b_e. \quad (4.2)$$

The quantities b_e and s are the *evolution* and *magnification biases*: a priori they do not necessarily follow from a cosmological model but are rather derived at a given redshift from

$$b_e(z) = -\frac{\partial \ln \bar{n}(z; < \bar{m})}{\partial \ln(1+z)} \quad \text{and} \quad s(z) = \left. \frac{\partial}{\partial m} \right|_{\bar{m}} \lg \bar{n}(z; < m) \quad (4.3)$$

with $\lg \equiv \log_{10}$, where $\bar{n}(z; < m)$ is the background comoving number density of the tracer population below a given absolute magnitude m , and \bar{m} is the absolute-magnitude threshold of any given tracer sample.

In equation (4.1), the first line is the familiar Newtonian redshift-space over-density field, which includes the linear RSD contribution albeit in the distant-observer limit; the second and third lines are relativistic corrections. These relativistic corrections do not indicate the fundamental behaviour of gravity is any different on cosmological scales (cf. the comoving density contrast Δ defined in appendix B); rather, they arise as the observed coordinates of galaxies do not reflect their true spatial positions because of geometric distortions, as the redshift and propagation of the photon and the observed comoving volume are perturbed. The second

line in equation (4.1) encapsulates Doppler-like terms associated with the peculiar velocity \mathbf{v} ; interestingly, the first term actually provides the mode-coupling RSD term ignored in the distant-observer limit (see eq. 1.31)—it corresponds to the term $(2 - 5s)/(\mathcal{H}\chi)$ in g_1 [199]. The third line contains potential terms, which include the effect of gravitational redshift; the ellipsis encompasses contributions from lensing magnification, time delay and the integrated Sachs–Wolfe (ISW) effect, which are omitted here since they are integrated terms involving the Bardeen potentials and cannot be easily included in a Cartesian power spectrum model. All of these corrections may affect cosmological parameter inference, as shown by recent studies of their relative importance in observable statistics such as the angular power spectrum or correlation function [195, 197, 203–205]. For the full derivation of relativistic corrections to first order, which takes into account the photon trajectory from the source to the observer, the reader is referred to refs. [123–125].

The evolution and magnification biases b_e and s both modulate part of the relativistic corrections. Without magnification, galaxies with intrinsic brightness below the threshold would not appear in the observed sample; with magnification or de-magnification, however, galaxies fainter or only slightly brighter than the threshold may enter or fade out of the sample, and the resulting change to the measured number count is thus described by a non-zero magnification bias s . The effect of evolution bias b_e can also be understood intuitively: relativistic perturbations change the photon redshift along its trajectory, leading to a misestimation of the background number density at the source redshift and thus altering the observed galaxy overdensity field, unless the background number density happens to be independent of redshift, i.e. $b_e = 0$.

In this work, we shall focus on the Doppler-like terms involving the peculiar velocity and the local potential terms only (i.e. all the terms shown explicitly in equation (4.1)), and consider their scale-dependent signature in the global plane-parallel limit where $\mu = \hat{\mathbf{k}} \cdot \hat{\mathbf{r}} = \hat{\mathbf{k}} \cdot \hat{\mathbf{n}}$ does not vary for a fixed global line of sight $\hat{\mathbf{n}}$ (see §1.2). Using the background and linearised Einstein field equations (A.13) and (B.9) for a flat Λ CDM universe written in the following form,

$$\mathbf{v} = -i \frac{\mathcal{H}}{k} f \delta_m \hat{\mathbf{k}}, \quad (4.4a)$$

$$\Phi = -\frac{3}{2} \left(\frac{\mathcal{H}}{k} \right)^2 \delta_m, \quad (4.4b)$$

$$\mathcal{H}^{-1} \Phi' = \left(\frac{\mathcal{H}'}{\mathcal{H}^2} - 1 \right) \left(\frac{\mathcal{H}}{k} \right)^2 f \delta_m - \Phi, \quad (4.4c)$$

where $\Phi = \Psi$ in the absence of anisotropic stress [126, 127, 194], equation (4.1) can be recast as

$$\delta(\mathbf{k}) = \left[b_1 + f\mu^2 + i \frac{\mathcal{H}}{k} g_1 f \mu + \left(\frac{\mathcal{H}}{k} \right)^2 g_2 \right] \delta_m(\mathbf{k}), \quad (4.5)$$

where redshift dependence is implicit, and a second dimensionless quantity has been introduced,

$$g_2 = -(b_e - 3)f + \left(\frac{\mathcal{H}'}{\mathcal{H}^2} - 1 \right) [g_1 + f - (2 - 5s)]. \quad (4.6)$$

By employing the Friedman equations (A.13) with negligible radiation, one can rewrite

$$\frac{\mathcal{H}'}{\mathcal{H}^2} = 1 - \frac{3}{2}\Omega_m; \quad (4.7)$$

the dimensionless quantities parametrising relativistic corrections are thus

$$g_1 = \left(3 - b_e - \frac{3}{2}\Omega_m\right) - (2 - 5s)\left(1 - \frac{1}{\mathcal{H}\chi}\right) \quad (4.8a)$$

$$\text{and } g_2 = \left(3 - b_e - \frac{3}{2}\Omega_m\right)f - \frac{3}{2}\Omega_m[g_1 - (2 - 5s)]. \quad (4.8b)$$

These quantities depend not only on the cosmological model parameters through the accelerating background expansion but also on the tracer sample in question through its evolution and magnification biases.

Therefore to determine the relativistic corrections in equation (4.1) or (4.5), two ingredients are needed: a background cosmological model and the tracer *luminosity function* (LF) $\phi(m, z)$ from which the underlying comoving number density

$$\bar{n}(z; < \bar{m}) = \int_{-\infty}^{\bar{m}} dm \phi(m, z) \quad (4.9)$$

below some absolute-magnitude threshold \bar{m} can be derived.

4.2 Luminosity function model fitting

Determining the tracer luminosity function is not only important for modelling relativistic corrections; it may also be a significant source of uncertainty for constraining PNG. In this work, we examine quasars as a single tracer for detecting f_{NL} thanks to their high tracer bias and redshift range, for which the scale-dependent signature in its large-scale bias should be enhanced (see eq. 1.60). By measuring the quasar evolution and magnification biases from the luminosity function, one can then derive relativistic corrections to the power spectrum multipoles as well as their uncertainties.¹

To this end, we consider the eBOSS QSO LF measurements reported in PD2016 (Tab. A.1 therein) for the redshift range $0.7 < z < 4.0$, which are corrected for observational systematics such as completeness and bandpass redshifting of spectra (also known as the *K*-correction). One of the currently adopted empirical quasar luminosity function models is the *pure luminosity evolution* (PLE) model [202, 207, 208],

$$\phi(m, z) = \frac{\phi_*}{10^{0.4(\alpha+1)[m-m_*(z)]} + 10^{0.4(\beta+1)[m-m_*(z)]}}, \quad (4.10)$$

which is a double power law with bright- and faint-end indices α and β that may differ depending on the redshift z relative to the pivot redshift $z_p = 2.2$. Here ϕ_* is the overall normalisation constant, and

$$m_*(z) = m_*(z_p) - \frac{5}{2} \left[k_1(z - z_p) + k_2(z - z_p)^2 \right] \quad (4.11)$$

¹ The same procedure also applies to tracers other than quasars, as long as one could determine their luminosity function; for instance, ref. [206] provides empirical luminosity functions for H α emitters, which could be used to constrain their evolution and magnification biases by propagating the fitted luminosity function parameters.

is the characteristic absolute magnitude, where k_1 and k_2 are redshift evolution parameters that can also differ between low redshift $z < z_p$ and high redshift $z > z_p$. Therefore this is a parametric model with 10 parameters, $\theta = \{\phi_*, m_*(z_p), \alpha_l, \beta_l, k_{1l}, k_{2l}, \alpha_h, \beta_h, k_{1h}, k_{2h}\}$, where subscripts ‘l’ and ‘h’ denote ‘low redshift’ and ‘high redshift’ respectively.²

4.2.1 Luminosity function likelihood

Without re-performing the iterative LF fitting procedure adopted by PD2016 on the raw QSO number count data, we take the likelihood inference approach outlined in ref. [206] for simplicity. For absolute magnitude and redshift bins (m_i, z_i) indexed by i , the quasar number count \hat{N}_i follows the Poisson distribution with variance $\text{var}[\hat{N}_i] = N_i$, which has the logarithmic PDF

$$\ln \mathbb{P}(\hat{N}_i | N_i) = \hat{N}_i \ln N_i - N_i - \ln \Gamma(\hat{N}_i). \quad (4.12)$$

The expected number count is given by

$$N_i = \langle \hat{N}_i \rangle = \iint_{\text{ith bin}} dz \frac{dV(z)}{dz} dm \phi_\theta(m, z), \quad (4.13)$$

where $\phi_\theta(m, z)$ is the PLE LF (eq. 4.10) with model parameters θ ,

$$dV(z) = 4\pi r^2 \frac{dr}{dz} dz \quad (4.14)$$

is the differential comoving volume, and the radial comoving distance is given by $r = \chi(z)$.

To obtain an approximate likelihood for the parametric LF model, one notes that the binned LF $\hat{\phi} \propto \hat{N}$ and thus an estimate for the uncertainty on $\ln \hat{\phi}$ is $\hat{N}^{-1/2}$. Expanding the PDF (eq. 4.12) around its maximum, one obtains the quadratic form

$$\ln \mathcal{L}(\theta) - \ln \mathcal{L}_{\max} \simeq -\frac{1}{2} \sum_i \frac{\zeta_i^2}{\hat{\sigma}_i^2}, \quad (4.15)$$

where $\hat{\sigma}_i^2 = 1/\hat{N}_i$ and

$$\zeta_i^2(\theta) = -2 \left[1 - \frac{\phi_\theta(m_i, z_i)}{\hat{\phi}_i} + \ln \frac{\phi_\theta(m_i, z_i)}{\hat{\phi}_i} \right] > 0. \quad (4.16)$$

Therefore the appropriate approximate likelihood for inferring the best-fitting LF model is given by

$$\ln \mathcal{L}(\theta) = \sum_i \frac{1}{\hat{\sigma}_i^2} \left[1 - \frac{\phi_\theta(m_i, z_i)}{\hat{\phi}_i} + \ln \frac{\phi_\theta(m_i, z_i)}{\hat{\phi}_i} \right], \quad (4.17)$$

where an additive normalisation constant has been omitted.

² Note that the letter α has a different usage outside the scope of this chapter, and except for the letter k with the subscripts specified here, k generally denotes a wavenumber.

Table 4.1. Posterior median estimates of the PLE model parameters for the eBOSS QSO LF measurements from PD2016.

Parameters	$\lg \phi_*$	$m_*(z_p)$	α	β	k_1	k_2
$z \in [0.68, 2.2)$	$-26.20^{+0.21}_{-0.20}$	$-5.76^{+0.09}_{-0.08}$	$-3.27^{+0.17}_{-0.19}$	$-1.40^{+0.06}_{-0.06}$	$-0.10^{+0.08}_{-0.09}$	$-0.40^{+0.06}_{-0.06}$
$z \in [2.2, 4.0]$			$-2.57^{+0.08}_{-0.09}$	$-1.21^{+0.10}_{-0.09}$	$-0.37^{+0.09}_{-0.09}$	$-0.01^{+0.06}_{-0.06}$

4.2.2 Best-fitting models

By sampling the model parameters from the posterior of the LF likelihood function (4.17) with the MCMC sampler ZEUS [209],³ we have re-fitted the PLE model for the eBOSS QSO measurements. Because of the exchange symmetry between the power-law indices α and β in equation (4.10), the constraint $\alpha < \beta$ has been imposed in the MCMC sampling to avoid a multimodal posterior distribution. The PLE model parameters are estimated by the posterior medians and reported in Table 4.1, with a reduced chi-square value of $105/77 \approx 1.36$ per degree of freedom.

We note that there appears to be some discrepancy between our fitted model parameters and the results in PD2016. In Figure 4.1, both best-fitting models are compared with the eBOSS QSO measurements. In all redshift bins, the two fitted models evaluated at the central redshift are in reasonable agreement with measurements, and are virtually indistinguishable across a wide magnitude range. Noticeable differences only appear either at the very faint end below the limiting absolute magnitude corresponding to the g -band apparent-magnitude cut $g = 22.5$,⁴ which is not constrained by any measurements, or at the very bright end, where uncertainties are comparatively large owing to smaller QSO sample size. Such discrepancies may be attributed to the fact that PD2016 were able to fit the raw QSO number counts whereas here we have only fitted the binned LF reported in their final results.⁵ As the next section will show, constraints on the relativistic corrections propagated from these best-fitting LF models are broadly statistically consistent and have no significant impact on the findings of our analysis.

4.3 Constraints on relativistic corrections

Having determined the QSO LF, we now proceed to constrain relativistic corrections to quasar clustering statistics by propagating the sampled LF parameters in the form of MCMC chains to evolution and magnification biases. To do so, we specify the *Planck* 2015 cosmology with $(h, \Omega_{\Lambda,0}, \Omega_{m,0}) = (0.6790, 0.6935, 0.3065)$ [211], which is a choice consistent with PD2016. We

³ The code is publicly available at github.com/minaskar/zeus.

⁴ The g -band apparent magnitude is converted to the absolute magnitude by $\bar{m}(z) = g - \mu(z) - [K(z) - K(z = 2)]$, where μ is the distance modulus, K is the K -correction function, and the normalisation redshift $z = 2.0$ is close to the median redshift of the eBOSS QSO sample in PD2016. Note that the letters μ and K denote different quantities outside the scope of the discussion here.

⁵ It is also worth mentioning that recently ref. [210] noted a possible error in the K -correction applied to the eBOSS QSO data sets by PD2016, which could have an impact on the fitted LF model.

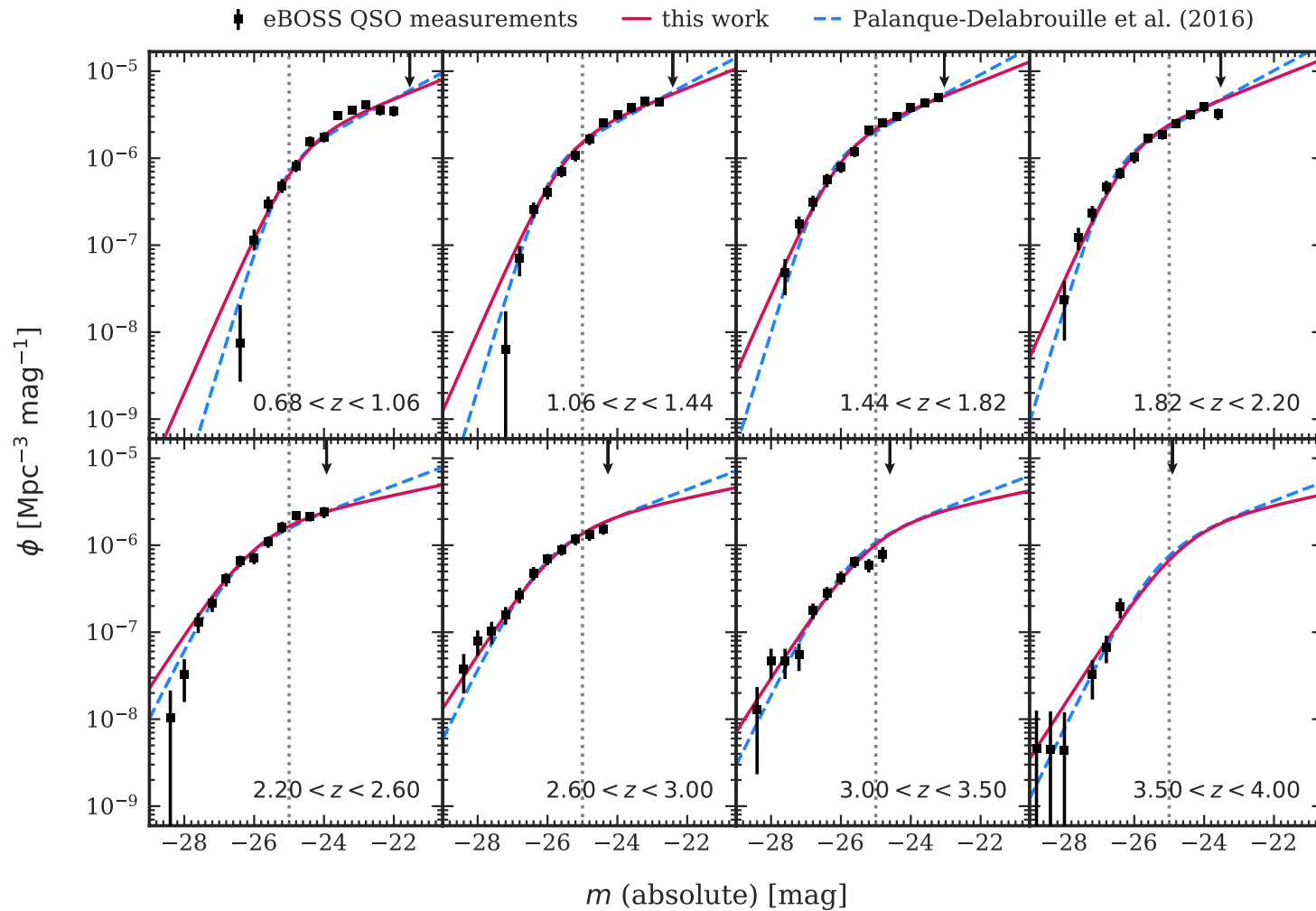


Figure 4.1. Best-fitting QSO LFs under the PLE model evaluated at the central redshift in eBOSS redshift bins. Measurements with uncertainties and the best-fitting model shown in dashed blue lines are taken from PD2016. The best-fitting model of this work (see Table 4.1) is shown in solid red lines. The downward pointing arrows mark the limiting absolute magnitude corresponding to the apparent-magnitude cut $g = 22.5$ in each redshift bin. The vertical dotted lines mark the absolute-magnitude threshold $\bar{m} = -25.0$ used in this work. Figure taken from Fig. 1 in Paper II.

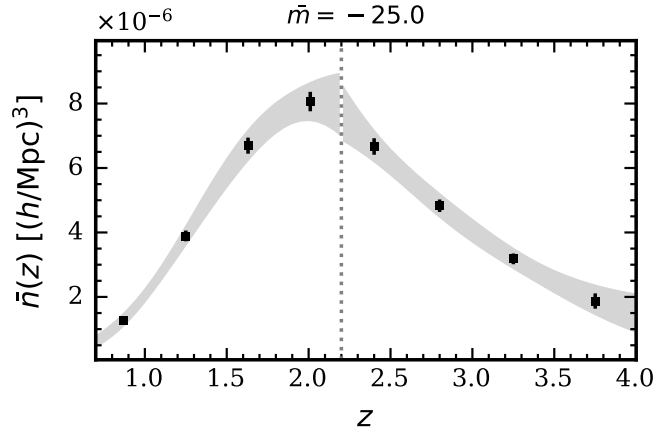


Figure 4.2. Derived measurements of the quasar comoving number density \bar{n} below the absolute-magnitude threshold $\bar{m} = -25.0$ as a function of redshift z from the best-fitting eBOSS QSO LF in this work (see Table 4.1). Data points with error bars are measurements within the 68 % credible interval for the eBOSS QSO redshift bins. The shaded grey regions show the 95 % credible interval. The vertical dotted line marks the pivot redshift $z_p = 2.2$. Figure taken from Fig. 2 in Paper II.

also specify a fiducial absolute-magnitude threshold $\bar{m} = -25.0$ based on the highest eBOSS QSO redshift bin.

4.3.1 Constraints on relativistic biases

To first compute the quasar comoving number density \bar{n} , equation (4.9) is numerically integrated with our best-fitting LF model $\phi_\theta(m, z)$ up to the absolute-magnitude threshold \bar{m} . In Figure 4.2, the derived measurements of \bar{n} from sampled LF parameters within the 95 % credible interval across the redshift range $0.7 < z < 4.0$ are shown; for the eBOSS QSO redshift bins, measurements of \bar{n} with error bars corresponding to the 68 % credible interval are also shown. The small apparent discontinuity in \bar{n} corresponds to the pivot redshift z_p , which divides some subsets of the combined eBOSS QSO data [PD2016]. The presence of the pivot redshift z_p is also a common feature of the empirical models currently used for the quasar luminosity function, where the model parameters can suddenly change. This may have possible links to the physics of quasar formation around that epoch in history and/or the fact that the double power-law form assumed for the quasar luminosity function is no longer adequate at higher redshifts, where it is more poorly constrained by current data [210, 212].

Next, evolution bias b_e and magnification bias s are computed by numerical differentiation with redshift step size $\Delta z = 0.001$ (see eq. 4.3). We have found that, based on the eBOSS QSO LF measurements, b_e can be an order of magnitude larger than s , although both appear in relativistic corrections at the same orders in equation (4.8). One interesting comparison one might make for b_e is with the analytic prediction from the universal mass function (UMF) of halos (see also appendix C), although the validity of this approach is only limited to tracer sample selection that is insensitive to halo merger history. The evolution bias predicted from the UMF is given by

$$b_e(z) = \delta_c f(z) [b_1(z) - 1], \quad (4.18)$$

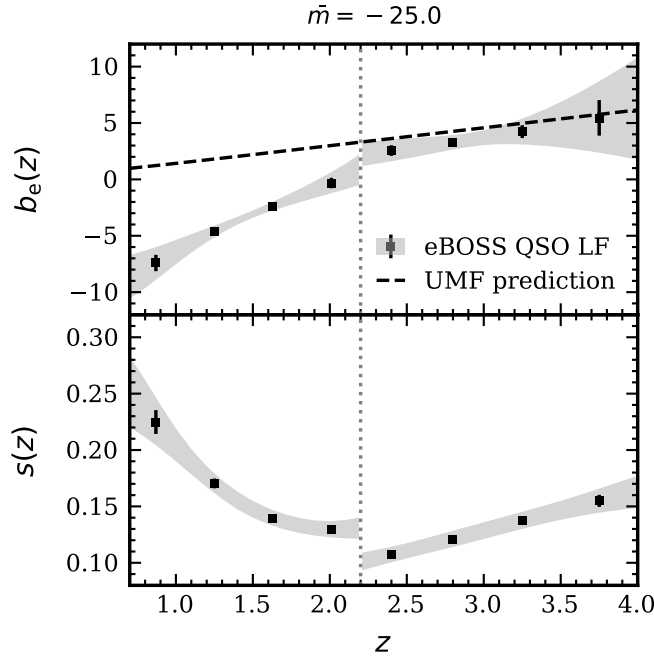


Figure 4.3. Derived measurements of evolution bias b_e and magnification bias s at the absolute-magnitude threshold $\bar{m} = -25.0$ from the best-fitting eBOSS QSO LF in this work (see Table 4.1). Data points with error bars are measurements within the 68 % credible interval for the eBOSS QSO redshift bins. The shaded grey regions show the 95 % credible interval. The vertical dotted lines mark the pivot redshift $z_p = 2.2$. The cause of the discontinuities at z_p in both b_e and s is unclear and could be attributed to unknown systematics in the high-redshift QSO sample [221]. Figure taken from Fig. 3 in Paper II.

where $\delta_c \approx 1.686$ is the critical over-density of spherical collapse (recall from appendix C) [127]. Here we consider a simple redshift evolution model for the quasar linear bias $b_1(z) = 1.2/D_1(z)$, the value of which increases from 1.7 to 4.7 almost linearly in the eBOSS QSO redshift range $0.7 < z < 4.0$. This bias model is based on the DESI baseline survey and does not account for possible luminosity dependence [213]. Based on power-law fitting to the observed quasar clustering amplitudes, studies have found that the luminosity dependence of quasar bias appears to be rather weak, at least at low and intermediate redshifts possibly because quasars reside in a broad range of halos of different masses [214–216]. However, some current models and observations hint at a greater level of luminosity dependence at higher redshifts and luminosity ranges, but such quasars are rare and the luminosity dependence of their bias can only be better constrained with larger data sets in the future [217–220].

In Figure 4.3, the derived measurements of b_e and s for $0.7 < z < 4.0$ within the 95 % credible interval and in eBOSS QSO redshift bins with 68 % level uncertainties are shown together with the UMF prediction. Similar to the constraints on comoving number density \bar{n} , uncertainties of b_e and s at each redshift are derived from sample values calculated from MCMC chains of the luminosity function parameters (which may differ on different sides of the pivot redshift z_p). It is apparent that, although the UMF prediction is in reasonable agreement with the measurements at high redshifts, it does not capture the behaviour of the negative evolution bias values below redshift $z \approx 2$. This is perhaps unsurprising given the limitation of the

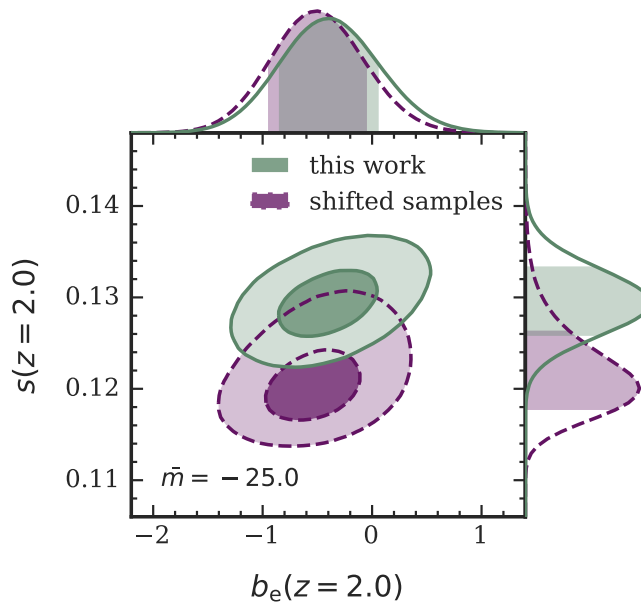


Figure 4.4. Constraints on evolution bias b_e and magnification bias s at redshift $z = 2.0$ for absolute-magnitude threshold $\bar{m} = -25.0$ from the eBOSS QSO LF measurements under the PLE model. The solid green contours show the 68 % and 95 % credible regions of the joint posterior distribution sampled from the likelihood function (eq. 4.17) (‘this work’). The dashed purple contours are from the same samples except shifted to coincide with the best-fitting PLE model parameters from PD2016 (‘shifted samples’). The shaded regions in the top and right panels show the 68 % credible intervals of the marginal posterior distributions. Figure taken from Fig. 4 in Paper II.

UMF prediction and the simplicity of the quasar bias model we have used. As is the case for comoving number density, there is an apparent discontinuity at the pivot redshift $z_p = 2.2$ in both b_e and s . However, these discontinuities are now large enough that even the 95 % uncertainty bounds are inconsistent across the pivot redshift. Unfortunately, we have checked that this problem persists with the LFs fitted by PD2016 and ref. [210], so it is not due to the different fitting procedures. Although the cause of these discontinuities has been attributed to the form of the empirical luminosity function, the largeness of the discrepancies may indicate unknown systematics in the eBOSS QSO sample at high redshifts, as noted by ref. [221]. Again, future survey data may hopefully be able to resolve this issue.

In section 4.2.2, we have also noted that our best-fitting LF under the PLE model is somewhat discrepant from that of PD2016 for the same underlying eBOSS QSO sample (possibly affected by unknown systematics), although their model parameter estimates have similar uncertainties. To investigate the impact of this on the measured relativistic bias parameters, we shift the MCMC chains of sampled LF parameters so that the shifted posterior median estimates coincide precisely with the best-fitting PLE parameters in PD2016. The shifted parameter samples are then propagated to constraints on b_e and s . Figure 4.4 shows that the joint b_e – s constraints from our original parameter samples and the shifted samples, for instance at redshift $z = 2.0$, are broadly consistent. This is especially the case for evolution bias b_e , which dominates the relativistic corrections over magnification bias s . We have also checked that the joint b_e – s constraints from the original and shifted samples are consistent at other redshifts,

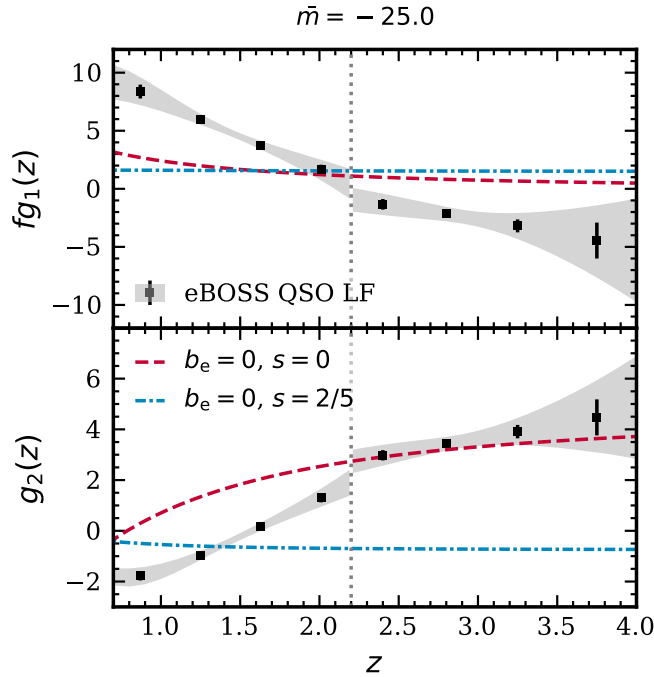


Figure 4.5. Derived measurements of the relativistic correction parameters fg_1 and g_2 (see eq. 4.8) at the absolute-magnitude threshold $\bar{m} = -25.0$ from the best-fitting eBOSS QSO LF in this work (see Table 4.1). The data points with error bars are measurements within the 68 % credible interval for the eBOSS QSO redshift bins, and the shaded region shows the 95 % credible interval. For comparison, the dashed red lines show the results with $(b_e, s) = (0, 0)$ and the dash-dotted blue lines with $(b_e, s) = (0, 2/5)$. The vertical dotted line marks the pivot redshift $z_p = 2.2$. Figure taken from Fig. 5 in Paper II.

e.g. $z = 0.87$ and 3.75 which correspond to the lowest and highest eBOSS QSO redshift bins respectively.

4.3.2 Constraints on relativistic correction parameters

In section 4.1, it is shown that relativistic corrections at $O(\mathcal{H}/k)$ and $O(\mathcal{H}^2/k^2)$ to the galaxy over-density field at different redshifts and scales are modulated by the parameters fg_1 and g_2 respectively, which can be constrained from the relativistic bias measurements obtained above under a given cosmological model. Figure 4.5 shows the derived bounds on fg_1 and g_2 within the 95 % credible interval and their measurements in eBOSS QSO redshift bins with 68 % level uncertainties. The discontinuities in the derived g_1 and g_2 have the same origin as those discussed previously. Both g_1 and g_2 as well as their uncertainties are dominated by contributions from evolution bias b_e , which can be an order of magnitude larger than magnification bias s , as shown in Figure 4.3. To assess the relative importance of b_e and s , we show in Figure 4.5 two particular fiducial cases: 1) $(b_e, s) = (0, 0)$, i.e. no account of the redshift evolution and luminosity dependence of the quasar number density; 2) $(b_e, s) = (0, 2/5)$, i.e. the comoving number density is constant and the common factor $(2 - 5s)$ in various relativistic correction terms vanishes, corresponding to the so-called ‘diffuse background’ scenario where the effects of lensing magnification and volume distortions partly cancel [127]. Comparisons with these

cases demonstrate that evolution bias b_e drives relativistic corrections at both low and high redshifts; unless $(2 - 5s)$ vanishes, terms containing the $(\mathcal{H}\chi)^{-1}$ factor are also important and increasingly so at lower redshift, especially towards $z = 0$ beyond the redshift range shown in the figures. This highlights the importance of including both b_e and s accurately in modelling relativistic corrections to galaxy clustering.

Having propagated quasar luminosity function measurements through to constraints on relativistic corrections, we shall investigate in the following section how they modify the quasar clustering power spectrum multipoles on large scales.

4.4 Scale-dependent modifications

In section 1.3, it is shown that in the presence of local primordial non-Gaussianity f_{NL} , the linear tracer bias b_1 receives a scale-dependent modification

$$\Delta b(k, z) = 3f_{\text{NL}}[b_1(z) - 1.6] \frac{1.27\delta_c\Omega_{\text{m},0}H_0^2}{c^2k^2 T(k) D_1(z)}, \quad (4.19)$$

where the speed of light c is restored, the tracer-dependent parameter is set to a fiducial value of $p = 1.6$ here for quasars, and the numerical factor arising from normalisation of the linear growth factor is $\mathcal{C}(\Omega_{\text{m},0}) \approx 1.27$ for the *Planck* 2015 Λ CDM cosmology (cf. eq. 1.60). As $k \rightarrow 0$, the transfer function $T(k) \rightarrow 1$ and $\Delta b \propto k^{-2}$, so the scale-dependent modification is enhanced on large scales.

It is also shown in section 1.2 that the global plane-parallel anisotropic power spectrum in the Kaiser RSD model is given by

$$P^{(\text{K})}(k, \mu) = (b + f\mu^2)^2 P_{\text{m}}(k), \quad (4.20)$$

which is equivalent to the combination of its Legendre multipoles of degrees $\ell = 0, 2$ and 4. With the total bias b now inclusive of both the scale-independent bias b_1 and the scale-dependent modification $\Delta b \propto k^{-2}$, the standard Kaiser multipoles $P_{\ell}^{(\text{K})}$ receive the following modifications,

$$\Delta P_0(k) = \left[\left(2b_1 + \frac{2}{3}f \right) \Delta b + \Delta b^2 \right] P_{\text{m}}(k), \quad (4.21a)$$

$$\Delta P_2(k) = \frac{4}{3}f \Delta b P_{\text{m}}(k). \quad (4.21b)$$

In contrast to the quadrupole which only receives a modification proportional to k^{-2} , the monopole receives modifications proportional to both k^{-2} and k^{-4} when $f_{\text{NL}} \neq 0$, whereas the hexadecapole receives no modifications as it does not depend on the tracer bias (see eq. 1.42).

In the case of relativistic corrections, it is shown in section 4.1 that they similarly leave a scale-dependent signature in the over-density field. By considering the two-point correlator of the Fourier clustering mode in equation (4.5), one sees that relativistic corrections change the Kaiser monopole and quadrupole by

$$\Delta P_0(k) = \left[\left(2b_1g_2 + \frac{2}{3}fg_2 + \frac{1}{3}f^2g_1^2 \right) \left(\frac{\mathcal{H}}{k} \right)^2 + g_2^2 \left(\frac{\mathcal{H}}{k} \right)^4 \right] P_{\text{m}}(k) \quad (4.22a)$$

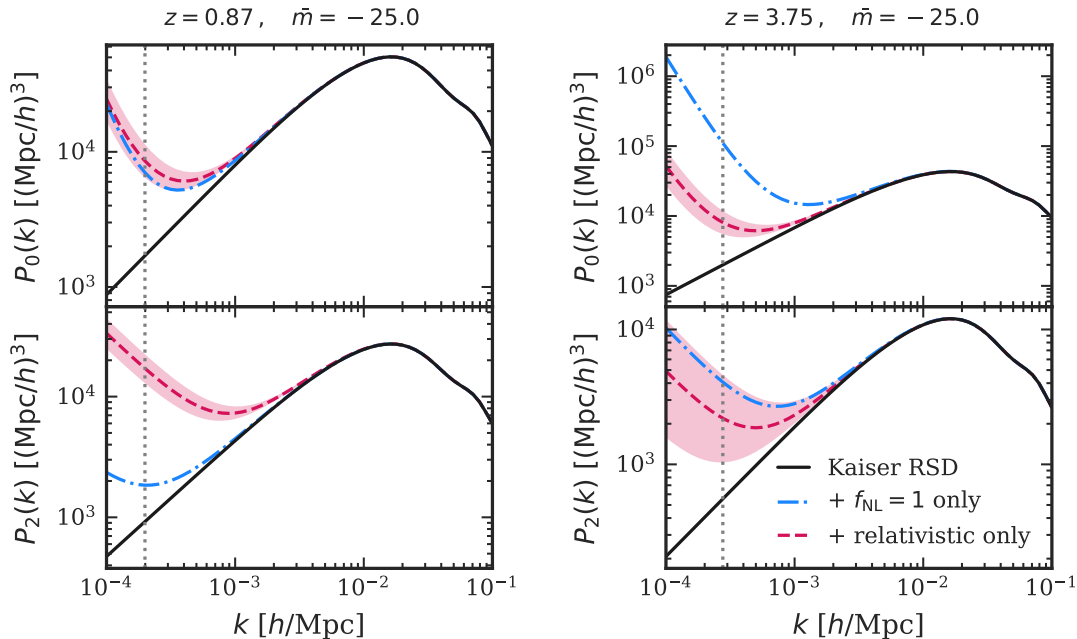


Figure 4.6. Quasar clustering power spectrum monopole P_0 and quadrupole P_2 , without and with scale-dependent modifications, at redshift $z = 0.87$ (left column) and $z = 3.75$ (right column) with the absolute-magnitude threshold $\bar{m} = -25.0$. The Kaiser RSD model is shown by the solid black lines. The effect of relativistic corrections without local PNG f_{NL} is shown by the dashed red lines with the shaded red regions showing the 95 % credible interval derived from the best-fitting eBOSS QSO LF in this work (see Table 4.1). The effect of $f_{\text{NL}} = 1$ (fiducial value) without relativistic corrections is shown by the dash-dotted blue lines. The vertical dotted lines mark the wavenumber $k = \mathcal{H}$ at horizon scale. Figure adapted from Fig. 6 in Paper II.

$$\text{and } \Delta P_2(k) = \frac{2}{3} (2fg_2 + f^2g_1^2) \left(\frac{\mathcal{H}}{k} \right)^2 P_m(k). \quad (4.22b)$$

By comparing equations (4.21) and (4.22), it is evident that relativistic corrections can mimic the effect of f_{NL} in both the power spectrum monopole and quadrupole on large scales; the extent to which relativistic corrections can wash out the f_{NL} signal depends on the precise amplitudes of the relativistic correction parameters g_1 and g_2 .

With the constraints on g_1 and g_2 obtained in the previous section, we can now make a concrete comparison between the power spectrum multipole modifications due to f_{NL} and due to relativistic corrections. To this end, we consider a fiducial value $f_{\text{NL}} = 1$ at which level different classes of inflation models can be distinguished [85, 86]; as before, the fiducial cosmology is the *Planck* 2015 Λ CDM model, the absolute-magnitude threshold remains $\bar{m} = -25.0$ for the eBOSS QSO sample, and $b_1(z) = 1.2/D(z)$ for the DESI baseline survey, which is used earlier in the UMF prediction for evolution bias b_e , remains the fiducial bias model.

In Figure 4.6, the Kaiser power spectrum multipoles for $k \in [10^{-4}, 10^{-1}] h \text{Mpc}^{-1}$ are shown with the aforementioned scale-dependent modifications at two redshifts, $z = 0.87$ and 3.75 , which correspond to the lowest and highest eBOSS QSO redshift bins respectively. At the lower redshift, relativistic effects dominate over the f_{NL} signal and obscure the PNG signature. At the higher redshift, although the relativistic modification is almost comparable to the

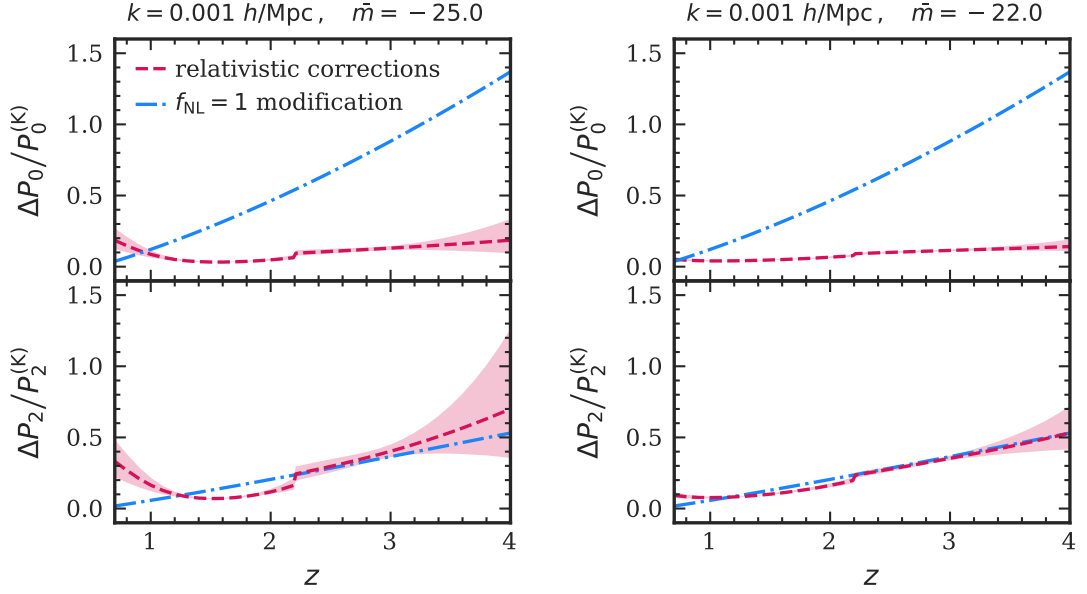


Figure 4.7. Scale-dependent modifications ΔP_ℓ to the quasar clustering power spectrum monopole and quadrupole as a fraction of the Kaiser RSD model $P_\ell^{(K)}$ at wavenumber $k = 0.001 \text{ h Mpc}^{-1}$ with the absolute-magnitude threshold $\bar{m} = -25.0$ (left column) or $\bar{m} = -22.0$ (right-hand column). Relativistic corrections without local PNG f_{NL} are shown by the dashed red lines with the shaded red regions showing the 95 % credible interval derived from the best-fitting eBOSS QSO LF in this work (see Table 4.1). Modifications due to $f_{\text{NL}} = 1$ (fiducial value) without relativistic corrections are shown by the dash-dotted blue lines. Figure adapted from Fig. 7 in Paper II.

effect of f_{NL} in the quadrupole, the f_{NL} signal is larger in the monopole. This offers a hint that, at least for the QSO sample, extending the upper redshift range may help mitigate some potential contamination of the f_{NL} signal from part of the relativistic corrections; however, we caution that lensing convergence and non-local potential terms have not been included in our analysis, and these integrated contributions might hamper the detection of local PNG again at higher redshifts [197, 203, 204].

To compare how the relative amplitudes of relativistic and PNG modifications evolve with redshift, in Figure 4.7 we show the resultant changes in power spectrum multipoles, ΔP_ℓ , as a fraction of the standard Kaiser multipoles $P_\ell^{(K)}$ without scale-dependent modifications, across the eBOSS QSO redshift range $0.7 < z < 4.0$ at a fixed wavenumber $k = 10^{-3} \text{ h Mpc}^{-1}$ —this is close to the largest scale which DESI and *Euclid* may probe [213, 222]. In addition to the fiducial absolute-magnitude threshold $\bar{m} = -25.0$, we also consider a less conservative threshold at $\bar{m} = -22.0$, which is the limiting magnitude of the lowest eBOSS QSO redshift bin. The discontinuities seen in the figure again have the same origin as those found in relativistic bias constraints in the previous section. The key findings from Figure 4.7 are:

- 1) For both monopole and quadrupole, relativistic corrections dominate over the effect of f_{NL} at low redshifts $z \lesssim 1$, and values of ΔP_ℓ due to relativistic effects and due to f_{NL} reach parity at some intermediate redshift below $z \approx 1.5$. The dominance of relativistic effects at lower redshifts is mainly driven by the large evolution bias b_e (see Fig. 4.3) and the geometric factor $(\mathcal{H}\chi)^{-1}$ when $s \neq 2/5$. If $b_e = 0$ and $s = 2/5$ (the

diffuse background scenario), relativistic effects would be much smaller than the f_{NL} signal overall;

- 2) Although relativistic corrections are comparable to the effect of f_{NL} in the quadrupole at most redshifts, the f_{NL} signal is stronger at higher redshifts in the monopole, mainly because of the redshift evolution of the scale-independent tracer bias b_1 and the fact that ΔP_0 due to f_{NL} contains contributions proportional to b_1^2 . If the tracer bias is constant, for instance $b_1 = 2$, then at higher redshifts, relativistic effects with $b_e \neq 0$ and $s \neq 2/5$ will wash out the f_{NL} signal, and can slightly reduce the f_{NL} signal even with $b_e = 0$ and $s = 2/5$;
- 3) Raising the absolute-magnitude threshold tends to reduce the relativistic corrections at all redshifts: we have found that both evolution and magnification biases decrease in absolute values with increasing magnitude threshold, suggesting that future surveys with sensitivity to detect more faint objects may also help with constructing tracer samples with subdued relativistic effects.

So far our comparison has focused on the relative strength of the PNG and relativistic signals, when on large scales cosmic variance may dominate both because of the small number of clustering modes. Depending on the size and maximum redshift of the survey, this can make the detection of either effect difficult though not entirely impossible [124, 125]. Nevertheless, if any excess power on large scales is measured, the analysis above is still useful for disentangling relativistic effects from the scale-dependent bias. Moreover, if one considers a multi-tracer approach where different samples with overlapping volumes are cross-correlated, then the impact of cosmic variance may be significantly reduced, making a clear detection much easier [223, 224]. The relativistic effects shown in the power spectrum above could in principle serve as a consistency test for relativistic gravitational theories: if all relativistic correction terms, including lensing and other integral contributions, are included in the modelling, the evolution and magnification biases can be fitted as free parameters and compared with the constraints obtained from a well-determined tracer luminosity function, although this is challenging in practice not least because the empirical luminosity function models are far from perfect as we saw in section 4.2. The dipole of the cross-power spectrum between two tracer samples is a more promising statistic, which also includes the relativistic bias terms constrained in this work: by parametrising the deviation from the Euler equation (see appendix B), one could use this technique to test the equivalence principle with surveys such as DESI in the near future [199–201].

Finally, it is worth mentioning that in the limit $k \rightarrow 0$, ref. [225] has recently argued that the full relativistic corrections actually vanish as a consequence of the equivalence principle, and thus they do not contaminate the PNG signature; the apparent divergence in ΔP_ℓ as $k \rightarrow 0$ in equation (4.22) is due to the exclusion of non-local contributions as well as contributions at the observer position. For finite clustering scales accessible to galaxy surveys, these relativistic effects do exist and should still be taken into account in a PNG analysis.

4.5 Summary and discussion

Motivated by recent studies of relativistic effects in LSS observations and the prospect of constraining PNG through galaxy redshift surveys to the level of precision competitive with CMB

experiments in the near future, we have sought to quantify relativistic corrections to the observed clustering statistics of quasars on very large scales. These corrections do not only depend on the cosmological expansion history, but also on the redshift evolution of the underlying quasar number density and its sensitivity to the luminosity threshold, which are parametrised by evolution bias b_e and magnification bias s . We have thus refitted the eBOSS QSO LF and derived measurements on both b_e and s , before propagating their constraints to relativistic corrections to the power spectrum multipoles. Our assessment of the impact of relativistic effects on the f_{NL} signature affirms the results of previous works mentioned at the beginning of the chapter, but this agreement is reached after a more realistic treatment for evolution and magnification bias contributions and in particular their uncertainties.

Indeed, relativistic corrections have been found to mimic scale-dependent bias modifications induced by f_{NL} , especially at low redshifts and in the power spectrum quadrupole. By using tracer samples at higher redshifts or with a fainter luminosity threshold, relativistic effects can be reduced to some extent. We have also found that, at least for the quasar population, the impact of evolution bias b_e and its uncertainties on clustering statistics is greater than that of magnification bias s . However, the latter also appears in lensing contributions to the observed galaxy over-density field, which have been neglected in this work along with other integrated terms involving the gravitational potential; these contributions can be important especially at higher redshifts and are best studied in future works with alternative statistics such as the angular power spectrum. Similarly, the plane-parallel limit for power spectrum multipoles has been taken to simplify our arguments, but wide-angle effects due to variation of the line of sight have been shown to be critical on very large scales [143, 144, 226, 227]. Therefore in a practical analysis, wide-angle corrections need to be included perturbatively [78, 142], or a spherical Fourier analysis could prove advantageous [3, 49, 228, 229], as the following chapter will demonstrate. Meanwhile, we have only considered quasars as a single tracer for detecting relativistic effects and the PNG signature; to extract maximal information from future LSS probes, a multi-tracer approach may prove beneficial as it can enhance the signal-to-noise ratio of certain cosmological parameters such as f_{NL} [223, 224].

Our findings have implications for future clustering measurements of the DESI QSO sample, which has an apparent-magnitude limit similar to the one considered in this work [213]. The forecast relativistic corrections can be an order of magnitude larger than the modifications induced by $f_{\text{NL}} = 1$ at wavenumbers $k = O(10^{-3}) h \text{Mpc}^{-1}$ at lower redshifts $z \lesssim 1$; at higher redshifts $z \gtrsim 2$, relativistic corrections remain comparable to or larger than the $f_{\text{NL}} = 1$ modification in the power spectrum quadrupole for absolute-magnitude thresholds up to $\bar{m} = -22$ at least. As seen in section 4.3, potential systematics in the quasar luminosity function can affect the relativistic bias parameters, and therefore the accurate determination of tracer luminosity functions is crucial to constraining relativistic corrections. We suggest that forward modelling from the tracer luminosity function to relativistic corrections should be fully included in future LSS analyses. For this purpose, we have made the code implemented in this work publicly available as a Python package, `HORIZONGROUND`.

Hybrid-Basis Fourier Analysis of Anisotropic Clustering

In linear cosmological perturbation theory, cosmic fluctuations are well described by Gaussian random fields if the initial conditions are also Gaussian. As one recalls from section 1.1, all statistical information is then encoded either in configuration space by the correlation function ξ or in Fourier space by the power spectrum P . Although the underlying matter distribution is expected to be spatially homogeneous and isotropic on large scales under the cosmological principle, the observed clustering of galaxies is not, since their spatial positions have to be inferred from the measured redshift and angular coordinates in the sky; this process can be affected by both RSDs and the AP effect, with apparent anisotropies induced around the line of sight in the observed galaxy distribution. In the seminal Kaiser RSD model introduced in section 1.2, the information contained in clustering measurements is then compressed into the anisotropic power spectrum $P(k, \mu)$ or equivalently its Legendre multipoles $P_\ell(k)$, which are defined with respect to a fixed line of sight \hat{n} in the distant-observer and global plane-parallel approximations; these summary statistics, as well as the underlying assumptions, continue to be adopted in extended non-linear RSD models.

Although such treatments have been justified for past surveys with small sky coverage, or for the analysis of N -body simulations confined to a Cartesian box which can be placed arbitrarily far away from the idealised observer, the intrinsic geometry of survey observations is spherical. For future large galaxy surveys such as DESI and *Euclid* covering almost a third of the sky (approximately 14000 and 15000 square degrees respectively [213, 222]), the observer is at the centre of the much wider cosmic volume being probed—it is impossible to directly measure the global plane-parallel $P(k, \mu)$ or $P_\ell(k)$. As discussed in section 2.2, discrepancies exist between the global plane-parallel predictions and the local plane-parallel estimators, the latter dependent on the choice of the pairwise line of sight \mathbf{d} which can vary considerably across the survey volume. Such discrepancies, known as the wide-angle effect, can be a significant systematic on large scales and degrade constraints on cosmological parameters [67, 142, 230]. It has recently been shown that even for BOSS-like data sets, wide-angle effects coupled to the

survey window can contribute up to 5 % uncertainties in the even power spectrum multipoles, and dominate the odd multipoles which are important to the search for relativistic effects [142].

To deal with this critical issue, new estimators for the two-point correlation function have been proposed, e.g. using expansions in bi-polar or tri-polar spherical harmonics in configuration space [143, 144, 226], which are valid for wide angular separations and also account for the mode-coupling term in the RSD operator (eq. 1.31) that is usually ignored; in Fourier space, perturbative wide-angle corrections have also been derived recently [78, 141, 142]. Yet despite these efforts, the Cartesian power spectrum analysis still encounters challenges on other fronts. For instance, the analysis is tomographic, which requires fine-tuned binning in redshift [e.g. 231, 232]; the covariance matrix, essentially a four-point function, is typically intractable analytically, so a large number of mock catalogues are needed for estimation, but this is computationally costly and often does not take cosmological dependence into account [1, 167], and introduces additional uncertainty into cosmological analyses (see §1.4). At a more fundamental level, the power spectrum analysis is based on the plane wave basis tied to a Cartesian coordinate system. Given much of the physics affecting galaxy clustering besides RSD (e.g. relativistic and light-cone effects) is along the radial line-of-sight direction, and many observational systematics also separate into radial and angular components (e.g. redshift errors and angular variations), a natural question arises as to whether there is an alternative approach to analysing anisotropic clustering on large scales that respects the symmetries of the problem [78].

Indeed, such an approach has been proposed before: the spherical Fourier analysis, also known as spherical harmonic analysis, was first laid down in refs. [49, 228, 233] and subsequently applied to the *IRAS* Point Source Catalog Redshift (PSCz) Survey [234–236], the 2-Micron All-Sky Redshift Survey (2MASS) [237, 238], and the final catalogue of the 2dFGRS [239]. The analysis procedure is in general more complex and computationally expensive, and thus less economical for past surveys covering a small sky fraction. In recent years, there has been renewed interest in this approach [e.g. 229, 240–245] due to the need for a methodology suited for future wide surveys that can probe much larger cosmological scales for studying PNG and relativistic effects. However, because of the computational cost and the difficulty in formulating non-linear galaxy clustering models in the spherical Fourier analysis, the Cartesian power spectrum analysis aided by the FFT is usually favoured.

Inspired by the use of hybrid estimators in CMB studies [177, 185, 246–250], we set out to investigate whether a hybrid-basis approach to likelihood inference from clustering measurements is viable; this is akin to *Planck*'s hybrid likelihoods, where the low- ℓ likelihood is directly built from map pixels and the high- ℓ likelihood is based on the compressed pseudo- C_ℓ estimator [251–253]. We propose a Fourier analysis that uses spherically decomposed statistics to describe anisotropic clustering on the largest scales in the survey, and switches to the Cartesian power spectrum analysis on comparatively smaller scales. The key advantage of this approach is that while the computational edge of the FFT is utilised when the clustering modes are numerous, the spherically decomposed clustering statistics can also faithfully capture physical and observational effects parallel and transverse to the line of sight without making any geometric approximations, which impact on measurements on very large scales in particular.

The following section will first introduce the spherical Fourier analysis extended from the earlier works of Heavens & Taylor [49] and Fisher et al. [228]. The reader may refer to chapter 2 for a review of the standard Cartesian power spectrum analysis. In the subsequent section, the

steps involved in constructing the likelihood function in the hybrid-basis analysis are laid out, and the technical aspects of likelihood evaluation are discussed. The applicability of the new methodology is then demonstrated by inferring the local PNG parameter f_{NL} from N -body simulations.

Throughout this chapter, the analysis of clustering statistics is performed in redshift space, so the subscript ‘s’ denoting redshift-space quantities will be omitted for brevity; in any case, the ambiguity between real- and redshift-space quantities can be eliminated by the context or the coordinates explicitly used.

5.1 Spherical Fourier analysis

5.1.1 Spherical Fourier–Bessel transform

The notion of Fourier transforms can be generalised to harmonic analysis, which describes a field $f(\mathbf{r})$ by its decomposition in an orthogonal basis of eigenfunctions of the Laplacian ∇^2 . Instead of adopting the Cartesian plane wave basis $\{|\mathbf{k}\rangle\}$ with $\langle \mathbf{r} | \mathbf{k} \rangle = e^{i\mathbf{k}\cdot\mathbf{r}}$ in the bra–ket notation, one could decompose clustering measurements in spherical coordinates $\mathbf{r} \equiv (r, \hat{\mathbf{r}})$ using the *spherical Fourier–Bessel* (SFB) transform,

$$f_{\ell m}(k) = \int d^3\mathbf{r} j_\ell(kr) Y_{\ell m}^*(\hat{\mathbf{r}}) f(\mathbf{r}), \quad (5.1a)$$

with the inverse transform given by

$$f(\mathbf{r}) = \frac{2}{\pi} \sum_{\ell m} \int dk k^2 j_\ell(kr) Y_{\ell m}(\hat{\mathbf{r}}) f_{\ell m}(k). \quad (5.1b)$$

The normalisation convention here differs slightly from that adopted in refs. [49, 78, 228, 233, 242]; in refs. [229, 240, 243, 245], the basis functions are chosen to be orthonormal with respect to an inner product integral with a one-dimensional measure; however, in practice, these definitions of the SFB transform are equivalent. If the Dirichlet boundary condition $f(\mathbf{r}) = 0$ is imposed on a sphere at radius $r = R$, the wavenumbers are then discretised,

$$k_{\ell n} = \frac{u_{\ell n}}{R}, \quad (5.2)$$

where $u_{\ell n}$ is the n th positive root of the spherical Bessel function j_ℓ . The associated tuple (ℓ, m_ℓ, n_ℓ) shall be referred to as the *spherical degree, order* and *depth* (though often with the subscript ‘ ℓ ’ suppressed for brevity), where $\ell \in \mathbb{N}$, $n \in \mathbb{N}_{>0}$ and $m = 0, \pm 1, \dots, \pm \ell$.

Therefore akin to the discrete Fourier transform over a regular Cartesian grid with periodic boundary conditions, a field f that vanishes outside some maximum radius R can be expanded in the SFB basis $\{|\mu\rangle \equiv |\ell_\mu, m_\mu, n_\mu\rangle\}$,

$$f(\mathbf{r}) = \sum_{\mu} \kappa_{\mu} f_{\mu} \langle \mathbf{r} | \mu \rangle \quad \text{with} \quad \langle \mathbf{r} | \mu \rangle = j_{\ell_{\mu}}(r) Y_{\ell_{\mu}}(\hat{\mathbf{r}}), \quad (5.3)$$

where the normalisation coefficient is given by

$$\kappa_{\ell n} = \frac{2}{R^3} j_{\ell+1}(u_{\ell n})^{-2}, \quad (5.4)$$

$f_\mu = f_{\ell_\mu m_\mu}(k_{\ell_\mu n_\mu})$, $j_\mu(r) \equiv j_{\ell_\mu}(k_{\ell_\mu n_\mu} r)$, and $Y_\mu(\hat{\mathbf{r}}) \equiv Y_{\ell_\mu m_\mu}(\hat{\mathbf{r}})$. In this basis, the Dirac and Kronecker delta functions are respectively represented by

$$\delta^{(\text{D})}(\mathbf{r}_1 - \mathbf{r}_2) = \sum_\mu \kappa_\mu \langle \mathbf{r}_1 | \mu \rangle \langle \mu | \mathbf{r}_2 \rangle \quad \text{and} \quad \delta_{\mu\nu}^{(\text{K})} = \kappa_\mu \int_{|\mathbf{r}| < R} d^3\mathbf{r} \langle \mu | \mathbf{r} \rangle \langle \mathbf{r} | \nu \rangle. \quad (5.5)$$

5.1.2 Spherical decomposition of clustering statistics

Since the SFB transform does not require the calculation of transverse distances from the angular position of galaxies, the RSD and AP effects can be modelled jointly by considering the redshift-space radial coordinate (cf. eq. 1.27)

$$s = \check{\chi}(z_{\text{obs}}) = \check{r} + \gamma u, \quad (5.6)$$

where we recall that a breve $\check{\cdot}$ denotes a fiducial comoving distance–redshift relation, $\check{r} = \check{\chi}(z)$, and the function

$$\gamma(z) = \frac{d\check{\chi}}{d\chi} = H \frac{d\check{\chi}}{dz} \quad (5.7)$$

measures the rescaling of the differential comoving distance compared to the true cosmology [244].

Similar to the derivation of the Kaiser RSD model in section 1.2, by using the number conservation law $n(\mathbf{r}, z) d^3\mathbf{r} = n(\mathbf{s}, z) d^3\mathbf{s}$ and equations (1.23)–(1.25), the SFB coefficient of the redshift-space number density $n(\mathbf{s}, z)$ can be written as

$$\begin{aligned} n_\mu &= \int d^3\mathbf{s} w(s) j_\mu(s) Y_\mu^*(\hat{\mathbf{s}}) n(\mathbf{s}, z) \\ &= \int d^3\mathbf{r} w(s) j_\mu(s) Y_\mu^*(\hat{\mathbf{r}}) M(\hat{\mathbf{r}}) \phi(r) \bar{n}(z) [1 + b_k(z) D_1(z) \delta_{\text{m},0}(\mathbf{r})], \end{aligned} \quad (5.8)$$

where we recall that a subscript ‘0’ denotes the value of a quantity at the current redshift $z = 0$. Here we assume the survey window function, $W = M\phi$, to be separable into the angular mask M and the radial selection function ϕ , and that the weighting scheme w is radial. If the angular mask M is simply a veto mask, then

$$f_{\text{sky}} = \int \frac{d^2\hat{\mathbf{r}}}{4\pi} M(\hat{\mathbf{r}}) \quad (5.9)$$

is the fraction of the sky observed, but more generally M can include completeness variations and angular weights, and f_{sky} becomes an effective sky fraction. The radial selection function ϕ may be normalised to the total number of galaxies in the survey volume,

$$N_g = 4\pi f_{\text{sky}} \int r^2 dr \phi(r) \bar{n}(z), \quad (5.10)$$

so that it is dimensionless. The subscript ‘ k ’ in the large-scale bias parameter b_k indicates that it may include scale-dependent modifications due to, for example, local PNG (see eq. 1.60). For a survey with maximum redshift z_{max} , the observed number density field vanishes outside $R = \check{\chi}(z_{\text{max}})$ and thus the wavenumbers are discretised. Using equation (5.6), one can expand

$$w(s) \simeq w(\check{r}) + \gamma u w'(\check{r}) \quad \text{and} \quad j_\mu(s) \simeq j_\mu(\check{r}) + \gamma u j'_\mu(\check{r}) \quad (5.11)$$

to first-order derivative terms, which are complete in linear perturbation theory as the expansion parameter is proportional to u . By considering the SFB expansion

$$\delta_{\mathbf{m},0}(\mathbf{r}) = \sum_{\mu} \kappa_{\mu} j_{\mu}(r) Y_{\mu}(\hat{\mathbf{r}}) (\delta_{\mathbf{m},0})_{\mu} \quad (5.12)$$

as well as the linearised continuity equation (B.21) written as

$$\mathbf{u} = -f \hat{\mathbf{r}} \cdot \nabla \nabla^{-2} \delta_{\mathbf{m}}, \quad (5.13)$$

one can recast

$$\begin{aligned} n_{\mu} = & \int d^3 \mathbf{r} w(\check{r}) j_{\mu}(\check{r}) Y_{\mu}^*(\hat{\mathbf{r}}) M(\hat{\mathbf{r}}) \phi(r) \bar{n}(z) \\ & + \int d^3 \mathbf{r} w(\check{r}) j_{\mu}(\check{r}) Y_{\mu}^*(\hat{\mathbf{r}}) M(\hat{\mathbf{r}}) \phi(r) \bar{n}(z) D_1(z) \sum_{\nu} \kappa_{\nu} j_{\nu}(r) Y_{\nu}(\hat{\mathbf{r}}) b_{k_{\nu}}(z) (\delta_{\mathbf{m},0})_{\nu} \\ & + \int d^3 \mathbf{r} (w j_{\mu})'(\check{r}) Y_{\mu}^*(\hat{\mathbf{r}}) M(\hat{\mathbf{r}}) \phi(r) \bar{n}(z) \gamma(z) \\ & \times f(z) D_1(z) \sum_{\nu} \frac{\kappa_{\nu}}{k_{\nu}^2} j'_{\nu}(r) Y_{\nu}(\hat{\mathbf{r}}) (\delta_{\mathbf{m},0})_{\nu}. \end{aligned} \quad (5.14)$$

This is a sum of three contributions, namely the background piece, the real-space perturbation piece and the RSD piece. By introducing the angular, radial and RSD coupling coefficients [49, 228, 242],

$$M_{\mu\nu} = \int d^2 \hat{\mathbf{r}} Y_{\mu}^*(\hat{\mathbf{r}}) M(\hat{\mathbf{r}}) Y_{\nu}(\hat{\mathbf{r}}), \quad (5.15a)$$

$$\Phi_{\mu\nu} = \kappa_{\nu} \int r^2 dr w(\check{r}) j_{\mu}(\check{r}) j_{\nu}(r) \phi(r) \bar{n}(z) \frac{b_{k_{\nu}}(z)}{b_{k_{\nu}}(0)} D_1(z), \quad (5.15b)$$

$$\Upsilon_{\mu\nu} = \frac{\kappa_{\nu}}{k_{\nu}^2} \int r^2 dr (w j_{\mu})'(\check{r}) j'_{\nu}(r) \phi(r) \bar{n}(z) \gamma(z) \frac{f(z)}{f(0)} D_1(z), \quad (5.15c)$$

one obtains a more compact expression,

$$n_{\mu} = \bar{n}_{\mu} + \sum_{\nu} M_{\mu\nu} \Phi_{\mu\nu} (b_0)_{\nu} (\delta_{\mathbf{m},0})_{\nu} + \sum_{\nu} M_{\mu\nu} \Upsilon_{\mu\nu} f_0 (\delta_{\mathbf{m},0})_{\nu}, \quad (5.16)$$

where $(b_0)_{\mu} \equiv b_{k_{\mu}}(0)$, $f_0 \equiv f(0)$ and

$$\bar{n}_{\mu} = \int d^3 \mathbf{r} w(\check{r}) j_{\mu}(\check{r}) Y_{\mu}^*(\hat{\mathbf{r}}) M(\hat{\mathbf{r}}) \phi(r) \bar{n}(z). \quad (5.17)$$

This shows that redshift-space galaxy clustering statistics in linear perturbation theory can be decomposed into spherical Fourier modes,

$$D_{\mu} = n_{\mu} - \bar{n}_{\mu} = \sum_{\nu} M_{\mu\nu} [(b_0)_{\nu} \Phi_{\mu\nu} + f_0 \Upsilon_{\mu\nu}] (\delta_{\mathbf{m},0})_{\nu}, \quad (5.18)$$

which satisfy $\langle D_{\mu} \rangle = 0$ with the two-point correlator

$$\langle D_{\mu} D_{\nu}^* \rangle = \sum_{\lambda} M_{\mu\lambda} M_{\nu\lambda}^* [(b_0)_{\lambda} \Phi_{\mu\lambda} + f_0 \Upsilon_{\mu\lambda}] [(b_0)_{\lambda} \Phi_{\nu\lambda} + f_0 \Upsilon_{\nu\lambda}] \kappa_{\lambda}^{-1} P_{\mathbf{m},0}(k_{\lambda}) + \langle \epsilon_{\mu} \epsilon_{\nu}^* \rangle, \quad (5.19)$$

where the second term

$$\langle \epsilon_\mu \epsilon_\nu^* \rangle = M_{\mu\nu} \int r^2 dr j_\mu(\check{r}) j_\nu(\check{r}) w(\check{r})^2 \phi(r) \bar{n}(z) \quad (5.20)$$

accounts for the shot noise contribution [49, 228, 244]. It is worth commenting here that the SFB clustering mode D_μ is constructed from fluctuations in the galaxy number density field directly, and it is not necessarily equivalent to the SFB coefficient of the observed over-density field δ , because the latter is defined as a fraction of the background number density which may have been spatially modulated. For the underlying matter density contrast δ_m at a fixed epoch z in real space, however, this is not an issue since it is statistically homogeneous.

5.1.3 Comparisons with the Cartesian power spectrum analysis

To motivate the hybrid-basis approach to galaxy clustering analysis in the next section, it is worth making connections as well as comparisons between the spherical and Cartesian clustering modes D_μ and $\delta(\mathbf{k})$ as well as the corresponding Fourier analyses.

First of all, both D_μ and $\delta(\mathbf{k})$ depend on galaxy bias with respect to the underlying matter distribution and capture anisotropic clustering due to the RSD and AP effects on linear scales. However, no geometric approximations have been made in obtaining D_μ , and the SFB coupling coefficients $Y_{\mu\nu}$ explicitly mix clustering modes at different wavenumbers as a reflection of RSDs; in contrast, $\delta(\mathbf{k})$ obtained in the distant-observer limit ignores the mode-coupling term in the Jacobian of the mapping from real space to redshift space (see eq. 1.31). That said, the preceding derivation of the anisotropic clustering model in the SFB basis is limited to linear perturbation theory, and the extension to non-linear clustering, though possible, is considerably more complex than the modelling of the Cartesian power spectrum.

Secondly, the clear distinction between angular and radial components in the spherical Fourier analysis offers a number of advantages:

- There is no ambiguity in the definition of the line of sight, which is free to vary across the entire sky, so wide-angle corrections are not needed (see §§1.2 and 2.2);
- Relativistic and light-cone effects, which affect galaxy clustering along the line of sight, can be more easily included (see also §4) [229]. Indeed, redshift evolution in the background number density, clustering amplitude, bias and growth rate is fully captured by the SFB couplings Φ and Y (eqs. 5.15b and 5.15c), where redshift dependence is integrated radially via the distance–redshift relation;
- It allows easier joint analyses with other probes such as the CMB, weak lensing and the ISW effect [e.g. 242, 254–256];
- The separation of angular and radial survey systematics allows individual clustering modes to be treated in isolation in analysis [78, 257].

Thirdly, the Cartesian power spectrum analysis is tomographic, i.e. it requires binning and averaging in redshift, with models of the power spectrum evaluated at some effective redshift in each bin (see §2.1). If the redshift bins are too narrow, large clustering modes along the line-of-sight can be missed and the level of shot noise is higher; too wide, then the effective redshift cannot capture any redshift evolution within the bin. Indeed, the spherical Fourier analysis has been shown to be more robust and optimal in this regard [e.g. 245, 257].

Lastly, the spherical Fourier analysis is based on individual clustering modes of the over-density field, whereas the Cartesian power spectrum analysis is based on the two-point correlator compressed from many clustering modes. This is an important distinction with several implications:

- In the case of the former, the distribution of SFB clustering modes D_μ is exactly Gaussian as long as cosmic fluctuations can be described by GRFs. Although an over-density field with non-zero PNG is not exactly Gaussian, the deviation is constrained to be small by current CMB measurements [88], and any signature of PNG will be reflected more in the amplitude of clustering statistics rather than the detailed probability distribution. Therefore, the GRF assumption serves as a useful null hypothesis for the detection of any non-Gaussianity. In the spherical Fourier analysis, all the cosmological information is encoded in the covariance matrix which is the two-point correlator (eq. 5.19) and analytically tractable. However, without the benefit of FFT algorithms, the computational cost of a spherical Fourier analysis can be considerable, not least because angular integration over spherical harmonics and radial integration over spherical Bessel functions need to be performed repeatedly for all the SFB couplings in equation (5.15) for different cosmological models. The next section will provide a more detailed account of the computational complexity of the spherical Fourier analysis;
- In contrast, the distribution of measured power spectrum multipoles is only approximately Gaussian when the number of clustering modes is large enough for the central limit theorem to hold. Models of the power spectrum, as well as measurements compressed from individual Cartesian clustering modes, can be efficiently computed using the Hankel transform and FFTs. However, the covariance matrix is now a four-point function, which usually has to be estimated from ideally 10^3 – 10^4 realistic mock catalogues and poses a significant computational challenge (see §1.4).

Despite some of these apparent differences between the spherical Fourier and Cartesian power spectrum analyses, in the simplest scenarios there are straightforward connections between the clustering modes D_μ and $\delta(\mathbf{k})$ thanks to the orthogonality of the SFB and plane wave bases. If the angular mask $M(\hat{\mathbf{r}}) = 1$, then the angular coupling coefficients (eq. 5.15a) reduce to $M_{\mu\nu} = \delta_{\ell_\mu, \ell_\nu}^{(K)} \delta_{m_\mu, m_\nu}^{(K)}$; if in addition, there are no radial selection, weighting, AP effects or redshift evolution (i.e. a fixed redshift is considered), then $M_{\mu\nu} \Phi_{\mu\nu} = \delta_{\mu\nu}^{(K)}$. Therefore, in the absence of RSDs (i.e. $Y_{\mu\nu} = 0$), the SFB two-point correlator reduces from an infinite series (eq. 5.19) to being diagonal,

$$\frac{\kappa_{\mu\nu}}{\bar{n}^2} \langle D_\mu D_\nu^* \rangle = \delta_{\mu\nu}^{(K)} \left[b_{k_\mu}^2 P_m(k_\mu) + \frac{1}{\bar{n}} \right], \quad (5.21)$$

which is essentially the isotropic power spectrum plus the Poisson shot noise. This also hints at a convergence check for the SFB coupling integrals (eq. 5.15) and the SFB two-point correlator series (eq. 5.19). Finally, for the sake of completeness, it is worth commenting here that under less restrictive assumptions than the ones above, the SFB two-point correlator can also be related to the tomographic angular power spectrum C_ℓ , but the reader is referred to e.g. refs. [78, 229, 245] for more detail.

5.1.4 Computational complexity

To determine how many discrete SFB clustering modes up to a given maximum wavenumber k_{\max} are present in a survey volume with boundary radius R , one needs to know the number of positive zeros $u_{\ell n}$ of the spherical Bessel function j_ℓ that satisfy $u_{\ell n} \leq k_{\max}R$. Using the asymptotic expansion $u_{\ell 1} \sim \ell$ as $\ell \rightarrow \infty$, the maximum spherical degree is estimated to be

$$\ell_{\max} \simeq k_{\max}R. \quad (5.22)$$

By considering another asymptotic expansion $u_{\ell n} \sim (n + \ell/2 - 1/4)\pi$ for a fixed spherical degree ℓ as $n \rightarrow \infty$ [258], one can estimate the maximum spherical depth by

$$n_{\max, \ell} \simeq \frac{k_{\max}R}{\pi} - \frac{\ell}{2} + \frac{1}{4} \leq \frac{k_{\max}R}{\pi} + \frac{1}{4}. \quad (5.23)$$

Since for each spherical degree there are $(2\ell + 1)$ equivalent spherical orders $m = -\ell, \dots, \ell$, the total number of SFB clustering modes is bounded above by

$$N_{\text{mode}} = \sum_{\ell=0}^{\ell_{\max}} (2\ell + 1)n_{\max, \ell} \leq \left(\frac{k_{\max}R}{\pi} + \frac{1}{4} \right) (k_{\max}R + 1)^2 \sim \frac{(k_{\max}R)^3}{\pi}, \quad (5.24)$$

but this bound is in general far from being saturated.

For SFB transforms of the survey and synthetic catalogues, if one constructs the clustering modes by direct summation, the number of computations is simply $(1 + \alpha^{-1})N_g N_{\text{mode}}$ where each unit of computation is an evaluation of the spherical Bessel and harmonic functions. The calculation of the SFB couplings (eq. 5.15) is more laborious: the number of angular coupling coefficients is

$$\left[\sum_{\ell=0}^{\ell_{\max}} (2\ell + 1) \right]^2 \sim (k_{\max}R)^4, \quad (5.25)$$

but this can be reduced by employing symmetry relations between the spherical harmonics; on the other hand, for both radial and RSD couplings, the number of coupling coefficients is

$$\left(\sum_{\ell=0}^{\ell_{\max}} n_{\max, \ell} \right)^2 \leq \left[(\ell_{\max} + 1) \left(\frac{k_{\max}R}{\pi} + \frac{1}{4} \right) \right]^2 \sim \frac{(k_{\max}R)^4}{\pi^2}, \quad (5.26)$$

and the number of shot noise integrals (eq. 5.20) is similar. Finally, the infinite series (eq. 5.19) requires at least N_{mode} terms for convergence, each of which is itself a product of the SFB coupling coefficients and the linear matter power spectrum.

Although the computational cost of angular coupling coefficients $M_{\mu\nu}$ seems to be the highest, like shot noise they are independent of the cosmological model and thus need to be calculated only once for a given survey. In contrast, the radial and RSD couplings can change with the cosmological model if redshift evolution is to be taken into account, so their evaluations are likely to be the most expensive steps in a full analysis.

5.2 Hybrid-basis likelihood inference

Having laid out the different aspects of spherical and Cartesian Fourier analyses based on the SFB and plane wave bases, we propose a hybrid-basis approach to cosmological parameter inference from galaxy clustering measurements:

- Since the survey geometry and other observational systematics have the biggest impact on the relatively few Fourier modes on the largest scales, one could use the spherical Fourier analysis to faithfully capture the physics of anisotropic galaxy clustering in linear perturbation theory, and construct from the SFB clustering modes D_μ the cosmological likelihood directly, which is exactly multivariate normal provided cosmic fluctuations are well described by GRFs. In this case, the covariance matrix is the two-point correlator, which contains all the cosmological information and can be computed analytically;
- On comparatively smaller scales, one could use the Yamamoto estimator \hat{P}_ℓ for power spectrum multipoles as the summary statistics (see §2.2), which can be efficiently computed using FFTs. Since \hat{P}_ℓ is compressed from a large number of clustering modes on these scales, its probability distribution is very close to being Gaussian by the central limit theorem;
- By combining the probability distributions of D_μ and \hat{P}_ℓ , one can then obtain a hybrid-basis likelihood for cosmological parameters.

This idea of adopting different statistics, either uncompressed or compressed from individual modes of fluctuations depending on the physical scale considered, is inspired by the use of hybrid estimators in CMB studies [249, 250]: one approach is to evaluate the likelihood function from the CMB map pixels directly, either searching for a quadratic or maximum likelihood estimator of the angular power spectrum [185, 246, 247] or Monte Carlo sampling the posterior surface [259–261]; another is to compress the map pixels into pseudo- C_ℓ estimators of the angular power spectra based on which an approximate likelihood can be constructed [177, 248]. This strategy has then been successfully applied by *Planck* to its cosmological likelihoods, which consist of a low- ℓ part ($\ell \leq 29$) based on the CMB temperature and polarisation map pixels, and a high- ℓ part ($\ell \geq 30$) based on the pseudo- C_ℓ estimator [251–253]. As far as we are aware, this approach has not been applied in any LSS settings before, which are arguably more nuanced as LSS data sets are intrinsically three-dimensional. This section will set out the basic steps involved in constructing the likelihood functions $\mathcal{L}(\theta)$ for cosmological parameters θ from D_μ and \hat{P}_ℓ , which are split at the *hybridisation wavenumber* k_{hyb} analogous to the ℓ split in *Planck* likelihoods, i.e. the SFB wavenumbers are restricted to $k_\mu \leq k_{\text{hyb}}$ and the power spectrum wavenumbers to $k_{\text{hyb}} < k < k_{\text{max}}$, where k_{max} is the overall maximum wavenumber in the analysis.

Spherical-basis likelihood. The data vector of SFB clustering modes, $\mathbf{D} = (D_\mu)$, is calculated from the survey and synthetic catalogues by direct summation over weighted Dirac delta contributions from each galaxy (cf. eq. 2.3),

$$D_\mu = \sum_{i=1}^{N_g} w(s_i) j_\mu(s_i) Y_\mu^*(\hat{\mathbf{s}}_i) - \alpha \sum_{i=1}^{N_s} w(s_i) j_\mu(s_i) Y_\mu^*(\hat{\mathbf{s}}_i). \quad (5.27)$$

Here the vector index μ can be ordered either ‘naturally’ by the tuple (ℓ_μ, m_μ, n_μ) , or ‘spectrally’ by the wavenumber k_μ and the spherical order m_μ . Under the GRF assumption for large-scale fluctuations in the galaxy distribution, and since the SFB transform is linear in the field, the spherical-basis data vector follows the *circularly-symmetric complex normal distribution* [262, 263], $\mathbf{D} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{S})$, where all cosmological parameter dependence is in the covari-

ance matrix $\mathbf{S}(\theta) = \text{cov}[\mathbf{D}]$ whose entries $(\mathbf{S})_{\mu\nu} = \langle D_\mu D_\nu^* \rangle$ are precisely the SFB two-point correlator (eq. 5.19).¹ Therefore the spherical-basis likelihood is given by

$$\mathcal{L}_{\text{sph}}(\theta; \mathbf{D}) = \frac{\exp[-\mathbf{D}^\dagger \mathbf{S}(\theta)^{-1} \mathbf{D}]}{|\pi \mathbf{S}(\theta)|}. \quad (5.28)$$

though there are a few technicalities in its practical evaluation:

- Based on the symmetry of the spherical harmonics, $Y_{\ell-m}^*(\hat{\mathbf{s}}) = (-1)^m Y_{\ell m}(\hat{\mathbf{s}})$, almost half of the SFB clustering modes D_μ can be calculated simply using $D_{\ell-mn} = (-1)^m D_{\ell mn}^*$;
- In order to evaluate \mathbf{S} accurately, the infinite series (eq. 5.19) must numerically converge. This may require additional SFB clustering modes with wavenumbers $k_\lambda > k_{\text{hyb}}$ to be included in the sum, and the appropriate truncation point in the series may need to be determined empirically, e.g. using the diagonal SFB two-point correlator $\langle |D_\mu|^2 \rangle$ (see eq. 5.21) as a diagnostic quantity;
- Since the data vector $\mathbf{D} \in \mathbb{C}^{N_{\text{mode}}}$ consists of N_{mode} uncompressed SFB clustering modes, dimensions of the covariance matrix $\mathbf{S} \in \mathbb{C}^{N_{\text{mode}} \times N_{\text{mode}}}$ can be prohibitively large that the inversion of \mathbf{S} becomes numerically unstable when it is not diagonal, i.e. when the SFB clustering modes are correlated. In this case, some eigenvalues of \mathbf{S} can be very close to zero, and with imperfect numerical precision the inverted matrix \mathbf{S}^{-1} may acquire large negative eigenvalues, posing a significant challenge to the sampling of the posterior distribution from \mathcal{L}_{sph} . A possible remedy to this problem is to apply a compression matrix $\mathbf{R} \in \mathbb{C}^{N'_{\text{mode}} \times N_{\text{mode}}}$ to both the data vector and the covariance matrix before evaluating \mathcal{L}_{sph} , i.e. by replacing $\mathbf{D} \mapsto \mathbf{R}\mathbf{D}$ and $\mathbf{S} \mapsto \mathbf{R}\mathbf{S}\mathbf{R}^\top$, where $N'_{\text{mode}} < N_{\text{mode}}$ and $\mathbf{R}\mathbf{R}^\top = \mathbf{I}$ is the identity matrix. One such compression method to ensure numerical stability is introduced below.

Cartesian-basis likelihood. The weighted field $F(\mathbf{s})$ is first computed from the survey and synthetic catalogues and Fourier transformed (see § 2.1), from which one can then form the data vector of power spectrum multipoles, $\hat{\mathbf{P}} = (\hat{P}_\ell(k_j))$, whose components are ordered by the multipole degree ℓ and then the binned wavenumber k_j . Being the two-point correlator of a GRF, the data vector $\hat{\mathbf{P}}$ should follow the hypo-exponential distribution (see § 3.1); however, by the central limit theorem, when the number of clustering modes contributing to $\hat{P}_\ell(k_j)$ is large, one may assume the multivariate normal distribution, $\hat{\mathbf{P}} \sim \mathbf{N}(\bar{\mathbf{P}}, \boldsymbol{\Sigma})$. The Cartesian-basis power spectrum likelihood is thus given by

$$\mathcal{L}_{\text{Cart}}(\theta; \hat{\mathbf{P}}) = |2\pi \boldsymbol{\Sigma}(\theta)|^{-1/2} \exp\left\{-\frac{1}{2} \left[\hat{\mathbf{P}} - \bar{\mathbf{P}}(\theta)\right]^\top \boldsymbol{\Sigma}(\theta)^{-1} \left[\hat{\mathbf{P}} - \bar{\mathbf{P}}(\theta)\right]\right\}. \quad (5.29)$$

Here $\boldsymbol{\Sigma}(\theta) = \text{cov}[\hat{\mathbf{P}}]$ is the covariance matrix, and $\bar{\mathbf{P}}(\theta) = \mathbb{E}[\hat{\mathbf{P}}]$ is the expectation of the power spectrum multipole estimator with components

$$\bar{P}_\ell(k_j) = \tilde{P}_\ell(k_j) + P_{\ell, \text{shot}}, \quad (5.30)$$

¹ In contrast to previous works [e.g. 49], we do not separate the spherical-basis data vector into real and imaginary parts which jointly follow the multivariate normal distribution.

where $\tilde{P}_\ell(k_j)$ is the window-convolved model of power spectrum multipoles and $P_{\ell,\text{shot}}$ is the shot noise contribution (see §2),

$$P_{\ell,\text{shot}} = \frac{1 + \alpha}{I} \int d^3\mathbf{s} \mathcal{L}_\ell(\hat{\mathbf{k}} \cdot \hat{\mathbf{s}}) w(s)^2 \bar{n}(s). \quad (5.31)$$

Similar to the spherical-basis likelihood, there are two technicalities related to the covariance matrix in evaluating the Cartesian-basis likelihood $\mathcal{L}_{\text{Cart}}$:

- The true covariance matrix Σ is usually analytically intractable, so it has to be replaced by an estimate $\hat{\Sigma}$ from mock catalogues. SH has shown that the appropriate distribution to use as the likelihood function is no longer multivariate normal but a modified Student's t -distribution (see §1.4). However, when the number of mock catalogues used to obtain the estimate $\hat{\Sigma}$ far exceeds the dimension of the data vector \hat{P} , the multivariate normal distribution remains an excellent approximation;
- Because of the high computational cost associated with generating a large number of mock catalogues, the covariance matrix estimate $\hat{\Sigma}$ is usually produced at fixed fiducial cosmological parameters $\check{\theta}$. To account for any parameter dependence, one could use the variance–correlation decomposition which allows a parameter-dependent estimate $\hat{\Sigma}(\theta)$ to be obtained from the fiducial estimate $\hat{\Sigma}_f$ by a rescaling (see §3),

$$\hat{\Sigma}(\theta) = \Lambda(\theta) \check{\Lambda}^{-1} \hat{\Sigma}_f \check{\Lambda}^{-1} \Lambda(\theta). \quad (5.32)$$

Here $\Lambda(\theta) = \text{diag}(\bar{P}_\ell(k_j))$ is a diagonal matrix with entries given by the window-convolved power spectrum multipole model, including the shot noise contribution, at cosmological parameters θ , and $\check{\Lambda}$ is the same diagonal matrix but evaluated at fiducial parameters $\check{\theta}$.

Hybrid-basis likelihood. In the idealised scenario where the data vectors D and \hat{P} are independent, the hybrid-basis likelihood is simply the product of the two above,

$$\mathcal{L}_{\text{hyb}}(\theta; D, \hat{P}) = \mathcal{L}_{\text{sph}}(\theta; D) \mathcal{L}_{\text{Cart}}(\theta; \hat{P}). \quad (5.33)$$

Unfortunately, when clustering modes of different wavenumbers k are mixed in the presence of survey window and RSD effects, this does not necessarily hold. In the *Planck* likelihood analyses, the correlation between low- ℓ and high- ℓ components poses a similar issue, and different hybridisation schemes are explored [251–253]; however, their results are not particularly sensitive to the hybridisation scheme and thus a sharp transition between low- ℓ and high- ℓ components can be adopted without accounting for their correlation. In this work, we make a similar assumption that the low- k spherical-basis and high- k Cartesian-basis likelihoods can be directly combined—this is justified if their correlation is weak when the mixing kernel is sufficiently narrow in k -space and if the joint probability distribution of D and \hat{P} is multivariate normal.² In practice, the correlation between the spherical-basis and Cartesian power spectrum data can be estimated from mock catalogues alongside the covariance matrix estimate for power spectrum multipoles, and one could attempt to de-correlate the combined data vector or reweight different data components before combining them, e.g. with a Bayesian hyperparametric method [264–266].

² Note that zero correlation does not necessarily imply independence between two multivariate normal random variables unless their joint probability distribution is also multivariate normal.

Spherical-basis data compression. As mentioned above, data compression may be necessary for the spherical Fourier analysis not only because of the high computational cost and but also for numerical stability. As such, previous applications to galaxy survey catalogues [e.g. 236, 239] had to adopt data compression techniques such as the Karhunen–Loève transform, which constructs optimal linear combinations of SFB clustering modes [267], to make the analysis feasible. It has been noted in earlier discussion that when dimensions of a matrix $\mathbf{S} \in \mathbb{C}^{N_{\text{mode}} \times N_{\text{mode}}}$ are large, matrix inversion may not be numerically stable—since one cannot determine the elements of \mathbf{S} arbitrarily precisely, small perturbations to \mathbf{S} can lead to spurious results for \mathbf{S}^{-1} . If \mathbf{S} is a covariance matrix, which must be positive-definite, the inversion procedure can introduce negative eigenvalues in \mathbf{S}^{-1} unless it happens to be diagonal. This can affect both likelihood calculations as well as Fisher forecasts for cosmological parameters, as noted in refs. [236, 239, 268]. At the root of this phenomenon is the *condition number* $\kappa(\mathbf{S})$, which gauges the sensitivity of precision in numerically inverted \mathbf{S}^{-1} to perturbations in the elements of \mathbf{S} ; for a covariance matrix, it is given by

$$\kappa(\mathbf{S}) = \frac{\lambda_1}{\lambda_{N_{\text{mode}}}}, \quad (5.34)$$

where $\lambda_1 > \dots > \lambda_{N_{\text{mode}}} > 0$ are its eigenvalues arranged in descending order [269]. For the covariance matrix \mathbf{S} of the SFB clustering modes, its eigenvalues are typically of similar orders of magnitude when \mathbf{S} is close to being diagonal, as they are related to the power spectrum (see eq. 5.21). However, when the SFB clustering modes become correlated because of the SFB coupling coefficients, \mathbf{S} is not diagonal and some eigenvalues are repelled towards zero, which can become unstable upon matrix inversion. This hints at a solution based on the *principal-component analysis*, where combinations of SFB modes corresponding to smaller eigenvalues are considered to have less cosmological information and thus discarded. A practical method for data compression is proposed as follows. One first considers the covariance matrix $\mathbf{S}(\check{\theta})$ evaluated at the fiducial cosmological parameters $\check{\theta}$, with its eigenvalue–eigenvector pairs $(\lambda_\beta, \mathbf{e}_\beta)$ arranged in descending order by eigenvalue. Next, one sets an acceptable condition number κ , and find the smallest eigenvalue $\lambda_\beta = \lambda_B$ such that $\lambda_1/\lambda_\beta \leq \kappa$. One can then define the fixed compression matrix

$$\mathbf{R} = (\mathbf{e}_1^\top, \dots, \mathbf{e}_B^\top) \in \mathbb{C}^{B \times N_{\text{mode}}}. \quad (5.35)$$

This satisfies the orthonormality condition $\mathbf{R}\mathbf{R}^\top = \mathbf{I}$, and can be used to replace

$$\mathbf{D} \mapsto \mathbf{R}\mathbf{D} \quad \text{and} \quad \mathbf{S}(\theta) \mapsto \mathbf{R}\mathbf{S}(\theta)\mathbf{R}^\top \quad (5.36)$$

in the spherical-basis likelihood (eq. 5.28). It is worth noting that the compression method proposed here, being essentially a principal-component analysis, is conceptually related to the Karhunen–Loève transform; however, the Karhunen–Loève transform used in e.g. ref. [236] considers the signal-to-noise eigenmodes of the Fisher information matrix for specific cosmological parameters, whereas here we are merely considering the eigenvalues of the covariance matrix indifferent to any particular parameters being estimated, as our main motivation for compression is to ensure numerical stability.

5.3 Application to N -body simulations

As an initial application of the hybrid-basis methodology, we focus on the real-space clustering of haloes in mock catalogues generated from N -body simulations with both Gaussian and non-Gaussian initial conditions, and attempt to recover the local PNG parameter f_{NL} , which imprints a scale-dependent signature in the tracer bias (see § 1.3), from large-scale clustering measurements. The next subsection describes properties of the mock catalogue and the intermediary quantities needed for likelihood evaluations, such as the window function, covariance matrix estimates and the SFB coupling coefficients; in the subsequent subsection, parameter constraints on f_{NL} and the scale-independent linear bias b_1 obtained from the hybrid-basis likelihood (eq. 5.33) and the Cartesian-basis power spectrum likelihood (eq. 5.29) are compared.

5.3.1 Mock catalogues, data products and model ingredients

Our halo mock catalogues are generated from a series of dark matter N -body simulations at a flat Λ CDM cosmology with $(h, \Omega_{m,0}, \Omega_{b,0}, \sigma_8) = (0.70, 0.27, 0.044, 0.80)$ —this is our fiducial cosmological model. The matter transfer function is first computed with the public code CAMB [270], which is then used to calculate initial conditions with the second-order Lagrangian perturbation theory (2LPT). The simulations are then seeded using the public code 2LPTIC which can generate initial conditions with the inclusion of non-zero local PNG [271, 272]. In total, 24 simulations with $f_{\text{NL}} = 0$ and 20 simulations with $f_{\text{NL}} = 100$ are produced, each in a $1 h^{-3} \text{Gpc}^3$ comoving box of 512^3 dark matter particles evolved from redshift $z = 32$ to $z = 1$ using the public code GADGET-2 [273]. We have not run the simulations down to redshift $z = 0$ because of computation time; indeed, future galaxy surveys probing f_{NL} on very large scales will mostly focus on the $z > 1$ Universe. Finally, dark matter haloes are identified within the mock catalogues using the public code AHF [274], with at least 36 particles per halo corresponding to a minimum halo mass of $M \approx 2.0 \times 10^{13} h^{-1} M_\odot$.

For this cosmological model, the scale-dependent modification received by the linear halo bias b_1 is

$$\Delta b(k, z) = 3f_{\text{NL}}[b_1(z) - 1] \frac{1.3\delta_c \Omega_{m,0} H_0^2}{c^2 k^2 T(k) D_1(z)}, \quad (5.37)$$

where the speed of light c is restored, the tracer-dependent parameter is set to $p = 1$ to match the mock catalogues, and the numerical factor arising from normalisation of the linear growth factor is $C(\Omega_{m,0}) \approx 1.3$ (cf. eq. 1.60).

To compare the hybrid-basis and Cartesian power spectrum analyses, two geometric set-ups are considered: the ‘full-sky’ set-up, in which only haloes within a sphere of comoving radius $R = 500 h^{-1} \text{Mpc}$ centred in the N -body simulations are used; the ‘partial-sky’ set-up, where the data sets of haloes are further restricted to the proportions covered by the footprint of the BOSS Data Release 12 Constant MASS (CMASS) North Galactic Cap (NGC) sample within the comoving radius range $100 h^{-1} \text{Mpc} \leq r \leq R$. The angular mask function M is constructed from the BOSS CMASS NGC random catalogue with HEALPIX pixelation $N_{\text{side}} = 32$ [275],³ as shown in Figure 5.1. This angular mask corresponds to a sky

³ The random catalogue is publicly available at data.sdss.org/sas/dr12/boos/lss/random0_DR12v5_CMASS_North.fits.gz.

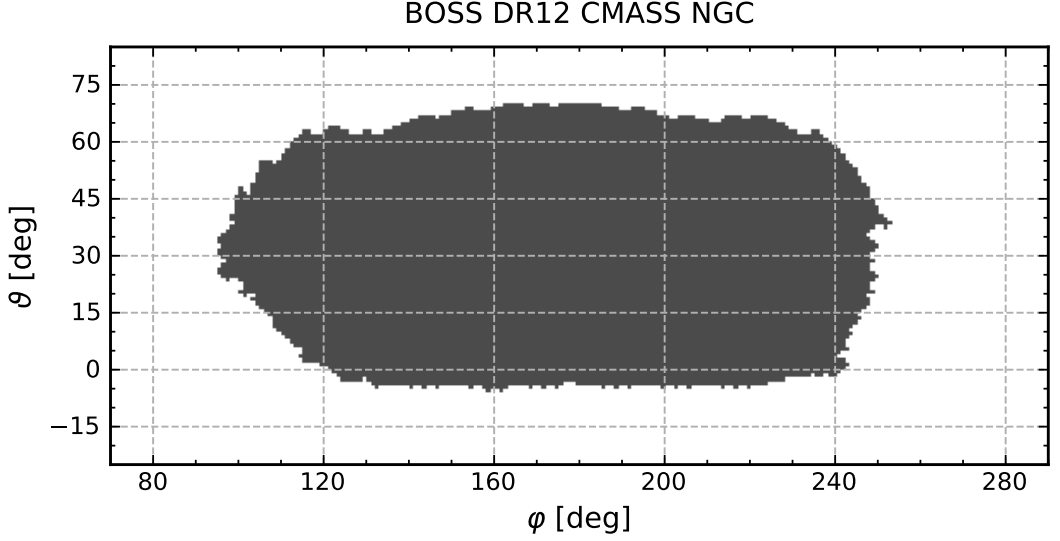


Figure 5.1. Survey angular mask $M(\hat{r})$ used in the partial-sky set-up. The angular mask function takes binary values (1 shown by the shaded region and 0 elsewhere) and is constructed from the BOSS Data Release 12 (DR12) CMASS NGC random catalogue with HEALPix pixelation $N_{\text{side}} = 32$. The vertical and horizon axes correspond to the polar and azimuthal angles (ϑ, φ) respectively in spherical coordinates. Figure taken from Fig. 1 in Paper III.

fraction of $f_{\text{sky}} \approx 0.2$ and is simply chosen to demonstrate the hybrid-basis approach with a realistic survey geometry.

For both full-sky and partial-sky set-ups, the hybrid-basis and Cartesian power spectrum analyses are performed with maximum wavenumber $k_{\text{max}} = 0.08 h \text{Mpc}^{-1}$, which is set to ensure that the halo bias remains scale-independent when $f_{\text{NL}} = 0$, since we have found the halo bias to be slightly scale-dependent even in the absence of PNG, and have had to set a relatively high minimum halo mass owing to the limited resolution of our simulations. As no anisotropy is expected from real-space halo clustering, only the power spectrum monopole is considered, which is estimated using the Yamamoto estimator (see §2.2) and binned in wavenumber with uniform width $\Delta k = 0.01 h \text{Mpc}^{-1}$. For the hybrid-basis analysis, the hybridisation wavenumber is pushed up to $k_{\text{hyb}} = 0.04 h \text{Mpc}^{-1}$ for which the computation times of spherical clustering statistics and the spherical-basis likelihood remain reasonable.

The intermediary data products and model ingredients, which are required for likelihood evaluations in addition to the hybrid-basis and Cartesian power spectrum data vectors \mathbf{D} and $\hat{\mathbf{P}}$, include the following:

- the SFB coupling coefficients $M_{\mu\nu}$ and $\Phi_{\mu\nu}$, which are numerically integrated from the angular mask M and the radial selection function ϕ (see eq. 5.15);
- the window function multipoles Q_ℓ used to convolve power spectrum models, which are determined from a synthetic random catalogue (see §2.3);
- the fiducial covariance matrix estimate $\hat{\Sigma}_f$ for the binned power spectrum monopole, which is obtained from a large number of synthetic random catalogues (see eq. 1.65);
- the compression matrix \mathbf{R} needed to reduce the dimensionality of \mathbf{D} and improve numerical stability in the partial-sky set-up, where the spherical-basis covariance matrix is non-diagonal.

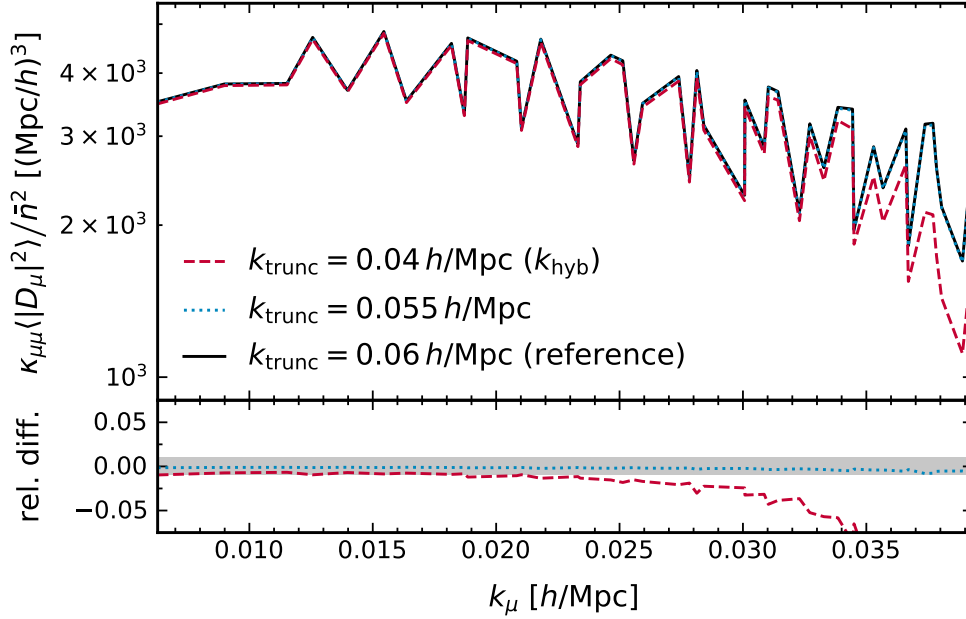


Figure 5.2. Convergence check of the normalised SFB autocorrelator, $\kappa_{\mu\mu} \langle |D_\mu|^2 \rangle / \bar{n}^2$, evaluated at each SFB wavenumber k_μ from the series (eq. 5.19) in the fiducial cosmological model with $f_{\text{NL}} = 0$ and $b_1 = 1$. The top panel shows partial sums of the series truncated at different wavenumbers $k_{\text{trunc}} = 0.04, 0.055, 0.06 h^{-1}$ Mpc, which respectively correspond to k_{hyb} (dashed red line), the actual wavenumber cut-off adopted in our analysis (dotted blue line) and the reference case (solid black line). The bottom panel shows the relative difference of partial sums compared to the reference case, with the shaded region marking deviations within $\pm 1\%$. Note that oscillations in the top panel are not due to numerical noise but reflect the behaviour of SFB clustering modes in the presence of window effects. Figure taken from Fig. 2 in Paper III.

First, while only SFB clustering modes with wavenumbers $k_\mu \leq k_{\text{hyb}}$ are included in the data vector \mathbf{D} , mode coupling in the partial-sky case means that more modes with wavenumbers $k_\mu > k_{\text{hyb}}$ must be included in the series (eq. 5.19) for convergence, and correspondingly additional angular and radial SFB coupling coefficients $M_{\mu\nu}$ and $\Phi_{\mu\nu}$ must be calculated. To check numerical convergence of the series, the normalised SFB autocorrelator $\kappa_{\mu\mu} \langle |D_\mu|^2 \rangle / \bar{n}^2$ is employed as a diagnostic quantity (see eq. 5.21). In Figure 5.2, partial sums of the series truncated at wavenumbers $k_{\text{trunc}} = 0.04, 0.055, 0.06 h^{-1}$ Mpc are compared, where the fiducial cosmological model is assumed with $f_{\text{NL}} = 0$ and $b_1 = 1$; it is worth commenting that oscillations of the SFB autocorrelator with the wavenumber are not due to numerical noise but are simply the behaviour of the SFB clustering modes in the presence of window effects. We have found that $k_{\text{trunc}} = 0.055 h^{-1}$ Mpc is sufficient to ensure numerical convergence at percent levels, which is then adopted as the series cut-off in our spherical Fourier analysis. In Figure 5.3, the dimensionless coupling coefficients $\text{Re } M_{\mu\nu}$ and $\Phi_{\mu\nu} / \bar{n}$ for wavenumbers $k_\mu \leq k_{\text{trunc}} = 0.055 h^{-1}$ Mpc are shown as matrices. The angular coupling coefficients $\text{Re } M_{\mu\nu}$ are arranged in the ‘natural’ order by the (ℓ, m) tuple and the radial coupling coefficients $\Phi_{\mu\nu}$ are arranged in the ‘spectral’ order by the wavenumber k_μ . The imaginary part $\text{Im } M_{\mu\nu}$ is not shown since it is close to zero for a binary-valued angular mask [49].

Secondly, the window function multipoles Q_ℓ are computed using FFTs from a synthetic

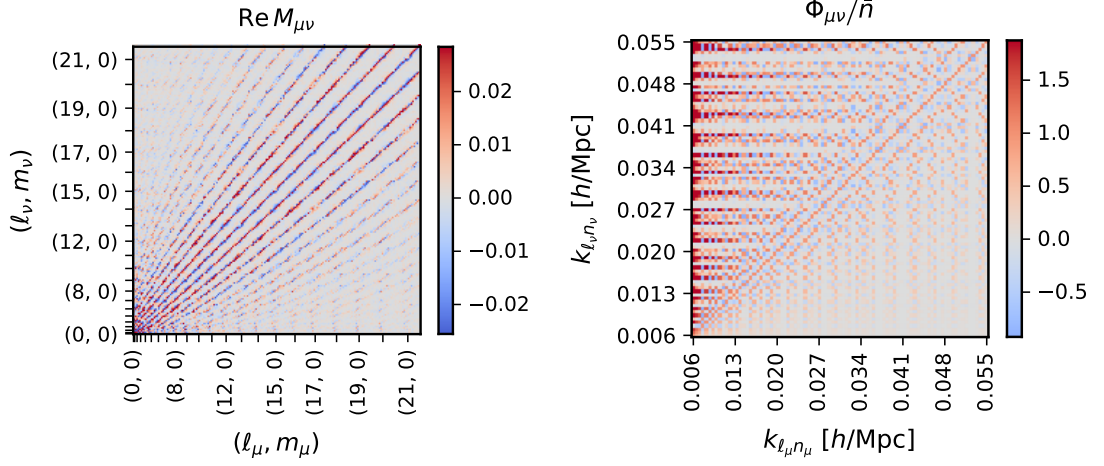


Figure 5.3. Dimensionless angular and radial SFB coupling coefficients $\text{Re } M_{\mu\nu}$ (left column) and $\Phi_{\mu\nu}/\bar{n}$ (right column) for wavenumbers $k_{\ell n} \leq k_{\text{trunc}} = 0.055 h^{-1} \text{ Mpc}$ in the partial-sky set-up. Angular coefficients $M_{\mu\nu}$ are ordered by spherical degree and order (ℓ, m) whereas radial coefficients $\Phi_{\mu\nu}$ are ordered by wavenumber $k_{\ell n}$. Figure taken from Fig. 3 in Paper III.

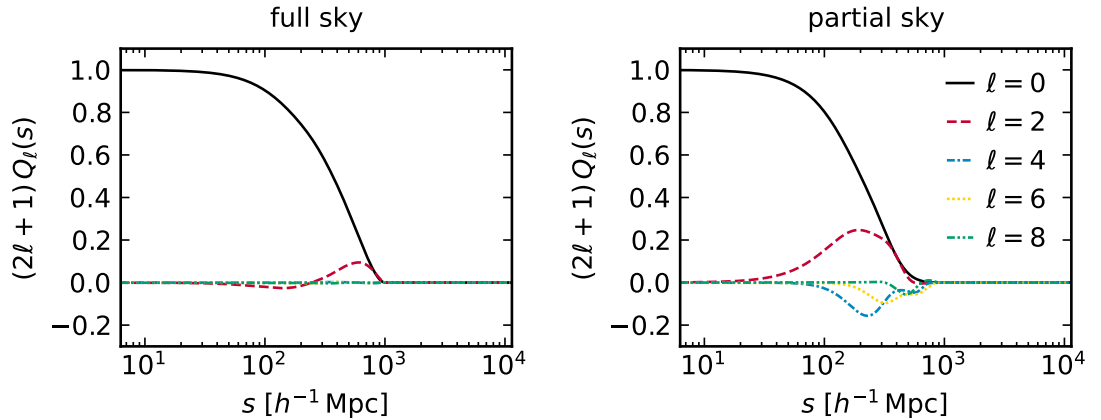


Figure 5.4. Window function multipoles Q_ℓ as functions of the comoving separation s , normalised to $Q_0(0) = 1$, in the full-sky (left column) and partial-sky (right column) set-ups. Each degree- ℓ multipole is multiplied by $(2\ell + 1)$ for visual differentiation. Figure adapted from Fig. 4 in Paper III.

random catalogue as described in section 2.3. The number density field of the synthetic catalogue is interpolated using the triangular-shaped cloud scheme (see e.g. ref. [131]) on a cubic grid with side length $L = 70 h^{-1} \text{ Mpc}$ and mesh number $N_{\text{grid}} = 768$. The large dimensions of the grid and the high mesh number allow one to compute the power spectrum across a wide range of scales without significant sample variance on very large scales or aliasing effects on very small scales. The power spectrum multipoles of the synthetic catalogue are then Hankel transformed to $Q_\ell(s)$. In Figure 5.4, the multipoles Q_ℓ in both full-sky and partial-sky set-ups are shown as functions of the comoving separation s ; in practice, only Q_0 is needed for our Cartesian power spectrum analysis since only the monopole is required for real-space clustering analysis.

Next, with the use of $N_{\text{rand}} = 2500$ synthetic random catalogues of 50 times the number

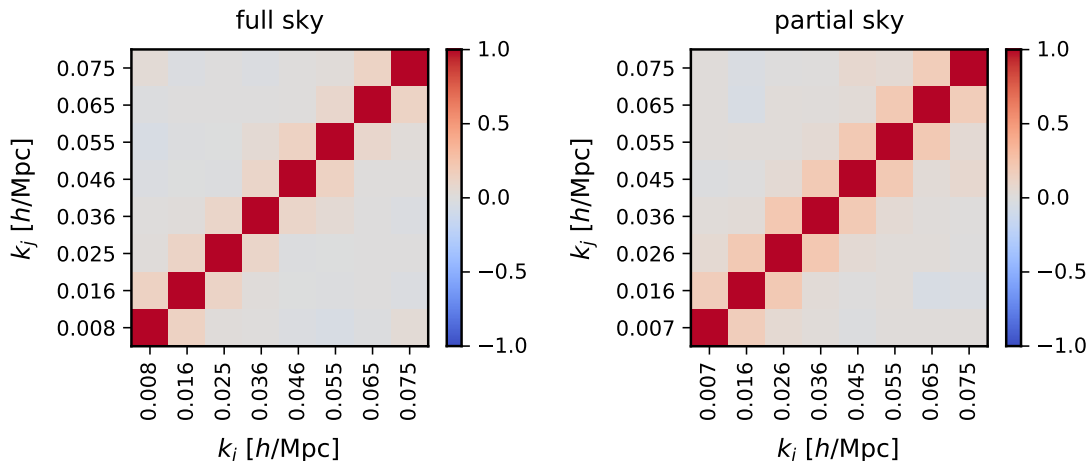


Figure 5.5. Estimated correlation matrices of the power spectrum monopole \hat{P}_0 in k -bins up to $k_{\max} = 0.08 h \text{Mpc}^{-1}$ in the full-sky (*left column*) and partial-sky (*right column*) set-ups. The wavenumber representing each bin is the average wavenumber of clustering modes in that bin. Figure taken from Fig. 5 in Paper III.

density of the halo mock catalogues (i.e. $\alpha = 0.02$), a fiducial covariance matrix estimate $\hat{\Sigma}_f$ is obtained for the binned power spectrum monopoles \hat{P}_0 . The use of unclustered random catalogues is justified as our analysis is restricted to linear scales, where the correlation between clustering modes is solely induced by the survey geometry rather than gravitational non-linearities. In Figure 5.5, the corresponding correlation matrices in the full-sky and partial-sky set-ups are shown, where the k -bins are represented by the average clustering mode wavenumber in each bin.

Finally, in our spherical Fourier analysis, the spherical-basis data \mathbf{D} consist of $N_{\text{mode}} = 456$ SFB clustering modes, so the dimensions of the covariance matrix $\mathbf{S} \in \mathbb{C}^{N_{\text{mode}} \times N_{\text{mode}}}$ are fairly large, which may render its inversion numerically unstable and result in divergence of the spherical-basis likelihood when evaluated along some particular direction in parameter space. In the full-sky set-up, \mathbf{S} is diagonal and well-conditioned; however, this is not the case for the partial-sky set-up, as SFB clustering modes become coupled. In the partial-sky set-up, the sky fraction is $f_{\text{sky}} \approx 0.2$, and together with the radial selection cut, the catalogue volume is about 18% of the full-sky comoving sphere, and hence the effective number of clustering modes is heuristically only $18\% \times N_{\text{mode}} \approx 82$. Therefore after the data compression procedure described in the previous section, we have decided to keep only 80 modes, corresponding to a conservative condition number $\kappa \approx 50$ for the compression matrix \mathbf{R} . Figure 5.6 shows the eigenvalue ratio λ_1/λ_β used as a proxy condition number, where $\{\lambda_\beta\}$ are the $N_{\text{mode}} = 456$ eigenvalues (arranged in descending order by magnitude) of the uncompressed spherical-basis covariance matrix $\check{\mathbf{S}}$ evaluated in the fiducial cosmology with $f_{\text{NL}} = 0$. It is evident that there are many positive eigenvalues λ_β which are orders of magnitude smaller than the largest eigenvalue λ_1 , and almost a third of all eigenvalues are negative. We also check whether the eigenvalue composition alters significantly if a different covariance matrix \mathbf{S} , e.g. evaluated at $f_{\text{NL}} = \pm 100$, is considered. The same figure shows that indeed the corresponding eigenvalues do not change very much except for the largest few and at the location where λ_β switches sign; the number of negative eigenvalues remains almost the same. This suggests that this data

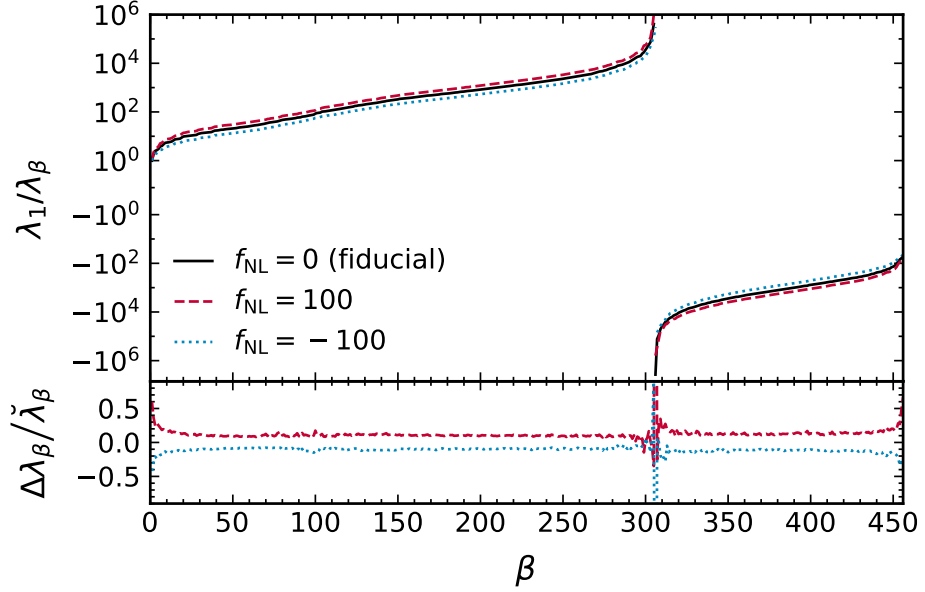


Figure 5.6. Proxy condition number λ_β/λ_1 of the spherical-basis covariance matrix with data compression, as a function of the index β for eigenvalues λ_β arranged in descending order by magnitude. In the top panel, λ_β/λ_1 is shown for eigenvalues of the fiducial covariance matrix $\check{\mathbf{S}}$ with $f_{\text{NL}} = 0$ (solid black line) as well as covariance matrices \mathbf{S} with $f_{\text{NL}} = \pm 100$ (dashed red line and dotted blue line respectively). In the bottom panel, the relative shift $\Delta\lambda_\beta$ in each eigenvalue compared to $\check{\lambda}_\beta$ of the fiducial covariance matrix $\check{\mathbf{S}}$ is shown. Note the sign change in the eigenvalue around $\beta = 305$. Figure adapted from Fig. 12 in Paper III.

compression method to stabilise the covariance matrix is robust.

5.3.2 Comparison of hybrid-basis and Cartesian power spectrum analyses

To justify the evaluation of the hybrid-basis likelihood as a product of the spherical-basis and Cartesian-basis likelihoods (eq. 5.33), it is necessary to check whether the correlation between low- k spherical-basis data \mathbf{D} and high- k Cartesian-basis data $\hat{\mathbf{P}}$ is sufficiently weak. To this end, the cross-correlation $\text{corr}[\mathbf{D}, \hat{\mathbf{P}}]$ is estimated from the aforementioned $N_{\text{rand}} = 2500$ synthetic random catalogues in both the full-sky and partial-sky set-ups, as shown in Figure 5.7. For SFB clustering modes $D_{\ell mn}$ of the same wavenumber $k_{\ell n}$ but different spherical orders m , the absolute cross-correlation value is averaged over these equivalent clustering modes. The cross-correlation indeed appears to be weak and there is no discernible evidence that particular SFB clustering modes are more strongly correlated with the power spectrum monopole in any particular wavenumber bin. For the full-sky case, the cross-correlation coefficient is consistently below 0.04 and, for the partial-sky case, below 0.06. Therefore we treat \mathbf{D} and $\hat{\mathbf{P}}$ as being effectively independent under the assumption that the joint distribution of $(\mathbf{D}, \hat{\mathbf{P}})$ is multivariate normal.

With all the data products and model ingredients described in the previous section, for each of the halo mock catalogues in the full-sky or partial-sky set-up we infer the local PNG parameter f_{NL} and the scale-independent linear bias b_1 jointly while keeping the fiducial cosmology fixed. Uniform priors for both parameters are chosen so that the posterior distribu-

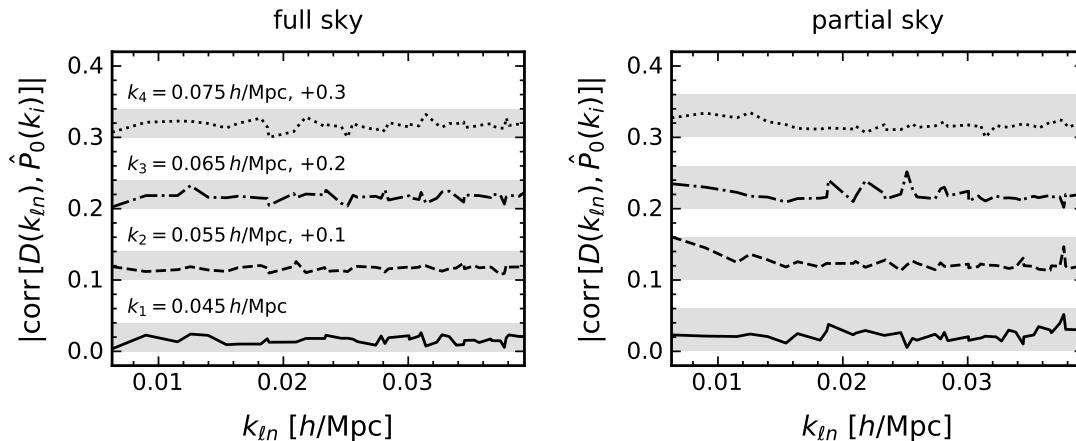


Figure 5.7. Estimates of the absolute value of cross-correlation coefficients $\text{corr}[D, \hat{P}]$ between the spherical-basis data D at wavenumbers $k_{\ell n} \leq k_{\text{hyb}}$ and the Cartesian-basis data \hat{P} (power spectrum monopole only) in wavenumber bins $k_i > k_{\text{hyb}}$ in the full-sky (left column) and partial-sky (right column) set-ups. The shaded grey stripes are of width 0.04 (left column) or 0.06 (right column) in cross-correlation value. Lines corresponding to the i th wavenumber bin are shifted up by $0.1(i - 1)$ in cross-correlation value for visual differentiation. Figure taken from Fig. 6 in Paper III.

tion is simply proportional to the likelihood function. As an example, Figure 5.8 shows the hybrid-basis posterior and its low- k and high- k components in the full-sky case. The results have been marginalised over 24 halo mock catalogues with $f_{\text{NL}} = 0$ by taking the average of the logarithmic posterior probabilities. The different orientations of the low- k and high- k posterior contours are mainly due to the different wavenumber ranges rather than differences in the spherical Fourier and Cartesian power spectrum analyses. In Figure 5.9, we compare the posterior distribution from the hybrid-basis likelihood with that from the Cartesian power spectrum likelihood for a few halo mock catalogues with $f_{\text{NL}} = 0$, showing individual results for each mock catalogue realisation without marginalisation or averaging. The constraints from the two likelihoods appear to be statistically consistent, though the posterior contours naturally scatter around $f_{\text{NL}} = 0$ owing to random data variation. In the rest of this section, measurements and parameter constraints will simply be presented as the marginalised results over different sets of mock catalogues rather than the combined results (which would have smaller uncertainties). This means that any data measurements presented hereafter have been averaged over equivalent mock catalogues, and so are the logarithmic posterior probabilities.

For the interested reader, here we provide some details of the actual computational cost involved in the spherical-basis likelihood analysis (see also §5.1.4). For the choice of the hybridisation wavenumber $k_{\text{hyb}} = 0.04 \text{ h Mpc}^{-1}$, the maximum spherical degree is $\ell_{\text{max}} = 15$. On a single processor core, the computation time of each SFB clustering mode D_μ by direct summation (eq. 5.27) is approximately 7×10^{-5} s per halo in the mock catalogues, or order of a day per mode for all galaxies in a DESI-like survey. The spherical coupling coefficients need to be computed once only for wavenumbers up to $k_{\text{trunc}} = 0.055 \text{ h Mpc}^{-1}$, since the background cosmology has been fixed at the fiducial model; this takes up to half an hour on 100 processor cores, and almost all the time is spent on the angular coupling coefficients $M_{\mu\nu}$ (eq. 5.15a) evaluated

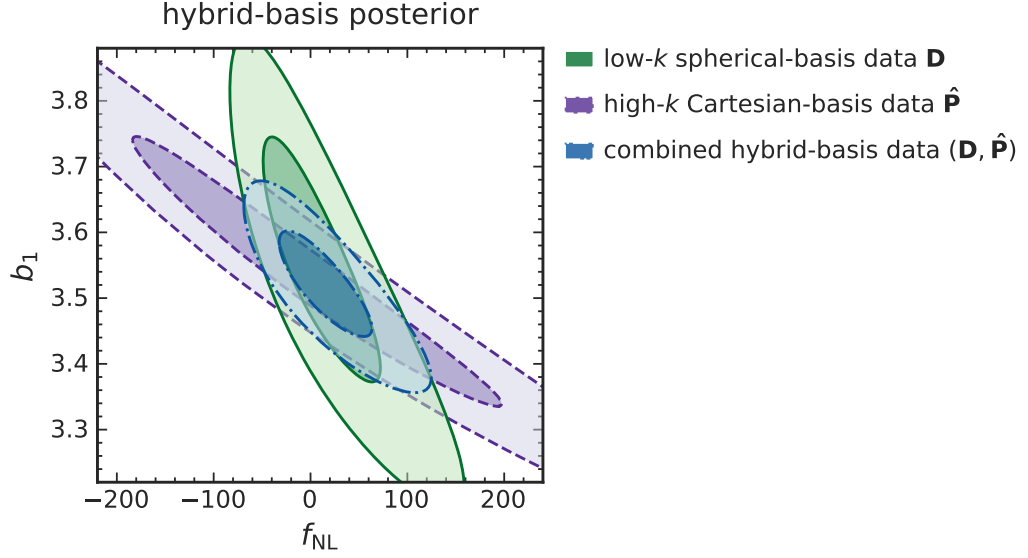


Figure 5.8. Joint posterior distribution of the local PNG parameter f_{NL} and the scale-independent linear bias b_1 from the hybrid-basis likelihood (eq. 5.33) (dash-dotted blue contours) as well as its low- k and high- k components based on the spherical-basis data \mathbf{D} (solid green contours) and the Cartesian-basis data $\hat{\mathbf{P}}$ (dashed purple contours). In this example, results have been marginalised over from 24 halo mock catalogues with $f_{\text{NL}} = 0$. The inner and outer contours in each set show the 1σ and 2σ credible regions. Figure taken from Fig. 7 in Paper III.

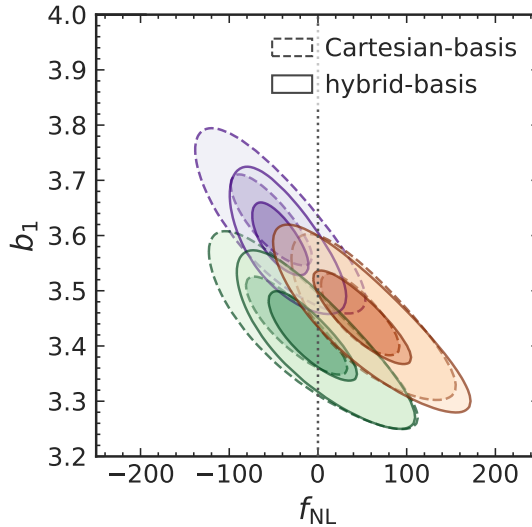


Figure 5.9. Joint posterior distribution of the local PNG parameter f_{NL} and the scale-independent linear bias b_1 from the hybrid-basis (solid contours) and Cartesian power spectrum (dashed contours) likelihoods from a subset of halo mock catalogues with $f_{\text{NL}} = 0$ in the full-sky set-up. The inner and outer contours in each coloured set show the 1σ and 2σ credible regions from a single halo mock catalogue.

Table 5.1. Posterior median estimates of f_{NL} and b_1 from the hybrid-basis and Cartesian power spectrum likelihood analyses, marginalised over halo mock catalogues with $f_{\text{NL}} = 0$ and $f_{\text{NL}} = 100$ in the full-sky and partial-sky set-ups. Uncertainties in both parameters correspond to the 68 % credible interval of the marginal posterior distribution.

Mock catalogues		Posterior median estimates of (f_{NL}, b_1)	
		Hybrid-basis analysis	Cartesian power spectrum analysis
Full sky	$f_{\text{NL}} = 0$	$(18^{+53}_{-44}, 3.51^{+0.08}_{-0.08})$	$(6^{+51}_{-43}, 3.52^{+0.08}_{-0.08})$
	$f_{\text{NL}} = 100$	$(117^{+63}_{-54}, 3.46^{+0.09}_{-0.09})$	$(89^{+57}_{-45}, 3.48^{+0.08}_{-0.08})$
Partial sky	$f_{\text{NL}} = 0$	$(11^{+146}_{-112}, 3.50^{+0.20}_{-0.20})$	$(-40^{+136}_{-122}, 3.54^{+0.22}_{-0.20})$
	$f_{\text{NL}} = 100$	$(111^{+158}_{-120}, 3.46^{+0.20}_{-0.20})$	$(62^{+134}_{-124}, 3.50^{+0.20}_{-0.18})$

using the HEALPIX pixelation scheme with $N_{\text{side}} = 256$. A single evaluation of the spherical-basis likelihood (eq. 5.28) takes just under a minute for the partial-sky set-up.⁴ If the hybridisation wavenumber were extended to $k_{\text{hyb}} = 0.05 h \text{ Mpc}^{-1}$, there would be 978 SFB clustering modes with $\ell_{\text{max}} = 20$; the total computation time of both SFB clustering modes and angular coupling coefficients would roughly double. However, we caution that these figures are for reference only: more sophisticated numerical algorithms exist (see the next section), and implementations in different programming languages may also have varying efficiencies.

In Figure 5.10(a), the full-sky $f_{\text{NL}}-b_1$ constraints are presented, which have been marginalised over the 24 halo mock catalogues with $f_{\text{NL}} = 0$ and 20 mock catalogues with $f_{\text{NL}} = 100$. The results from the hybrid-basis and Cartesian power spectrum likelihood analyses are in good agreement. Similarly, in Figure 5.10(b), constraints are shown for the same mock catalogues in the partial-sky set-up with the BOSS-like angular mask and the radial selection cut. As above, results from the hybrid-basis and Cartesian power spectrum analyses are statistically consistent.

To compare the best-fitting parameters from the different posterior distributions, the posterior median estimates for f_{NL} and b_1 are tabulated in Table 5.1, with uncertainties given by the 68 % credible intervals of their marginal posterior distributions. Finally, in Figures 5.11(a) and 5.11(b), we directly compare the window-convolved models of the power spectrum monopole inferred from the posterior distributions in the hybrid-basis and Cartesian power spectrum analyses with the measurements averaged over the different sets of halo mock catalogues.

In both the full-sky and partial-sky set-ups, the recovered models from both analyses are in good agreement with the measurements, with the 68 % credible intervals of the inferred models comparable to the measurement uncertainties given by the estimated covariance matrix. However, it is worth noting that the measurement uncertainties are derived under the assumption that the power spectrum estimates follow the multivariate normal distribution, whereas the credible intervals of the inferred models are obtained from the non-normal posterior distributions of f_{NL} and b_1 .

⁴ For the full-sky set-up where a number of simplifications can be made (see §5.1.3), the covariance matrix \mathbf{S} is diagonal, so the spherical-basis likelihood can be evaluated more quickly.

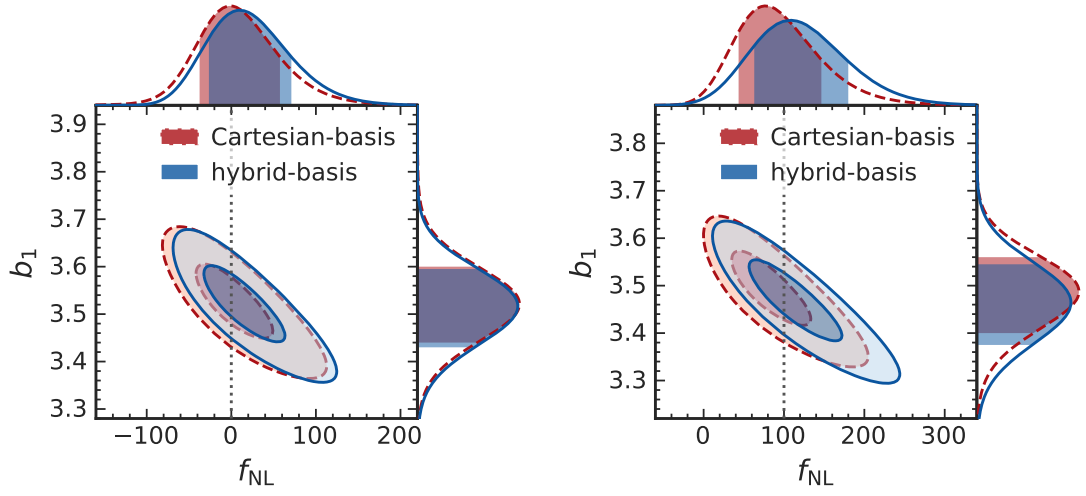


Figure 5.10(a). *Full-sky $f_{\text{NL}}-b_1$ constraints from the hybrid-basis and Cartesian power spectrum likelihood analyses of halo mock catalogues with $f_{\text{NL}} = 0$ (left column) and $f_{\text{NL}} = 100$ (right column).* In the main panels, 1σ and 2σ credible regions of the joint posterior distribution are shown by the shaded contours, and the vertical dotted lines mark the true f_{NL} values. The top and side panels show the marginal posterior distributions for f_{NL} and b_1 respectively, with the shaded regions showing the 1σ credible interval. Hybrid-basis analysis results are coloured in blue and marked by solid lines, whereas Cartesian power spectrum analysis results are coloured in red and marked by dashed lines. Figure taken from Fig. 8 in Paper III.

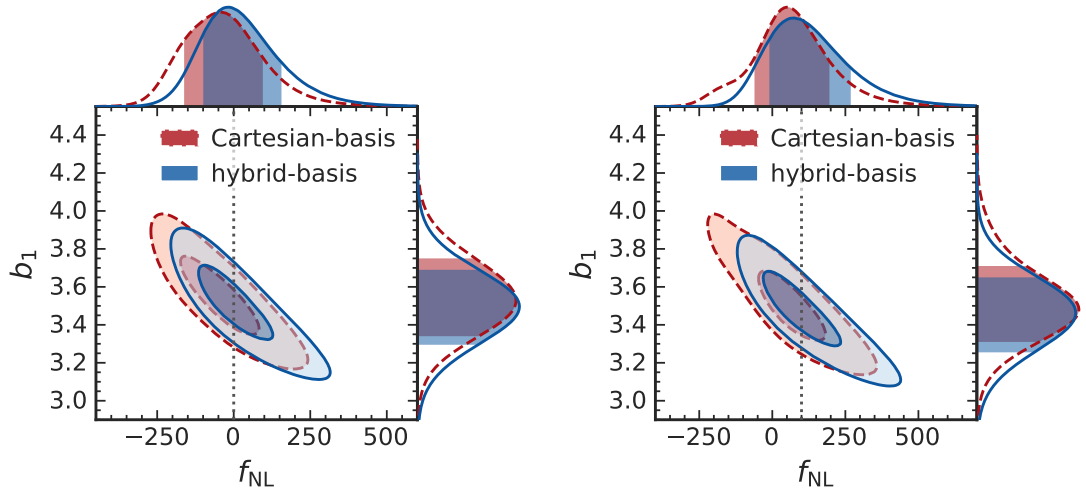


Figure 5.10(b). *Partial-sky $f_{\text{NL}}-b_1$ constraints from the hybrid-basis and Cartesian power spectrum likelihood analyses of halo mock catalogues with $f_{\text{NL}} = 0$ (left column) and $f_{\text{NL}} = 100$ (right column).* In the main panels, 1σ and 2σ credible regions of the joint posterior distribution are shown by the shaded contours, and the vertical dotted lines mark the true f_{NL} values. The top and side panels show the marginal posterior distributions for f_{NL} and b_1 respectively, with the shaded regions showing the 1σ credible interval. Hybrid-basis analysis results are coloured in blue and marked by solid lines, whereas Cartesian power spectrum analysis results are coloured in red and marked by dashed lines. Figure taken from Fig. 9 in Paper III.

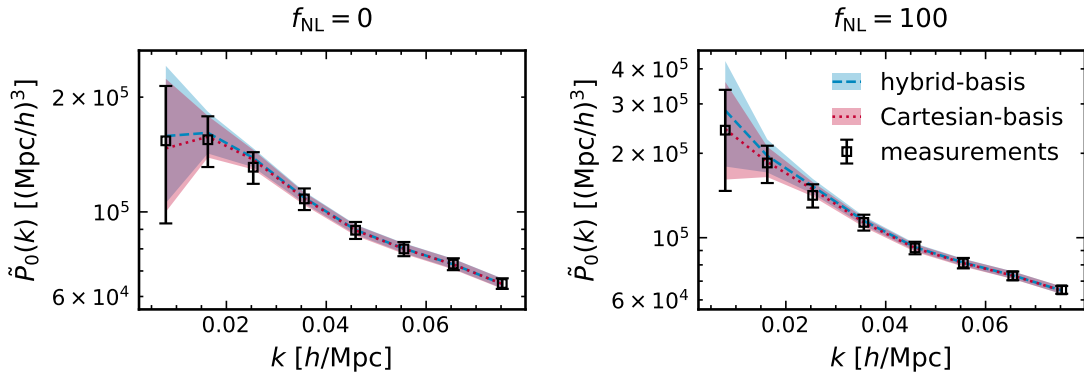


Figure 5.11(a). Inferred models of the window-convolved power spectrum monopole \tilde{P}_0 from the joint posterior distribution of f_{NL} and b_1 compared to the measurements averaged from halo mock catalogues with $f_{\text{NL}} = 0$ (left column) and $f_{\text{NL}} = 100$ (right column) in the *full-sky* set-up. Measurement uncertainties are obtained from the estimated covariance matrix. The shaded regions show the 68 % credible interval of the inferred models from the hybrid-basis likelihood (dashed blue lines) and the Cartesian-basis power spectrum likelihood (dotted red lines). Figure taken from Fig. 10 in Paper III.

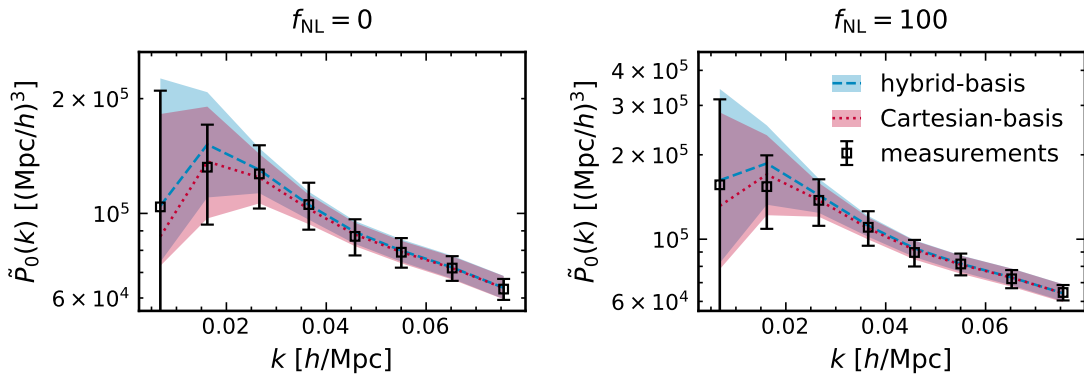


Figure 5.11(b). Inferred models of the window-convolved power spectrum monopole \tilde{P}_0 from the joint posterior distribution of f_{NL} and b_1 compared to the measurements averaged from halo mock catalogues with $f_{\text{NL}} = 0$ (left column) and $f_{\text{NL}} = 100$ (right column) in the *partial-sky* set-up. Measurement uncertainties are obtained from the estimated covariance matrix. The shaded regions show the 68 % credible interval of the inferred models from the hybrid-basis likelihood (dashed blue lines) and the Cartesian-basis power spectrum likelihood (dotted red lines). Figure taken from Fig. 11 in Paper III.

5.4 Summary and discussion

With access to huge cosmic volumes, future galaxy redshift surveys have the potential to probe cosmological physics close to the horizon scale. On such large scales, next-generation missions such as DESI and *Euclid* are forecast to provide local PNG with uncertainties $\sigma_{f_{\text{NL}}} \approx 5$ competitive to the *Planck* result, and relativistic effects in galaxy clustering can also be possibly detected [104–106, 201]. However, various large-scale systematics in the survey, if unaccounted for in the likelihood analysis, threaten to degrade or bias these parameter constraints.

In chapter 2, we have reviewed redshift-space galaxy clustering that is commonly mod-

elled by the anisotropic power spectrum $P(k, \mu)$ or the equivalent Legendre multipoles $P_\ell(k)$ derived in the distant-observer and global plane-parallel approximations by assuming a fixed line of sight $\hat{\mathbf{n}}$. However, the actual line of sight \mathbf{d} varies across the survey volume, so the power spectrum multipoles are commonly estimated with the FFT-based Yamamoto estimator \hat{P}_ℓ derived in the local plane-parallel approximation. The discrepancy between the global plane-parallel prediction and the local plane-parallel estimator, known as the wide-angle effect, makes a significant contribution to systematic errors on large scales when coupled with the survey window function. Although wide-angle corrections have recently been derived, at a more fundamental level, the Cartesian power spectrum analysis is based on Fourier modes decomposed in the plane wave basis, which does not match the spherical geometry of survey observations.

A more natural description of redshift-space galaxy clustering on large scales is the spherical Fourier analysis based on the discrete SFB clustering modes D_μ , since many of the physical and observational effects affect clustering measurements parallel and transverse to the line of sight differently. In this chapter, we have extended previous works (e.g. refs. [49, 228, 244]) to include the AP effect, redshift evolution and scale-dependent galaxy bias on linear scales coherently; further extensions to our model, such as the inclusion of relativistic corrections (see also §4) and scale-dependent linear growth rate in modified gravity theories, should be reasonably straightforward. Although the spherical Fourier analysis offers many advantages such as a clear separation between radial and angular components as well as being fully three-dimensional (i.e. no tomographic binning in redshift z) as discussed in section 5.1.3, it is computationally expensive especially when confronted with huge data sets from future surveys and harder to relate to current models of non-linear galaxy clustering.

Inspired by the hybrid estimator approach used in CMB studies, we have proposed an analogous hybrid-basis approach to analysing LSS observations: below some hybridisation wavenumber k_{hyb} chosen for a given survey, anisotropic galaxy clustering can be accurately described by a spherical Fourier analysis using SFB modes; above k_{hyb} , one switches to the standard Cartesian power spectrum analysis. This approach has some major benefits: no geometric approximations are needed on large scales where a small number of clustering modes can be particularly affected by the survey geometry, and the likelihood directly constructed from SFB clustering modes is exactly Gaussian with an analytically tractable covariance matrix; on smaller scales, the large number of clustering modes can be compressed into power spectrum multipoles, which are computationally fast to evaluate with FFTs and can be related to non-linear galaxy clustering models (e.g. the TNS model for RSD [75]), while the likelihood is now well approximated by a multivariate normal distribution thanks to the central limit theorem.

As a first step in demonstrating the applicability of the hybrid-basis approach, we have analysed real-space clustering statistics of halo mock catalogues from a series of N -body simulations with both Gaussian and non-Gaussian initial conditions. By performing likelihood analysis on the local PNG parameter f_{NL} and the scale-independent halo bias b_1 , we have found that the hybrid-basis approach yields statistically consistent results with those from the Cartesian power spectrum analysis. However, one would expect that, when applied to more realistic scenarios that include RSD and light-cone effects, the hybrid-basis approach will *outperform* the standard analysis for multiple reasons: first, it is in the presence of anisotropic

clustering around the line of sight that the plane-parallel approximations start to break down, unless one corrects for wide-angle effects; secondly, when the clustering measurements span a wide redshift range, the tomographic analysis of power spectrum in multiple redshift bins is less optimal than the fully three-dimensional spherical Fourier analysis; thirdly, the spherical Fourier analysis is better suited for treating angular systematics in future wide surveys. In forthcoming works, the hybrid-basis approach shall be extensively tested with data sets that include RSD and redshift evolution effects, and cosmological parameters other than f_{NL} , such as the linear growth rate f , will also be included in the likelihood analysis. In these more complex scenarios, we will also need to investigate the sensitivity of the hybrid-basis analysis to the hybridisation scale k_{hyb}^{-1} , and whether the low- k and high- k components can still be treated as being effectively independent in the hybrid-basis likelihood; this requires more careful characterisation of the correlation between the low- k spherical and high- k Cartesian data, as well as statistical methods for combining the two regimes when their correlation cannot be neglected. Finally, ref. [276] has recently proposed a modified SFB basis that is tailored for even more realistic survey geometries such as a spherical shell or cap, which can also be implemented in the hybrid-basis framework.

To serve follow-up studies, we have released our public code `HARMONIA` as a Python package designed for both handling catalogue data and modelling clustering statistics in spherical and Cartesian Fourier bases. Optimal weighting schemes for the spherical Fourier analysis [e.g. 49], though not covered in this work, can be readily implemented using the code. In the future, more sophisticated computational algorithms that can accelerate the SFB transform may also be incorporated into this framework [e.g. 240].

Conclusion

The study of the large-scale structure of the Universe has expanded considerably over the past century from curated galaxy catalogues to the wide- and deep-field galaxy surveys of the current era. In 1926, Hubble attempted to test the uniformity of the structure of the Universe through merely four hundred ‘extra-galactic nebulae’ [277], and three years later, he was able to observationally extract the Hubble–Lemaître law [278] that was predicted by Lemaître in 1927 [279].¹ With the Lick catalogue and the Zwicky catalogue [281, 282], which contained tens of thousands of galaxies, Peebles and collaborators were able to demonstrate the power-law behaviour of the angular two-point correlation function of galaxy clustering in the 1970s [e.g. 283]. However, three-dimensional clustering measurements were only made possible later with systematic field surveys such as the pioneering Center for Astrophysics (CfA) Redshift Survey [284] and the *IRAS* mission [285]. The advent of the revolutionary multi-fibre spectrographs has brought the 2dFGRS, the 6-degree Field Galaxy Redshift Survey (6dFGRS) and the SDSS to fruition, with the overall catalogue size in the order of millions including the very recently completed eBOSS [24, 33, 286]. As discussed in chapter 1, these surveys have already helped establish the Λ CDM model as the standard model of cosmology by complementing CMB and SN observations, and in the course of the last two to three decades, the field of LSS cosmology has steadily matured. With powerful probes such as BAO and RSD and the combination of multiple tracer species with $O(10^7)$ – $O(10^8)$ redshift measurements, the eagerly anticipated DESI and *Euclid* surveys will in the coming decade produce constraints on fundamental cosmological parameters that are competitive in comparison with CMB measurements as represented by the *Planck* results [105, 213, 222, 287].

Needless to say, with great instrumental and statistical power comes great responsibility to carry out accurate LSS analyses when ever more refined theoretical models are confronted with increasingly large data sets. As is often discussed in the recent literature, the current cosmological tensions may hint at physical processes beyond the standard Λ CDM model [e.g.

¹ Using Hubble’s catalogue from 1926, Lemaître in fact made a measurement of ‘the coefficient of expansion’ (i.e. the Hubble constant) though noted the large scatter in the results. He omitted the discussion of his determination of the Hubble constant in the translation of his original work into ref. [17] in 1931 [280].

288–291], but it is crucial to control for systematic effects present in the survey pipeline from the very source of observational data to the final statistical inference of cosmological models.² This is indeed the ethos of this thesis, which has reconsidered the different aspects of the prevailing galaxy clustering analysis methodologies; these are reviewed summarily in the following section before the conclusion over future studies and the outlook of LSS cosmology.

6.1 Summary review

Ideally, systematic errors should be limited to levels below the statistical uncertainties in any experiment. These ‘systematics’ can come from a number of places in a galaxy survey analysis: corrupted or incomplete survey data, e.g. due to instrumental limitations such as fibre collisions [e.g. 292–294]; inaccurate modelling of cosmological observables such as the neglect of relativistic corrections; inappropriate mathematical and statistical prescriptions connecting measurement data and theoretical predictions, e.g. wrong probability distributions used for clustering measurements.³ After laying down the theoretical foundation and providing the broad context in the introductory chapters (§§1 and 2) and a series of appendices, we have examined in this thesis various assumptions and treatments in the current framework of galaxy clustering analysis, which may become significant systematics in the future. We have also proposed solutions that address these problems, with a particular focus on the signature of the local-type primordial non-Gaussianity in galaxy clustering, which parametrises the initial conditions for structure formation and serves as an important discriminant for inflationary models. Although each methodology proposed in this thesis has been tested and validated with a focus on the PNG parameter f_{NL} , they are also applicable to other cosmological parameters and are especially relevant to clustering analyses on very large scales.

Chapter 3 investigates the issues of the underlying non-Gaussian likelihood for power spectrum measurements and the cosmological parameter dependence of covariance matrices. Likelihood fitting to two-point clustering statistics made from galaxy surveys almost invariably assumes a multivariate normal distribution for the measurements, with justification based on the central limit theorem given the large number of clustering modes. However, this asymptotic normality assumption cannot hold on the largest survey scales where the number of modes is small. At the same time, a fixed covariance matrix estimate made with expensive mock catalogues produced at a fiducial cosmological model is typically used in the evaluation of the likelihood function, but this assumption has too been demonstrated to be inadequate for statistical inference in previous works. We have therefore devised a simple univariate Gaussianisation scheme, namely the Box–Cox transformation, that results in an approximately normal data distribution, with a variance–correlation decomposition of the covariance matrix to account for its cosmological parameter dependence in linear perturbation theory. By comparing with the standard likelihood assuming asymptotic normality and fixed covariance matrices, we have found from Monte Carlo simulations that our novel likelihood derived with the Gaussi-

² ☞ “Cosmologists are often in error but seldom in doubt.” A quote attributed to Lev Landau by Simon Singh in *Big Bang* (2004).

³ The classification boundary between these types of systematics is not always clear; for instance, the lack of consideration for observational effects in measurements may also be regarded as inadequacy in the theoretical modelling.

anisotropy and covariance matrix rescaling procedure gives more accurate constraints on the local PNG parameter f_{NL} . This improvement fundamentally comes from the superior approximation aimed to recover the full shape of the true posterior distribution which is inaccessible in realistic survey analyses. A streamlined survey analysis pipeline is then presented for straightforward application of the new methodology for forthcoming large galaxy surveys such as DESI and *Euclid*.

We have then focused on relativistic modelling of clustering statistics in chapter 4, which is particularly relevant to the next-generation surveys with access to scales approaching the Hubble horizon, since relativistic corrections are typically suppressed on sub-horizon scales. Relativistic effects in clustering observations have been shown to introduce scale-dependent corrections to the Newtonian galaxy over-density field on large scales, which may hamper the detection of f_{NL} through the scale-dependent tracer bias. The amplitude of relativistic corrections depends not only on the cosmological background expansion, but also on the redshift evolution and sensitivity to the luminosity threshold of the tracer population being examined, as parametrised by evolution bias b_e and magnification bias s . These parameters cannot be accurately predicted in general because of complications such as halo merger history, and thus an empirical approach to determining their values is required. In this work, we have refitted the luminosity function for the eBOSS QSO sample, which has a large tracer bias and a high redshift range that are beneficial to PNG detection, and propagated the results to b_e and s . We have thereby derived constraints on relativistic corrections to the quasar clustering power spectrum multipoles, with more realistic uncertainties compared to previous works assuming fixed fiducial values for the evolution and magnification bias parameters b_e and s . Our results have shown that, for clustering measurements to be made with the DESI QSO sample at wavenumbers $k = O(10^{-3}) h \text{Mpc}^{-1}$, relativistic corrections can overwhelm the expected f_{NL} signature at low redshifts $z \lesssim 1$ and remain comparable to $f_{\text{NL}} \approx 1$ in the power spectrum quadrupole at redshifts $z \gtrsim 2.5$. Although one could mitigate their impact by adjusting the redshift range or the luminosity threshold of the tracer sample being considered, we suggest that, for future surveys probing large cosmic volumes, relativistic corrections should be forward-modelled from the tracer luminosity function including its uncertainties.

In chapter 5, we have turned our attention to the mathematical description of cosmic fluctuations observed through the survey window: the choice of Fourier basis for decomposing clustering statistics that matches the geometry of observations. The Cartesian power spectrum analysis of anisotropic galaxy clustering based on the plane wave basis makes a number of assumptions, including the distant-observer and plane-parallel approximations, that will no longer be valid on very large scales in future surveys and may degrade cosmological constraints. Inspired by the hybrid likelihoods used in CMB studies by *Planck*, we have proposed an approach that utilises a hybrid basis: on the largest scales, clustering statistics are decomposed into spherical Fourier–Bessel modes which respect the natural survey geometry and fully capture physical and observational effects along the line of sight such as redshift-space distortions and light-cone effects; on smaller scales with far more clustering modes that are much less affected by the aforementioned systematics, one retains the computational edge of the FFT-based power spectrum analysis which also naturally extends to non-linear scales. The direct use of individual clustering modes in constructing the likelihood function on large scales also circumvents the issues of non-Gaussian likelihoods and covariance matrix estimation. We have

validated this methodology by testing on real-space halo clustering in N -body simulations, and publicly released our code HARMONIA for future investigation and application.

6.2 Future work and outlook

Based on the discussions in the preceding chapters, there are various avenues for further research, which also include ideas that are more explorative in nature.

A unified framework for clustering analyses. In the hybrid-basis approach advocated in this thesis, we have emphasised the utilisation of the spherical Fourier–Bessel basis for analysing cosmic density fluctuations on the largest scales in wide- and deep-field surveys. Not only does this eliminate wide-angle effects present in a Cartesian power spectrum analysis, but the natural separation between radial and angular clustering components also makes it a superior analytic tool for the full modelling of relativistic corrections including the lensing terms ignored in chapter 4, where the plane-parallel power spectrum multipoles have been adopted as the observables for simplicity. Indeed, ref. [229] has already made the detailed translation from relativistic corrections to the SFB clustering modes, though in its derivation these modes are not discretised according to any survey boundary conditions. Therefore one could envisage a unified framework for Fourier clustering analyses combining all the methodologies considered in this thesis: the full relativistic treatment of galaxy clustering with the SFB decomposition is used on the largest survey scales subject to limited computational resources and the requirement to have systematic errors below the levels of statistical uncertainties; on intermediate linear scales, the Gaussianisation and covariance matrix rescaling strategy is still useful for constructing a power spectrum likelihood function that reproduces the true cosmological posterior distribution as closely as possible. Following the work presented in chapter 5, testing the hybrid-basis methodology with RSD effects included is a natural next step, and the correlation between spherical and Cartesian Fourier clustering measurements, as well as how to appropriately combine them, also warrants further investigation. Finally, efficient numerical algorithms, e.g. the implementation of the SFB transform, are worth exploring to reduce the computational cost of the hybrid-basis approach.

Multi-tracer approach. The potential of next-generation galaxy surveys to detect primordial non-Gaussianity lies with the unprecedented cosmic volume they observe which avails the probe of ultra-large-scale fluctuations. Unfortunately, clustering modes on such scales are few and thus cosmic variance dominates, which diminishes the prospect of detection. Ref. [223] has made use of the insight that on large scales the PNG signature is in the non-stochastic tracer bias, and by cross-correlating two or more tracers with contrasting bias values, one can significantly reduce the impact of cosmic variance in the underlying matter distribution. The same idea is also applicable to measuring redshift-space distortion parameters such as the ratio f/b_1 between the linear growth rate f and the linear bias b_1 [224]. In addition, as manifested in equation (4.5), the cross-correlation between clustering modes of two tracers with different evolution and magnification biases will have a non-vanishing imaginary part, an insight noted by ref. [295] and suggested as a method for detecting relativistic effects. However, it should be

mentioned that to eliminate or reduce cosmic variance, the multiple tracers considered must have overlapping volume so that their samples trace the same patch of the underlying matter density field, and ideally the difference in their bias values should be sufficiently large; there is also the practical challenge that tracers may have different survey systematics. Since these early works, the multi-tracer technique has spurred a large number of studies involving both Fisher forecasts and survey measurements, which have shown that it may help reduce the uncertainties on various parameters by a factor of a few [see e.g. 296, 297]; one such example is the recently completed eBOSS analysis [298], which has been able to combine the LRG and ELG samples in the redshift range $0.6 < z < 1.0$ with a significant sky coverage of 9500 deg^2 . Since most of these studies have considered the power spectrum or other two-point clustering statistics as the observable, it would be novel and exciting to explore whether this technique can be embedded in the SFB analysis based on individual, uncompressed clustering modes as elaborated in chapter 5. These efforts may ultimately lead to a hybrid-basis multi-tracer Fourier clustering analysis framework serving future-generation surveys even beyond DESI and *Euclid*, e.g. the Spectro-Photometer for the History of the Universe, Epoch of Reionization, and Ices Explorer (SPHEREx) [299].

Application of random matrix theory. The ubiquity of the Gaussian likelihood, as discussed in chapters 1 and 3, stems from the fact that asymptotic normality is a *universality law* as warranted by the central limit theorem. Such universality laws do not only exist for random variables or vectors, but also for random matrices as predicted by *random matrix theory*, which was first introduced by Wishart [300] and popularised by Wigner [301–303]. Typically, universality laws in random matrix theory state that as the dimensions of a random matrix become large, the distribution of its eigenvalues converges to some specific spectrum. In cosmology, random matrix theory has already been applied to the study of multi-field inflation models, where the Hessian matrix of the field potential can be described as random matrices, and the dynamics of the system of fields can be analysed even when the number of fields becomes very large [e.g. 304–306]. Random matrix theory might also be of use for more data-oriented LSS analyses. For example, the observed galaxy over-density field may be contaminated by sources of galactic origins such as dust extinction, and foreground removal methods are sometimes based on the covariance matrix of clustering modes at the map level to remove the contaminant components, e.g. mode subtraction or mode de-projection [307–309]. Random matrix theory may offer alternative methods that could in principle work particularly well when the number of clustering modes, and thus the dimensions of the covariance matrix, are large; in this case, eigenvalues of the sample covariance matrix that fall outside the spectrum predicted by random matrix theory may be regarded as unexpected ‘signals’, i.e. they could be contaminant components. Similar random matrix techniques based on the comparison between a predicted eigenvalue spectrum and the observed one are common for cleansing estimated covariance matrices in modern portfolio theory in quantitative finance [e.g. 310], though it remains to be investigated how such methods could be practically utilised for improving covariance matrix estimation or removing systematic effects from foreground contaminants in clustering measurements.



The field of large-scale structure studies will undoubtedly be at the forefront in the so-called ‘golden era’ of observational cosmology. Such optimism is founded on the rich data sets that multiple generations of galaxy surveys have brought and will continue to bring in the future, and one only needs to look to the history of the cosmic microwave background for further convincing: from the prediction by Alpher and Hermann in 1948 [21] to the accidental discovery by Penzias and Wilson in 1965 [22] and then to the first CMB map produced by *COBE* in the 1990s, it took nearly four decades; yet in the last two decades, *WMAP* and *Planck* have measured the CMB temperature to exquisite precision, with minuscule fluctuations in CMB anisotropies and polarisation used to firmly establish the Λ CDM paradigm and constrain cosmological parameters to sub-percent levels [14]. Since the groundbreaking work laid down by Peebles almost half a century ago for which he was awarded the Nobel Prize in 2019 [31], and after multiple decades of endeavour from countless others, the field of large-scale structure cosmology too has flourished and is no longer data-starved. Although history is viewed in retrospect, one could still proclaim with confidence that the golden era has only just begun.

Appendices



A

Cosmological Background

The materials presented in this appendix are mostly based on refs. [47, 312, 313] except where referenced otherwise.



FLRW metric. Under the cosmological principle, the background space-time must admit a metric g_{ab} with the line element in the form of

$$ds^2 = g_{\mu\nu} dX^\mu dX^\nu = -dt^2 + a(t)^2 \gamma_{ij} dx^i dx^j, \quad (A.1)$$

where the four-vector $X^\mu = (X^0, \mathbf{x})$, a is a scale factor evolving with physical time t , and the spatial metric γ_{ij} is maximally-symmetric²—it can only describe a spherical, Euclidean or hyperbolic space with positive, zero or negative constant Gaussian curvature K respectively,³

$$\gamma_{ij} dx^i dx^j = \frac{dr^2}{1 - Kr^2} + r^2 d\Omega^2, \quad (A.2)$$

where r and Ω are the radial and solid-angle coordinates. The FLRW line element is thus given by

$$ds^2 = -dt^2 + a(t)^2 \left(\frac{dr^2}{1 - Kr^2} + r^2 d\Omega^2 \right), \quad (A.3)$$

and by rescaling the length dimension, one may fix $a = 1$ at the current epoch. It is often instructive to consider the *conformal time* τ with

$$d\tau = \frac{dt}{a} \quad (A.4)$$

¹ The metric signature follows the $(-, +, \dots)$ convention, and natural units are adopted in which the speed of light in vacuum and the gravitational constant are unity, $c = G = 1$. Greek indices μ, ν, \dots indicate tensor components and Latin indices a, b, \dots are abstract indices.

² That is, the spatial metric has the maximum number of Killing vectors as warranted by homogeneity and isotropy. See e.g. ref. [314] for proof.

³ This curvature is proportional to the Riemannian three-curvature of the spatial submanifold.

and remove the coordinate singularity in r using

$$d\chi = \frac{dr}{\sqrt{1 - Kr^2}}, \quad (\text{A.5})$$

so that $ds^2 = a(\tau)^2(-d\tau^2 + d\Sigma_K^2)$ where

$$d\Sigma_K^2 = d\chi^2 + S_K(\chi)^2 d\Omega^2 \quad (\text{A.6a})$$

measures *comoving distances* and

$$S_K(\chi) = \begin{cases} K^{-1/2} \sin(K^{1/2}\chi) & (K > 0) \\ \chi & (K = 0) \\ |K|^{-1/2} \sinh(|K|^{1/2}\chi) & (K < 0) \end{cases}. \quad (\text{A.6b})$$

In the case of zero spatial curvature, $K = 0$, the FLRW metric is simply conformal to the Minkowski metric η_{ab} , i.e. $g_{\mu\nu} = a(\tau)^2 \eta_{\mu\nu}$.

Redshift and the Hubble expansion. Since there are no cosmic clocks that can directly measure the time interval between two events, it is more practical to consider the *cosmological redshift* z of the electromagnetic spectrum of photons travelling from a distant source to the observer as a proxy temporal coordinate, with the observer at the current epoch $z = 0$. The cosmological redshift is related to the scale factor by

$$a = (1 + z)^{-1}, \quad (\text{A.7})$$

which can be derived by considering the frequency change between emission and reception of photons travelling along a null geodesic between the source and observer, whose comoving distance is constant. The expansion of the Universe resulting in this redshift can be measured using the *Hubble parameter* defined by

$$H = \frac{d \ln a}{dt} = \frac{\dot{a}}{a}, \quad (\text{A.8a})$$

or equivalently, the *conformal Hubble parameter*

$$\mathcal{H} = \frac{d \ln a}{d\tau} = \frac{a'}{a} = aH. \quad (\text{A.8b})$$

Here an overdot $\dot{\cdot}$ denotes a cosmic-time derivative, and a thick prime \cdot' denotes a conformal-time derivative. In practice, the *reduced Hubble parameter* h is often used to report measurements of the Hubble parameter of today,

$$h = H_0 / (100 \text{ km s}^{-1} \text{ Mpc}^{-1}), \quad (\text{A.9})$$

where \cdot_0 denotes the value of a quantity at the current redshift $z = 0$.

Dynamics of a Λ CDM universe. The major components of a Λ CDM universe are idealised as *perfect fluids* without shear or viscosity, whose density ρ and pressure P are related by an *equation of state*,

$$P = w\rho, \quad (A.10)$$

where w is a dimensionless parameter; for example, $w = 0$ for pressureless matter, $w = -1$ for dark energy as a cosmological constant Λ , and $w = 1/3$ for radiation. The corresponding stress–energy tensor is given by

$$T_{ab} = (\rho + P)U_a U_b + P g_{ab}, \quad (A.11)$$

where U^a is the fluid four-velocity. By energy–momentum conservation, $\nabla^a T_{ab} = 0$, one can infer a mass conservation law that is analogous to the first law of thermodynamics for an adiabatic process,

$$\dot{\rho} + 3H(\rho + P) = 0; \quad (A.12)$$

thus with the equation of state for a perfect fluid, $\rho \propto a^{-3(1+w)}$. This result also naturally follows from the Einstein field equation, $G_{ab} = 8\pi T_{ab}$, since the contracted Bianchi identity $\nabla^a G_{ab} = 0$ implies energy–momentum conservation.⁵ Indeed, if one considers the purely temporal and spatial components of the Einstein field equation, the *Friedmann equations*,

$$H^2 = \frac{8\pi}{3}\rho - \frac{K}{a^2}, \quad (A.13a)$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi}{3}(\rho + 3P). \quad (A.13b)$$

would emerge and lead to equation (A.12). It is evident that the dynamics of the Hubble expansion depends on the matter–energy contents of the Universe as parametrised by the *critical density*

$$\rho_c = \frac{3H^2}{8\pi}, \quad (A.14)$$

which corresponds to the density in the absence of spatial curvature, $K = 0$. The *density parameter* Ω_I for each individual component I is given by

$$\Omega_I = \frac{\rho_I}{\rho_c}, \quad \text{with} \quad \Omega_\Lambda = \frac{\Lambda}{3H^2} \quad \text{and} \quad \Omega_K = -\frac{K}{a^2 H^2}. \quad (A.15)$$

Therefore it is identically true by construction that $\sum_I \Omega_I = 1$, and the (first) Friedmann equation (A.13a) becomes

$$H^2 = H_0^2 E^2, \quad \text{where} \quad E^2 = \sum_{I \neq K} \Omega_{I,0} a^{-3(1+w)} + \frac{\Omega_{K,0}}{a^2}. \quad (A.16)$$

⁴ Only in this appendix and the next does P denote pressure in addition to the power spectrum, though confusion is unlikely to arise.

⁵ The Einstein tensor (also known as the trace-reversed Ricci curvature tensor), $G_{ab} = R_{ab} - (R/2)g_{ab}$, is constructed from the Ricci curvature tensor R_{ab} and the Ricci curvature scalar R for the metric g_{ab} .

Cosmological distances. There are several distance measures that could have been introduced earlier in the discussion of the FLRW metric, but it is more useful to express them now in terms of redshift z and the density parameter(s) Ω defined above. The *radial*, or *line-of-sight*, *comoving distance* is defined by

$$D_C(z) = \chi = D_{H,0} \int_0^z \frac{dz}{E(z)}, \quad (\text{A.17a})$$

and the *transverse*, or *metric*, *comoving distance* is

$$D_M(z) = S_K(\chi) = \begin{cases} \frac{D_{H,0}}{\sqrt{\Omega_{K,0}}} \sin \left[\sqrt{\Omega_{K,0}} \frac{D_C(z)}{D_{H,0}} \right] & (\Omega_{K,0} > 0) \\ D_C(z) & (\Omega_{K,0} = 0) \\ \frac{D_{H,0}}{\sqrt{|\Omega_{K,0}|}} \sinh \left[\sqrt{|\Omega_{K,0}|} \frac{D_C(z)}{D_{H,0}} \right] & (\Omega_{K,0} < 0) \end{cases}, \quad (\text{A.17b})$$

where $D_H(z) = H(z)^{-1}$ is the Hubble distance. Sometimes, the *angular diameter distance* and *luminosity distance* are also used, which are respectively given by

$$D_A(z) = D_M(z)/(1+z) \quad \text{and} \quad D_L(z) = (1+z) D_M(z). \quad (\text{A.18})$$

It is worth noting that, whilst χ is introduced in equation (A.5) as a spatial coordinate variable that is formally independent of any temporal coordinate variable such as z , the comoving distances defined above by considering photon null geodesics with $ds^2 = 0$ relate the two. In particular, in the case of zero spatial curvature $K = 0$, these distance measures as well as the radial coordinate r all reduce to χ , i.e. $\chi = r = D_C(z) = D_M(z)$. As such, χ is often treated as a function of redshift, and the relationship $r = \chi(z)$ is known as the *distance–redshift relation*, with r now denoting the comoving distance between an observer and a source galaxy that emits the photons. At low redshift, this reduces to an approximate relation known as the *Hubble–Lemaître law*, $z \approx H_0 D$, where D is the physical distance to a galaxy at redshift z [278].

B

Linearised Cosmological Perturbations

The materials presented in this appendix are mostly based on refs. [47, 312, 315–317] except wherever referenced otherwise. Since the spatial-curvature density parameter $\Omega_{K,0}$ is observed to be consistent with zero, a flat FLRW background space-time is assumed.



Metric perturbations. In general, perturbations to the space-time metric can be scalars, vectors or tensors, which decouple at linear order. Since vector perturbations correspond to vorticity, which is diluted away with cosmological expansion unless actively sourced, and tensor perturbations are associated with gravitational waves, the focus here is on scalar perturbations to the flat FLRW metric which now takes the general form

$$ds^2 = a(\tau)^2 \left\{ - (1 + 2A) d\tau^2 + B_{,i} dx^i d\tau + \left[(1 + 2C)\delta_{ij} + \left(\partial_i \partial_j - \frac{1}{3} \delta_{ij} \nabla^2 \right) E \right] dx^i dx^j \right\}, \quad (\text{B.1})$$

where $\cdot_{,\mu}$ denotes a partial derivative with respect to the coordinate vector X^μ , and δ_{ab} is the Kronecker delta tensor. Under a general coordinate transformation,

$$X^\mu \mapsto X^\mu + (T, L^i), \quad (\text{B.2})$$

the scalar perturbations A, B, C and E transform as

$$A \mapsto A - T' - \mathcal{H}T, \quad (\text{B.3a})$$

$$B \mapsto B + T - L', \quad (\text{B.3b})$$

$$C \mapsto C - \mathcal{H}T - \nabla^2 L / 3, \quad (\text{B.3c})$$

$$E \mapsto E - L. \quad (\text{B.3d})$$

The freedom in the choice of a coordinate system means that there are physically equivalent gauges: in *synchronous gauge* with $A = B = 0$, the time component is unperturbed; in *spatially*

flat gauge with $C = E = 0$, the spatial components are unperturbed; in *longitudinal gauge* with $B = E = 0$, the spatial perturbation can be identified with the gravitational potential in Newtonian gravity, and thus this gauge is also known as the *Newtonian gauge*. To avoid any ambiguity, gauge-invariant observables can be constructed: the *Bardeen potentials*

$$\Psi = A + \mathcal{H}(B - E') + (B - E')' \quad (\text{B.4a})$$

$$\text{and } \Phi = -C - \mathcal{H}(B - E') + \frac{1}{3}\nabla^2 E \quad (\text{B.4b})$$

are a prime example. In Newtonian gauge, the perturbed flat FLRW metric then takes the following form:

$$ds^2 = a(\tau)^2 \left[-(1 + 2\Psi) d\tau^2 + (1 - 2\Phi) \delta_{ij} x^i x^j \right]. \quad (\text{B.5})$$

Matter perturbations. Similar to the metric perturbations above, the stress–energy tensor admits scalar perturbations in the form of

$$\delta T_{\mu\nu} = (\delta\rho + \delta P) \bar{U}_\mu \bar{U}_\nu + (\bar{\rho} + \bar{P}) \left(\bar{U}_\mu \delta U_\nu + \delta U_\mu \bar{U}_\nu \right) + \delta P \delta_{\mu\nu} + \left(\partial_i \partial_j - \frac{1}{3} \delta_{ij} \nabla^2 \right) \Pi, \quad (\text{B.6})$$

where $\bar{\cdot}$ denotes a background quantity, $\delta\cdot$ denotes a perturbation quantity and Π is associated with anisotropic stress. Since the four-velocity for massive particles satisfies $U^a U_a = -1$ and the unperturbed $\bar{U}^\mu = a^{-1} \delta_0^\mu$, the four-velocity perturbation is $\delta U_\mu = a^{-1} (-A, \partial_i V)$ for some scalar perturbation V associated with the three-velocity. Under the coordinate transformation (eq. B.2),

$$\delta\rho \mapsto \delta\rho - T\bar{\rho}', \quad (\text{B.7a})$$

$$\delta P \mapsto \delta P - T\bar{P}', \quad (\text{B.7b})$$

$$V \mapsto V + L', \quad (\text{B.7c})$$

$$\Pi \mapsto \Pi. \quad (\text{B.7d})$$

There are two common matter gauges: in *uniform-density gauge*, $\delta\rho = B = 0$; and in *comoving gauge*, $V = B = 0$. A gauge-invariant observable can similarly be constructed: the *comoving density contrast* is defined as

$$\Delta = \frac{\delta\rho}{\bar{\rho}} + \frac{\bar{\rho}'}{\bar{\rho}} (V + B), \quad (\text{B.8})$$

which is simply the density contrast $\delta = \delta\rho/\bar{\rho}$ in comoving gauge.¹

Linearised field equation. By equating metric and matter perturbations, the linearised Einstein field equation yields at first order in, for example, Newtonian gauge the following system of equations:

$$\Phi - \Psi = \Pi, \quad (\text{B.9a})$$

$$\nabla^2 \Phi = 4\pi a^2 \bar{\rho} \delta + 3\mathcal{H}(\Phi' + \mathcal{H}\Phi), \quad (\text{B.9b})$$

¹ Only in this appendix and the next does δ denote density contrast in general instead of the over-density field of tracer number counts.

$$\Phi' + \mathcal{H}\Phi = -4\pi a^2 (\bar{\rho} + \bar{P})V, \quad (\text{B.9c})$$

$$\Phi'' + 3\mathcal{H}\Phi' + (2\mathcal{H}' + \mathcal{H}^2)\Phi = 4\pi a^2 \delta P. \quad (\text{B.9d})$$

In particular, the first equation here means the Bardeen potentials are identical in the absence of anisotropic stress, $\Pi = 0$; the second and third equations combine to give the *Poisson equation*,

$$\nabla^2 \Phi = 4\pi a^2 \bar{\rho} \Delta, \quad (\text{B.10})$$

if the conservation law (eq. A.12) is also employed. The energy–momentum conservation law $\nabla^a T_{ab} = 0$ yields the continuity and Euler equations,

$$\delta' + (1+w)(\nabla \cdot \mathbf{v} - 3\Phi') + 3\mathcal{H}(c_s^2 - w)\delta = 0, \quad (\text{B.11a})$$

$$\mathbf{v}' + \mathcal{H}\left(1 - 3\frac{\bar{\rho}'}{\bar{\rho}}\right)\mathbf{v} = -\frac{c_s^2}{1+w}\nabla\delta - \nabla\Psi, \quad (\text{B.11b})$$

where $c_s = (\delta P/\delta\rho)^{1/2}$ is the speed of sound and $\mathbf{v} = \nabla V$ the peculiar velocity. However, not all these equations are independent, since energy–momentum conservation is equivalent to the contracted Bianchi identity.

Growth of matter density perturbations. Equipped with the linearised perturbation equations above, one can readily derive the equation governing the growth of matter density perturbations in the late-time Universe with negligible radiation. Neglecting the pressure perturbation, $\delta P \approx 0$, since there are no fluctuations in dark energy as a cosmological constant, equations (B.9d) and (B.10) combine to give

$$\Delta'' + \mathcal{H}\Delta' + (\mathcal{H}' - \mathcal{H}^2)\Delta = 0. \quad (\text{B.12})$$

By changing the time variable from conformal time τ to the scale factor a [318], this equation becomes

$$\frac{d^2(\Delta/H)}{da^2} + 3\frac{d\ln(aH)}{da}\frac{d(\Delta/H)}{da} = 0. \quad (\text{B.13})$$

There are two solutions: the decaying one corresponds to $\Delta \propto H$ which is decreasing, and the growing solution can be found by integrating the equation above once,

$$\Delta \propto H \int_{a_i} \frac{da}{a^3 H(a)^3}, \quad (\text{B.14})$$

where the subscript ‘i’ denotes some initial epoch deep in the matter-dominated era. The *linear growth factor* D_1 simply relates the value of Δ ($\approx \delta_m$ on sub-horizon scales) at a given epoch to its value at some initial (or the current) epoch, so it takes the following form:

$$D_1(z) = \lambda H(z) \int_z^{z_i} \frac{1+\tilde{z}}{H(\tilde{z})^3} d\tilde{z}, \quad (\text{B.15})$$

where λ is a normalisation factor yet to be determined. Interestingly, deep in the matter-dominated era,

$$H(z)^2 = H_0^2 E(z)^2 \approx H_0^2 \Omega_{m,0}(1+z)^3, \quad (\text{B.16})$$

and thus

$$D_1(z) = \frac{2}{5} \frac{\lambda}{H_0^2 \Omega_{m,0}} a^{-3/2} \left(a^{5/2} - a_i^{5/2} \right) \approx \frac{2}{5} \frac{\lambda}{H_0^2 \Omega_{m,0}} a \quad (\text{B.17})$$

is proportional to the scale factor—as a result, one normalisation convention is $\lambda_{\text{early}} = (5/2)H_0^2\Omega_{m,0}$, labelled with the subscript ‘early’, so that $D_1(z) = a(z)$ in the matter-dominated era. However, the normalisation convention with $D_1(z=0) = 1$ is sometimes more convenient. The quadrature solution above evaluates to

$$D_1(0) = \frac{\lambda}{H_0^2} \left[\frac{(1 - \Omega_{m,0})^{-3/2}}{2} (1+z)^2 {}_2F_1\left(\frac{2}{3}, \frac{3}{2}; \frac{5}{3}; -\frac{\Omega_{m,0}}{1-\Omega_{m,0}}(1+z)^3\right) \right]_0^{z_i}, \quad (\text{B.18})$$

where ${}_2F_1$ denotes the ordinary hypergeometric function; by expanding this result as a series around $z_i = \infty$ and keeping the leading order, one obtains $\lambda_{\text{late}} = C \lambda_{\text{early}}$, labelled with the subscript ‘late’, where the ratio of the two different normalisation factors is

$$C \approx \frac{2}{5} \Omega_{m,0}^{-1} \left[\pi^{-1/2} \Gamma\left(\frac{5}{6}\right) \Gamma\left(\frac{5}{3}\right) (1 - \Omega_{m,0})^{-5/6} \Omega_{m,0}^{-2/3} - \frac{1}{2} (1 - \Omega_{m,0})^{-3/2} {}_2F_1\left(\frac{2}{3}, \frac{3}{2}; \frac{5}{3}; -\frac{\Omega_{m,0}}{1-\Omega_{m,0}}\right) \right]^{-1}, \quad (\text{B.19})$$

and Γ is the gamma function. This normalisation factor is independent of the initial epoch and, unsurprisingly, dependent on the matter density parameter. Another important quantity parametrising the evolution of matter density perturbations is the *linear growth rate*,

$$f = \frac{d \ln D_1}{d \ln a}, \quad (\text{B.20})$$

whose measurements self-evidently provide constraints on the matter density parameter. Although a precise formula for f can be readily derived from the quadrature solution (eq. B.15), there are some convenient approximations, e.g. $f \approx \Omega_m^\gamma$ for some power law index γ [10]. With the introduction of the growth rate f , the continuity equation (B.11a) becomes

$$\nabla \cdot \mathbf{v} = -\mathcal{H} f \delta_m \quad (\text{B.21})$$

for pressureless, dust-like matter.

Matter power spectrum. As favoured by current observations, many models of inflation predict an almost scale-invariant primordial power spectrum for scalar perturbations,

$$P_{\mathcal{R}}(k) = \frac{2\pi^2}{k^3} A_s k^{n_s-1}, \quad (\text{B.22})$$

where A_s and n_s are the *scalar perturbation amplitude* and *spectral index*, and

$$\mathcal{R} = -C + \frac{1}{3} \nabla^2 E - \mathcal{H}(V + B) \quad (\text{B.23})$$

is the *comoving curvature perturbation*.² Sometimes, the *uniform-density curvature perturbation*

$$\zeta = -C + \frac{1}{3} \nabla^2 E + \mathcal{H} \frac{\delta\rho}{\bar{\rho}'} \quad (\text{B.24})$$

² Scale invariance here refers to the dimensionless power spectrum defined with respect to $\ln k$ rather than k , and corresponds to $n_s = 1$. Any pivot scale is absorbed into A_s here; see e.g. refs. [14, 101].

is used as the curvature variable instead of \mathcal{R} . To propagate from the primordial perturbations to the (linear) matter power spectrum P_m at a late-time epoch, one needs a *transfer function* T that encodes intermediate physical processes,³

$$P_m(k, z) = D_1(z)^2 T(k)^2 P_{\mathcal{R}}(k). \quad (\text{B.25})$$

Here the linear growth factor D_1 is factored in separately from T , which processes primordial curvature perturbations into density perturbations at some initial epoch of structure formation, which then begin scale-independent linear evolution on large scales. The precise computation of the transfer function is challenging since there is a mixture of massive and relativistic particles at early times, and different particle species interact through complex scattering processes. However, some simplified scale dependence of the transfer function can be derived for different eras dominated by a single fluid component (e.g. the radiation-dominated or matter-dominated era) [see e.g. 47, 312], and detailed fitting formulæ exist [e.g. 97, 186]. In general, an approach based on relativistic kinetic theory is required, and one needs to solve a system of coupled Einstein–Boltzmann equations numerically, for which there are existing public codes such as CAMB [319, 320] and CLASS [321].

³ Care needs to be taken when referring to a transfer function in a specific context as to what perturbations are being ‘transferred’.

Non-Linear Structure Formation

The materials presented in this appendix are mostly based on refs. [68, 98, 312, 322, 323] except where referenced otherwise.



To understand the issue of galaxy bias, it is imperative to address the environment in which galaxies reside—haloes, which are stable, gravitationally-bound objects that have collapsed in over-dense regions and undergone virialisation. This is a highly non-linear process, yet it is still possible to make reasonable predictions relating the distribution of haloes to the underlying matter distribution, at least on large scales.

Spherical collapse. In the idealised spherical-collapse scenario of halo formation, one considers a spherical region of physical radius R_0 at some initial epoch in a matter-dominated universe. Perturbations within this spherical region evolve in a fashion independent of the region outside, a consequence of Birkhoff's theorem, the relativistic analogue to Newton's shell theorem. It follows that the solution for both the interior and exterior regions has the scalar factor $a(t) = (3H_0 t/2)^{1/2}$ from the Friedmann equation (A.13a) with matter domination, and the physical radius R and cosmic time t are linked by the following parametric solution:

$$R = R_0 A(1 - \cos \theta), \quad t = B(\theta - \sin \theta), \quad (\text{C.1a})$$

with

$$A = \frac{\Omega_{m,0}}{2(\Omega_{m,0} - 1)} \quad \text{and} \quad B = \frac{\Omega_{m,0}}{2H_0(\Omega_{m,0} - 1)^{3/2}}. \quad (\text{C.1b})$$

Expanding both R and t to first order in parameter θ , one finds

$$R \approx R_0 \frac{A}{2} \left(\frac{6t}{B} \right)^{2/3} \left[1 - \frac{1}{20} \left(\frac{6t}{B} \right)^{2/3} \right], \quad (\text{C.2})$$

and thus the over-density evolves approximately as

$$\delta \approx \frac{3}{20} \left(\frac{6t}{B} \right)^{2/3}. \quad (\text{C.3})$$

The *critical over-density* for spherical collapse is defined when the parametric solution first reaches $R = 0$ for $t > 0$, i.e. $\theta = 2\pi$, so that

$$\delta_c = \frac{3}{20}(12\pi)^{2/3} \approx 1.686. \quad (\text{C.4})$$

Naturally, the final condition $R = 0$ is never reached because the linearised solution breaks down, as clearly indicated by an over-density value $\delta > 1$. Nonetheless, δ_c serves as an indicator when one might expect a halo to have already collapsed. By employing the virial theorem, in particular applied to the turn-around point $\theta = \pi$ when R stops expanding and starts contracting, one can deduce that the virialised halo density is $18\pi^2 \approx 178$ times the background density. The value of the critical over-density above is derived in a matter-dominated universe and associated with a particular epoch, and more generally it will evolve in redshift with other cosmological parameters; a common approximation in practice is to extrapolate its value from above with the linear growth factor D_1 ,

$$\delta_c(z) \approx \frac{1.686}{D_1(z)}, \quad (\text{C.5})$$

though methods to compute $\delta_c(z)$ in more general cosmological models exist [see e.g. 324].

Halo mass function. In order to associate a spherical region of over-density and a corresponding mass to each point in a random field, one could consider a smoothed matter over-density field that has been filtered by some isotropic function W with a characteristic scale R ,

$$\delta_R(\mathbf{r}) = \int d^3\mathbf{x} W(\mathbf{r} - \mathbf{x}; R) \delta(\mathbf{x}), \quad (\text{C.6})$$

where the filter is normalised as $\int d^3\mathbf{x} W(\mathbf{x}; R) = 1$. For a homogeneous and isotropic GRF δ with power spectrum P , the variance of the smoothed field δ_R is then

$$\langle \delta_R^2 \rangle = \int \frac{k^2 dk}{2\pi^2} W(kR)^2 P(k), \quad (\text{C.7})$$

where convolution becomes multiplication in Fourier space. An important cosmological parameter, σ_8 , is thereby defined as

$$\sigma_8^2 = \int \frac{k^2 dk}{2\pi^2} W_{\text{TH}}(kR)^2 P_{\text{lin}}(k) \quad \text{with} \quad R = 8 h^{-1} \text{ Mpc} \quad (\text{C.8})$$

for the linear power spectrum (labelled with the subscript ‘lin’), where the filter is a rectangular function, also known as a ‘top-hat’ filter,

$$W_{\text{TH}}(\mathbf{r}; R) = \frac{3}{4\pi R^3} \Theta(1 - r/R) \quad (\text{C.9a})$$

$$\text{with} \quad W_{\text{TH}}(kR) = \frac{3}{(kR)^3} [\sin(kR) - kR \cos(kR)], \quad (\text{C.9b})$$

where Θ is the Heaviside function. The σ_8 parameter thus represents the root-mean-square linear density fluctuation within an $8 h^{-1}$ Mpc radius, and acts to normalise the power spectrum. For a given filter W , one could also associate a mass to a point,

$$M(\mathbf{r}; R) = V_R \int d^3\mathbf{x} W(\mathbf{r} - \mathbf{x}; R) \rho(\mathbf{x}), \quad (\text{C.10})$$

where $V_R = \int d^3\mathbf{x} [W(\mathbf{x}; R)/\max_{\mathbf{r}} W(\mathbf{r}; R)] = [\max_{\mathbf{r}} W(\mathbf{r}; R)]^{-1}$ is defined as the filter volume. By the linear relation between δ and M , it follows that $\sigma_M^2 = \langle \delta_R^2 \rangle$, where

$$\sigma_M^2 = \frac{\langle M - \langle M \rangle \rangle^2}{\langle M \rangle^2}. \quad (\text{C.11})$$

For any given point \mathbf{r} , one could consider $\delta_R(\mathbf{r})$ as a random variable parametrised by R or equivalently M , i.e. δ_R is described by a continuous stochastic process indexed by R^{-1} , with $R = \infty$ decreasing to $R = 0$. In the *Press–Schechter model*, a halo is identified when δ_R first exceeds the spherical-collapse critical over-density δ_c , which has the following conditional probability [325]:

$$\mathbb{P}(\delta_R > \delta_c | M) = \int_{\delta_c}^{\infty} d\eta \frac{1}{\sqrt{2\pi\sigma_M^2}} \exp\left(-\frac{1}{2} \frac{\eta^2}{\sigma_M^2}\right) = \frac{1}{2} \operatorname{erfc}\left(\frac{\nu}{2}\right), \quad (\text{C.12})$$

where erfc is the complementary error function, and

$$\nu = \frac{\delta_c}{\sigma_M} \quad (\text{C.13})$$

is the *peak height*, or *significance*, parameter. One can then translate this into a *halo mass function* representing the comoving halo number density per unit mass,

$$n_h(M) = -2 \frac{\bar{\rho}}{M} \frac{d}{dM} \mathbb{P}(\delta > \delta_c | M) = -\sqrt{\frac{2}{\pi}} \nu e^{-\nu^2/2} \frac{\bar{\rho}}{M^2} \frac{d \ln \sigma_M}{d \ln M}, \quad (\text{C.14})$$

where $\bar{\rho}$ is the background density (equivalently, the smoothed density in the limit $R \rightarrow \infty$). Since this function only depends on redshift, halo mass and generic cosmological parameters, it is also known as the *universal mass function* (UMF). The important factor of 2 here was originally added ad hoc in ref. [325] to ensure that the halo mass function is appropriately normalised, since the probability (eq. C.12) is not unity in the limit $\nu \rightarrow 0$, when all regions should be enclosed in haloes of arbitrarily small masses.¹ However, it can be argued as follows: δ_R may reach δ_c at some value of R_c^{-1} but then subsequently drops below, and this has a mirrored trajectory after R_c^{-1} that is reflected about δ_c ; physically, this corresponds to an under-dense region that is inside a collapsed region of a larger radius, and a halo should still be identified.² This is essentially the argument made in the excursion set formalism [327], which succeeds the Press–Schechter model.

Halo bias. Given that haloes originate from collapsed density peaks in a background of fluctuations, it is conceivable that halo clustering behaves differently in comparison with the

¹ However, it should be noted that equation (C.12) in itself is a bona fide distribution function where formally $\nu \in (-\infty, 0)$ should be allowed for a normal distribution; rather, it is the physical connection to halo formation that is under scrutiny.

² This argument based on a mirror trajectory is somewhat similar to the coupling technique used to prove e.g. the convergence theorem of Markov chains [326]; indeed, if the function W is chosen to be the sharp- k filter (i.e. a top-hat filter in k -space), then δ_R is Markovian [98].

underlying matter distribution. Therefore it is instructive to consider the ‘peak–background split’ [96, 97], where one writes the matter over-density field as

$$\delta = \delta_l + \delta_s, \quad (\text{C.15})$$

a sum of long-wavelength background fluctuations, denoted by the subscript ‘l’, and short-wavelength peak fluctuations denoted by the subscript ‘s’, which should not affect large-scale halo clustering. In regions where δ_l is high, the over-density barrier to halo formation is effectively lowered, and correspondingly the peak height ν is decreased,

$$\nu = \frac{\delta_c}{\sigma_M} \mapsto \check{\nu} = \frac{\delta_c - \delta_l}{\sigma_M}. \quad (\text{C.16})$$

Expanding the halo mass function in δ_l , one would then find

$$n_h(\check{\nu}) = \bar{n}_h(\nu) + \delta_l \left. \frac{\partial n_h(\check{\nu})}{\partial \delta_l} \right|_{\check{\nu}=\nu} + O(\delta_l^2), \quad (\text{C.17})$$

so that the local halo over-density field is

$$\delta_{h,l}^{(L)} = \frac{n_h(\check{\nu})}{\bar{n}_h(\nu)} - 1 = b_1^{(L)} \delta_l + O(\delta_l^2), \quad (\text{C.18})$$

where the linear *Lagrangian bias* (denoted with the superscript ‘(L)’) is given by

$$b_1^{(L)} = \frac{1}{\bar{n}_h} \frac{\partial \bar{n}_h}{\partial \nu} \frac{\partial \check{\nu}}{\partial \delta_l} = \frac{\nu^2 - 1}{\delta_c}. \quad (\text{C.19})$$

Assuming haloes are comoving with the underlying matter, one can relate the Lagrangian bias to the usual (Eulerian) bias by local halo number conservation [98],

$$1 + \delta_{h,l} = (1 + \delta_l) \left(1 + \delta_{h,l}^{(L)} \right), \quad (\text{C.20})$$

so to first order, the large-scale linear *Eulerian bias* for haloes is

$$b_1 = b_1^{(L)} + 1 = \frac{\nu^2 - 1}{\delta_c} + 1. \quad (\text{C.21})$$

Bibliography

- [1] **MS Wang**, WJ Percival, S Avila, R Crittenden and D Bianchi, 2019. ‘Cosmological inference from galaxy-clustering power spectrum: Gaussianization and covariance decomposition’, *Mon. Not. Roy. Astron. Soc.* **486** (1), 951 [1811.08155] [Paper I].
- [2] **MS Wang**, F Beutler and D Bacon, 2020. ‘Impact of relativistic effects on the primordial non-Gaussianity signature in the large-scale clustering of quasars’, *Mon. Not. Roy. Astron. Soc.* **499** (2), 2598 [2007.01802] [Paper II].
- [3] **MS Wang**, S Avila, D Bianchi, R Crittenden and WJ Percival, 2020. ‘Hybrid-basis inference for large-scale galaxy clustering: combining spherical and Cartesian Fourier analyses’, *J. Cosmol. Astropart. Phys.* **2020** (10), 022 [2007.14962] [Paper III].
- [4] CR Harris, KJ Millman, SJ van der Walt, R Gommers, P Virtanen, D Cournapeau et al., 2020. ‘Array programming with NumPy’, *Nature* **585**, 357 [2006.10256].
- [5] (SciPy 1.0 Contributors) P Virtanen, R Gommers, TE Oliphant, M Haberland, T Reddy, D Cournapeau et al., 2020. ‘SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python’, *Nat. Methods* **17**, 261 [1907.10121].
- [6] (Astropy Collaboration) TP Robitaille, EJ Tollerud, P Greenfield, M Droettboom, E Bray, T Aldcroft et al., 2013. ‘Astropy: A community python package for astronomy’, *Astron. Astrophys.* **558**, A33 [1307.6212].
- [7] (Astropy Collaboration) AM Price-Whelan, BM Sipőcz, HM Günther, PL Lim, SM Crawford, S Conseil et al., 2018. ‘The Astropy Project: Building an Open-science Project and Status of the v2.0 Core Package’, *Astron. J.* **156** (3), 123 [1801.02634].
- [8] N Hand, Y Feng, F Beutler, Y Li, C Modi, U Seljak et al., 2018. ‘nbodykit: An Open-source, Massively Parallel Toolkit for Large-scale Structure’, *Astron. J.* **156** (4), 160 [1712.05834].
- [9] JD Hunter, 2007. ‘Matplotlib: A 2D graphics environment’, *Comput. Sci. Eng.* **9** (3), 90.
- [10] PJE Peebles, *The Large-Scale Structure of the Universe*. Princeton University Press, 1980.
- [11] SDSS Collaboration, ‘Cosmology Results from eBOSS’, 2020. [🔗 sdss.org/science/cosmology-results-from-eboss](https://sdss.org/science/cosmology-results-from-eboss), retrieved 2020-09-14.
- [12] National Academy of Science Colloquia, ‘Physical Cosmology’, 1992. [🔗 nasonline.org/programs/nas-colloquia/completed_colloquia/physical-cosmology.html](https://nasonline.org/programs/nas-colloquia/completed_colloquia/physical-cosmology.html), retrieved 2020-09-14.

- [13] JB Birks (ed.), *Rutherford at Manchester*. Heywood and Company Ltd, 1962.
- [14] (Planck Collaboration) N Aghanim, Y Akrami, F Arroja, M Ashdown, J Aumont, C Baccigalupi et al., 2020. 'Planck 2018 results I. Overview and the cosmological legacy of Planck', *Astron. Astrophys.* **641**, A1 [1807.06205].
- [15] A Friedmann, 1922. 'Über die Krümmung des Raumes', *Z. Phys.* **10**, 377.
- [16] A Friedmann, 1924. 'Über die Möglichkeit einer Welt mit konstanter negativer Krümmung des Raumes', *Z. Phys.* **21**, 326.
- [17] AG Lemaître, 1931. 'A Homogeneous Universe of Constant Mass and Increasing Radius accounting for the Radial Velocity of Extra-galactic Nebulae', *Mon. Not. Roy. Astron. Soc.* **91** (5), 483.
- [18] HP Robertson, 1935. 'KINEMATICS AND WORLD-STRUCTURE', *Astrophys. J.* **82**, 284.
- [19] AG Walker, 1937. 'On Milne's Theory of World-Structure', *Proc. Lond. Math. Soc.* **s2-42** (1), 90.
- [20] AH Guth, 1981. 'The Inflationary Universe: A Possible Solution to the Horizon and Flatness Problems', *Phys. Rev. D* **23**, 347.
- [21] RA Alpher and R Herman, 1948. 'Evolution of the Universe', *Nature* **162** (4124), 774.
- [22] AA Penzias and RW Wilson, 1965. 'A MEASUREMENT OF EXCESS ANTENNA TEMPERATURE AT 4080 Mc/s.', *Astrophys. J.* **142**, 419.
- [23] (SDSS Collaboration) DJ Eisenstein, I Zehavi, DW Hogg, R Scoccimarro, MR Blanton, RC Nichol et al., 2005. 'Detection of the Baryon Acoustic Peak in the Large-Scale Correlation Function of SDSS Luminous Red Galaxies', *Astrophys. J.* **633** (2), 560 [astro-ph/0501171].
- [24] (2dFGRS Team) S Cole, WJ Percival, JA Peacock, P Norberg, CM Baugh, CS Frenk et al., 2005. 'The 2dF Galaxy Redshift Survey: power-spectrum analysis of the final data set and cosmological implications', *Mon. Not. Roy. Astron. Soc.* **362** (2), 505 [astro-ph/0501174].
- [25] (2dFGRS Team) WJ Percival, CM Baugh, J Bland-Hawthorn, T Bridges, R Cannon, S Cole et al., 2001. 'The 2dF Galaxy Redshift Survey: the power spectrum and the matter content of the Universe', *Mon. Not. Roy. Astron. Soc.* **327** (4), 1297 [astro-ph/0105252].
- [26] AG Riess, AV Filippenko, P Challis, A Clocchiatti, A Diercks, PM Garnavich et al., 1998. 'Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant', *Astrophys. J.* **116** (3), 1009 [astro-ph/9805201].
- [27] S Perlmutter, G Aldering, G Goldhaber, RA Knop, P Nugent, PG Castro et al., 1999. 'Measurements of Ω and Λ from 42 High-Redshift Supernovae', *Astrophys. J.* **517** (2), 565 [astro-ph/9812133].
- [28] RA Alpher, H Bethe and G Gamow, 1948. 'The Origin of Chemical Elements', *Phys. Rev.* **73**, 803.
- [29] G Steigman, 2007. 'Primordial Nucleosynthesis in the Precision Cosmology Era', *Annu. Rev. Nucl. Part. Sci.* **57** (1), 463 [0712.1100].
- [30] C Pitrou, A Coc, J-P Uzan and E Vangioni, 2018. 'Precision big bang nucleosynthesis with improved Helium-4 predictions', *Phys. Rep.* **754**, 1 [1801.08023].

- [31] BD Fields, KA Olive, T-H Yeh and C Young, 2020. ‘Big-Bang Nucleosynthesis after Planck’, *J. Cosmol. Astropart. Phys.* 2020 (03), 010 [1912.01132].
- [32] (2dFGRS Team) G Efstathiou, S Moody, JA Peacock, WJ Percival, C Baugh, J Bland-Hawthorn et al., 2002. ‘Evidence for a non-zero and a low matter density from a combined analysis of the 2dF Galaxy Redshift Survey and cosmic microwave background anisotropies’, *Mon. Not. Roy. Astron. Soc.* 330 (2), L29 [astro-ph/0109152].
- [33] (eBOSS Collaboration) S Alam, M Aubert, S Avila, C Balland, JE Bautista, MA Bershadsky et al., 2020. ‘The Completed SDSS-IV extended Baryon Oscillation Spectroscopic Survey: Cosmological Implications from Two Decades of Spectroscopic Surveys at the Apache Point Observatory’, *Phys. Rev. D* 103 (8), 083533 [2007.08991].
- [34] PJE Peebles and JT Yu, 1970. ‘PRIMEVAL ADIABATIC PERTURBATION IN AN EXPANDING UNIVERSE’, *Astrophys. J.* 162, 815.
- [35] A Albrecht, RA Battye and J Robinson, 1997. ‘The Case against Scaling Defect Models of Cosmic Structure Formation’, *Phys. Rev. Lett.* 79 (24), 4736 [astro-ph/9707129].
- [36] R Durrer, M Kunz and A Melchiorri, 2002. ‘Cosmic structure formation with topological defects’, *Phys. Rep.* 364 (1), 1 [astro-ph/0110348].
- [37] S Nadathur, WJ Percival, F Beutler and HA Winther, 2020. ‘Testing Low-Redshift Cosmic Acceleration with Large-Scale Structure’, *Phys. Rev. Lett.* 124 (22), 221301 [2001.11044].
- [38] L Verde, T Treu and AG Riess, 2019. ‘Tensions between the early and late universe’, *Nat. Astron.* 3 (10), 891 [1907.10625].
- [39] M Douspis, L Salvati and N Aghanim, 2018. ‘On the Tension between Large Scale Structures and Cosmic Microwave Background’, *Proc. Sci.* 335, 037 [1901.05289].
- [40] JL Bernal, L Verde and AG Riess, 2016. ‘The trouble with h_0 ’, *J. Cosmol. Astropart. Phys.* 2016 (10), 019 [1607.05617].
- [41] C Alcock and B Paczyński, 1979. ‘An evolution free test for non-zero cosmological constant’, *Nature* 281, 358.
- [42] N Kaiser, 1987. ‘Clustering in real space and in redshift space’, *Mon. Not. Roy. Astron. Soc.* 227 (1), 1.
- [43] WLW Sargent and EL Turner, 1977. ‘A STATISTICAL METHOD FOR DETERMINING THE COSMOLOGICAL DENSITY PARAMETER FROM THE REDSHIFTS OF A COMPLETE SAMPLE OF GALAXIES’, *Astrophys. J.* 212, L3.
- [44] B Jain and P Zhang, 2008. ‘Observational tests of modified gravity’, *Phys. Rev. D* 78, 063503 [0709.2375].
- [45] Y-S Song and WJ Percival, 2009. ‘Reconstructing the history of structure formation using redshift distortions’, *J. Cosmol. Astropart. Phys.* 2009 (10), 004 [0807.0810].
- [46] WJ Percival and M White, 2009. ‘Reconstructing the history of structure formation using redshift distortions’, *Mon. Not. Roy. Astron. Soc.* 393 (1), 297 [0808.0003].
- [47] JA Peacock, *Cosmological Physics*. Cambridge University Press, 1998.

- [48] F Bernardeau, S Colombi, E Gaztañaga and R Scoccimarro, 2002. ‘Large-scale structure of the Universe and cosmological perturbation theory’, *Phys. Rep.* **367** (1–3), 1 [astro-ph/0112551].
- [49] AF Heavens and AN Taylor, 1995. ‘A spherical harmonic analysis of redshift space’, *Mon. Not. Roy. Astron. Soc.* **275** (2), 483 [astro-ph/9409027].
- [50] L Isserlis, 1918. ‘On a Formula for the Product-Moment Coefficient of any Order of a Normal Frequency Distribution in any Number of Variables’, *Biometrika* **12** (1–2), 134.
- [51] GC Wick, 1950. ‘The Evaluation of the Collision Matrix’, *Phys. Rev.* **80**, 268.
- [52] H Wilbraham, 1848. ‘On a certain periodic function’, *Camb. Dublin Math. J.* **3**, 198.
- [53] E Hewitt and RE Hewitt, 1979. ‘The Gibbs-Wilbraham phenomenon: An episode in fourier analysis’, *Arch. Hist. Exact Sci.* **21**, 129.
- [54] G Rácz, I Szapudi, I Csabai and L Dobos, 2018. ‘Compactified cosmological simulations of the infinite universe’, *Mon. Not. Roy. Astron. Soc.* **477** (2), 1949 [1711.04959].
- [55] N Wiener, 1930. ‘Generalized harmonic analysis’, *Acta Math.* **55**, 117.
- [56] A Khintchine, 1934. ‘Korrelationstheorie der stationären stochastischen Prozesse’, *Math. Ann.* **109**, 604.
- [57] RJ Adler, *The Geometry of Random Fields*. Wiley, 1981.
- [58] HA Feldman, N Kaiser and JA Peacock, 1994. ‘POWER-SPECTRUM ANALYSIS OF THREE-DIMENSIONAL REDSHIFT SURVEYS’, *Astrophys. J.* **426**, 23 [astro-ph/9304022] [FKP].
- [59] P Coles and B Jones, 1991. ‘A lognormal model for the cosmological mass distribution’, *Mon. Not. Roy. Astron. Soc.* **248**, 1.
- [60] HJ Mo and SDM White, 1996. ‘An analytic model for the spatial clustering of dark matter haloes’, *Mon. Not. Roy. Astron. Soc.* **282** (2), 347 [astro-ph/9512127].
- [61] R Casas-Miranda, HJ Mo, RK Sheth and G Boerner, 2002. ‘On the distribution of haloes, galaxies and mass’, *Mon. Not. Roy. Astron. Soc.* **333** (4), 730 [astro-ph/0105008].
- [62] T Baldauf, U Seljak, RE Smith, N Hamaus and V Desjacques, 2013. ‘Halo stochasticity from exclusion and nonlinear clustering’, *Phys. Rev. D* **88**, 083507 [1305.2917].
- [63] AJS Hamilton, ‘Linear Redshift Distortions: A Review’, in *The Evolving Universe*, D Hamilton (ed.), pp. 185–275, Kluwer Academic 1998, DOI:10.1007/978-94-011-4960-0_17 [astro-ph/9708102].
- [64] S Zaroubi and Y Hoffman, 1996. ‘CLUSTERING IN REDSHIFT SPACE: LINEAR THEORY’, *Astrophys. J.* **462**, 25 [astro-ph/9311013].
- [65] S Raychaudhury and WC Saslaw, 1996. ‘The Observed Distribution Function of Peculiar Velocities of Galaxies’, *Astrophys. J.* **461**, 514 [astro-ph/9602001].
- [66] RK Sheth and A Diaferio, 2001. ‘Peculiar velocities of galaxies and clusters’, *Mon. Not. Roy. Astron. Soc.* **322** (4), 901 [astro-ph/0009166].

- [67] J Yoo and U Seljak, 2014. ‘Wide-angle effects in future galaxy surveys’, *Mon. Not. Roy. Astron. Soc.* **447** (2), 1789 [1308.1093].
- [68] JA Peacock and SJ Dodds, 1994. ‘Reconstructing the linear power spectrum of cosmological mass fluctuations’, *Mon. Not. Roy. Astron. Soc.* **267** (4), 1020 [astro-ph/9311057].
- [69] PJE Peebles, 1976. ‘A Cosmic Virial Theorem’, *Astrophys. Space Sci.* **45** (1), 3.
- [70] WE Ballinger, JA Peacock and AF Heavens, 1996. ‘Measuring the cosmological constant with redshift surveys’, *Mon. Not. Roy. Astron. Soc.* **282** (3), 877 [astro-ph/9605017].
- [71] H Magira, YP Jing and Y Suto, 2000. ‘Cosmological Redshift-Space Distortion on Clustering of High-Redshift Objects: Correction for Nonlinear Effects in the Power Spectrum and Tests with N -Body Simulations’, *Astrophys. J.* **528** (1), 30 [astro-ph/9907438].
- [72] C Park, MS Vogeley, MJ Geller and JP Huchra, 1994. ‘POWER SPECTRUM, CORRELATION FUNCTION, AND TESTS FOR LUMINOSITY BIAS IN THE CfA REDSHIFT SURVEY’, *Astrophys. J.* **431**, 569.
- [73] S Cole, KB Fisher and DH Weinberg, 1995. ‘Constraints on Omega from the *IRAS* redshift surveys’, *Mon. Not. Roy. Astron. Soc.* **275** (2), 515 [astro-ph/9412062].
- [74] R Scoccimarro, 2004. ‘Redshift-space distortions, pairwise velocities, and nonlinearities’, *Phys. Rev. D* **70**, 083007 [astro-ph/0407214].
- [75] A Taruya, T Nishimichi and S Saito, 2010. ‘Baryon acoustic oscillations in 2D: Modeling redshift-space power spectrum from perturbation theory’, *Phys. Rev. D* **82**, 063522 [1006.0699].
- [76] Y Zheng and Y-S Song, 2016. ‘Study on the mapping of dark matter clustering from real space to redshift space’, *J. Cosmol. Astropart. Phys.* **2016** (08), 050 [1603.0010].
- [77] Z Vlah and M White, 2019. ‘Exploring redshift-space distortions in large-scale structure’, *J. Cosmol. Astropart. Phys.* **2019** (03), 007 [1812.02775].
- [78] E Castorina and M White, 2018. ‘Beyond the plane-parallel approximation for redshift surveys’, *Mon. Not. Roy. Astron. Soc.* **476** (4), 4403 [1709.09730].
- [79] (BOSS Collaboration) F Beutler, H-J Seo, S Saito, C-H Chuang, AJ Cuesta, DJ Eisenstein et al., 2016. ‘The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: anisotropic galaxy clustering in Fourier space’, *Mon. Not. Roy. Astron. Soc.* **466** (2), 2242 [1607.03150].
- [80] (BOSS Collaboration) S Alam, M Ata, S Bailey, F Beutler, D Bizyaev, JA Blazek et al., 2017. ‘The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: cosmological analysis of the DR12 galaxy sample’, *Mon. Not. Roy. Astron. Soc.* **470** (3), 2617 [1607.03155].
- [81] (eBOSS Collaboration) R Neveux, E Burtin, A de Mattia, A Smith, AJ Ross, J Hou et al., 2020. ‘The Completed SDSS-IV extended Baryon Oscillation Spectroscopic Survey: BAO and RSD measurements from the anisotropic power spectrum of the quasar sample between redshift 0.8 and 2.2’, *Mon. Not. Roy. Astron. Soc.* **499** (1), 210 [2007.08999].

- [82] (eBOSS Collaboration) H Gil-Marín, JE Bautista, R Paviot, M Vargas-Magaña, S de la Torre, S Fromenteau et al., 2020. ‘The Completed SDSS-IV extended Baryon Oscillation Spectroscopic Survey: measurement of the BAO and growth rate of structure of the luminous red galaxy sample from the anisotropic power spectrum between redshifts 0.6 and 1.0’, *Mon. Not. Roy. Astron. Soc.* **498** (2), 2492 [2007.08994].
- [83] W Hu and N Sugiyama, 1995. ‘Toward understanding CMB anisotropies and their implications’, *Phys. Rev. D* (6), 2599 [astro-ph/9411008].
- [84] X Xu, AJ Cuesta, N Padmanabhan, DJ Eisenstein and CK McBride, 2013. ‘Measuring d_a and h at $z = 0.35$ from the SDSS DR7 LRGs using baryon acoustic oscillations’, *Mon. Not. Roy. Astron. Soc.* **431** (3), 2834 [1206.6732].
- [85] J Maldacena, 2003. ‘Non-gaussian features of primordial fluctuations in single field inflationary models’, *J. High Energy Phys.* 2003 (05), 013 [astro-ph/0210603].
- [86] M Alvarez, T Baldauf, JR Bond, N Dalal, R de Putter, O Doré et al., 2014. ‘Testing Inflation with Large Scale Structure: Connecting Hopes with Reality’, arXiv eprint [1412.4671].
- [87] M Biagetti, 2019. ‘The Hunt for Primordial Interactions in the Large-Scale Structures of the Universe’, *Galaxies* 7 (3), 71 [1906.12244].
- [88] (Planck Collaboration) Y Akrami, F Arroja, M Ashdown, J Aumont, C Baccigalupi, M Ballardini et al., 2020. ‘Planck 2018 results. IX. Constraints on primordial non-Gaussianity’, *Astron. Astrophys.* **641**, A9 [1905.05697].
- [89] T Falk, R Rangarajan and M Srednicki, 1993. ‘THE ANGULAR DEPENDENCE OF THE THREE-POINT CORRELATION FUNCTION OF THE COSMIC MICROWAVE BACKGROUND RADIATION AS PREDICTED BY INFLATIONARY COSMOLOGIES’, *Astrophys. J.* **403**, L1 [astro-ph/9208001].
- [90] P Creminelli and M Zaldarriaga, 2004. ‘A single-field consistency relation for the three-point function’, *J. Cosmol. Astropart. Phys.* 2004 (10), 006 [1211.0083].
- [91] X Chen, 2010. ‘Primordial Non-Gaussianities from Inflation Models’, *Adv. Astron.* 2010, 1 [1002.1416].
- [92] T Takahashi, 2014. ‘Primordial non-Gaussianity and the inflationary Universe’, *Prog. Theor. Exp. Phys.* 2014 (6), .
- [93] N Dalal, O Doré, D Huterer and A Shirokov, 2008. ‘Imprints of primordial non-Gaussianities on large-scale structure: Scale-dependent bias and abundance of virialized objects’, *Phys. Rev. D* **77**, 123514 [0710.4560].
- [94] S Matarrese and L Verde, 2008. ‘The Effect of Primordial Non-Gaussianity on Halo Bias’, *Astrophys. J.* **677** (2), L77 [0801.4826].
- [95] A Slosar, C Hirata, U Seljak, S Ho and N Padmanabhan, 2008. ‘Constraints on local primordial non-Gaussianity from large scale structure’, *J. Cosmol. Astropart. Phys.* 2008 (08), 031 [0805.3580].
- [96] N Kaiser, 1984. ‘ON THE SPATIAL CORRELATIONS OF ABELL CLUSTERS’, *Astrophys. J.* **284**, L9.

- [97] JM Bardeen, JR Bond, N Kaiser and AS Szalay, 1986. 'THE STATISTICS OF PEAKS OF GAUSSIAN RANDOM FIELDS', *Astrophys. J.* **304**, 15.
- [98] V Desjacques, D Jeong and F Schmidt, 2018. 'Large-scale galaxy bias', *Phys. Rep.* **733**, 1 [1611.09787].
- [99] F Schmidt and M Kamionkowski, 2010. 'Halo clustering with nonlocal non-Gaussianity', *Phys. Rev. D* **82** (10), 103002 [1008.0638].
- [100] T Giannantonio, C Porciani, J Carron, A Amara and A Pillepich, 2012. 'Constraining primordial non-Gaussianity with future galaxy surveys', *Mon. Not. Roy. Astron. Soc.* **422** (4), 2854 [1109.0958].
- [101] (*Planck* Collaboration) N Aghanim, Y Akrami, M Ashdown, J Aumont, C Baccigalupi, M Ballardini et al., 2020. 'Planck 2018 results. VI. Cosmological parameters', *Astron. Astrophys.* **641**, A6 [1807.06209].
- [102] A Barreira, G Cabass, F Schmidt, A Pillepich and D Nelson, 2020. 'Galaxy bias and primordial non-Gaussianity: insights from galaxy formation simulations with IllustrisTNG', *J. Cosmol. Astropart. Phys.* **2020** (12), 013 [2006.09368].
- [103] (eBOSS Collaboration) E Castorina, N Hand, U Seljak, F Beutler, C-H Chuang, C Zhao et al., 2019. 'Redshift-weighted constraints on primordial non-Gaussianity from the clustering of the eBOSS DR14 quasars in fourier space', *J. Cosmol. Astropart. Phys.* **2019** (9), 010 [1904.08859].
- [104] A Font-Ribera, P McDonald, N Mostek, BA Reid, H-J Seo and A Slosar, 2014. 'DESI and other Dark Energy experiments in the era of neutrino mass measurements', *J. Cosmol. Astropart. Phys.* **2014** (05), 023 [1308.4164].
- [105] L Amendola, S Appleby, A Avgoustidis, D Bacon, T Baker, M Baldi et al., 2018. 'Cosmology and fundamental physics with the Euclid satellite', *Living Rev. in Relativ.* **21** (2), 1 [1606.00180].
- [106] E-M Mueller, WJ Percival and R Ruggeri, 2018. 'Optimizing primordial non-Gaussianity measurements from galaxy surveys', *Mon. Not. Roy. Astron. Soc.* **485** (3), 4160 [1702.05088].
- [107] G Casella and R Berger, *Statistical Inference*. Brooks/Cole, 2001.
- [108] D Spiegelhalter, 'Part IB Statistics'. University of Cambridge Mathematical Tripos, 2017. [↗ statslab.cam.ac.uk/Dept/People/djsteaching/teaching17.html](https://statslab.cam.ac.uk/Dept/People/djsteaching/teaching17.html), retrieved 2020-10-16.
- [109] DS Sivia and J Skilling, *Data Analysis: A Bayesian Tutorial*. Oxford University Press, 2nd ed., 2006.
- [110] R Trotta, 2008. 'Bayes in the sky: Bayesian inference and model selection in cosmology', *Contemp. Phys.* **49** (2), 71 [0803.4089].
- [111] R Trotta, 2017. 'Bayesian Methods in Cosmology', arXiv eprint [1701.01467].
- [112] LE Padilla, LO Tellez, LA Escamilla and JA Vazquez, 2019. 'Cosmological parameter inference with Bayesian statistics', arXiv eprint [1903.11127].
- [113] ET Jaynes, 1957. 'Information Theory and Statistical Mechanics', *Phys. Rev.* **106**, 620.

- [114] Y Li, S Singh, B Yu, Y Feng and U Seljak, 2019. ‘Disconnected covariance of 2-point functions in large-scale structure’, *J. Cosmol. Astropart. Phys.* **01** (01), 016 [1811.05714].
- [115] D Wadekar and R Scoccimarro, 2020. ‘The galaxy power spectrum multipoles covariance in perturbation theory’, *Phys. Rev. D* **102** (12), 123517 [1910.02914].
- [116] A Taylor, B Joachimi and T Kitching, 2013. ‘Putting the precision in precision cosmology: How accurate should your data covariance matrix be?’, *Mon. Not. Roy. Astron. Soc.* **432** (3), 1928 [1212.4359].
- [117] J Hartlap, P Simon and P Schneider, 2007. ‘Why Your Model Parameter Confidences Might Be Too Optimistic. Unbiased Estimation of the Inverse Covariance Matrix’, *Astron. Astrophys.* **464** (1), 399 [astro-ph/0608064].
- [118] A Gupta and D Nagar, *Matrix Variate Distributions*. Chapman and Hall/CRC, 2000.
- [119] S Dodelson and MD Schneider, 2013. ‘The Effect of Covariance Estimator Error on Cosmological Parameter Constraints’, *Phys. Rev. D* **88**, 063537 [1304.2593].
- [120] (BOSS Collaboration) WJ Percival, AJ Ross, AG Sánchez, L Samushia, A Burden, R Crittenden et al., 2014. ‘The clustering of Galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: Including covariance matrix errors’, *Mon. Not. Roy. Astron. Soc.* **439** (3), 2531 [1312.4841].
- [121] E Sellentin and AF Heavens, 2016. ‘Parameter inference with estimated covariance matrices’, *Mon. Not. Roy. Astron. Soc.* **456** (1), L132 [1511.05969] [SH].
- [122] E Sellentin and AF Heavens, 2016. ‘Quantifying lost information due to covariance matrix estimation in parameter inference’, *Mon. Not. Roy. Astron. Soc.* **464** (4), 4658 [1609.00504].
- [123] J Yoo, AL Fitzpatrick and M Zaldarriaga, 2009. ‘New perspective on galaxy clustering as a cosmological probe: General relativistic effects’, *Phys. Rev. D* **80**, [0907.0707].
- [124] C Bonvin and R Durrer, 2011. ‘What galaxy surveys really measure’, *Phys. Rev. D* **84**, 063505 [1105.5280].
- [125] A Challinor and A Lewis, 2011. ‘Linear power spectrum of observed source number counts’, *Phys. Rev. D* **84**, 043516 [1105.5292].
- [126] M Bruni, R Crittenden, K Koyama, R Maartens, C Pitrou and D Wands, 2012. ‘Disentangling non-gaussianity, bias, and general relativistic effects in the galaxy distribution’, *Phys. Rev. D* **85**, 041301 [1106.3999].
- [127] D Jeong, F Schmidt and CM Hirata, 2012. ‘Large-scale clustering of galaxies in general relativity’, *Phys. Rev. D* **85**, 023504 [1107.5427].
- [128] S Camera, R Maartens and MG Santos, 2015. ‘Einstein’s legacy in galaxy surveys’, *Mon. Not. Roy. Astron. Soc.* **451** (1), L80 [1412.4781].
- [129] D Alonso, P Bull, PG Ferreira, R Maartens and MG Santos, 2015. ‘ULTRA-LARGE-SCALE COSMOLOGY IN NEXT-GENERATION EXPERIMENTS WITH SINGLE TRACERS’, *Astrophys. J.* **814** (2), 145 [1505.07596].

- [130] J Fonseca, S Camera, MG Santos and R Maartens, 2015. ‘HUNTING DOWN HORIZON-SCALE EFFECTS WITH MULTI-WAVELENGTH SURVEYS’, *Astrophys. J.* **812** (2), L22 [1507.04605].
- [131] YP Jing, 2005. ‘Correcting for the Alias Effect When Measuring the Power Spectrum Using a Fast Fourier Transform’, *Astrophys. J.* **620** (2), 559 [astro-ph/0409240].
- [132] E Sefusatti, M Crocce, R Scoccimarro and HMP Couchman, 2016. ‘Accurate estimators of correlation functions in Fourier space’, *Mon. Not. Roy. Astron. Soc.* **460** (4), 3624 [1512.07295].
- [133] WJ Percival, L Verde and JA Peacock, 2004. ‘Fourier analysis of luminosity-dependent galaxy clustering’, *Mon. Not. Roy. Astron. Soc.* **347** (2), 645 [astro-ph/0306511].
- [134] RE Smith and L Marian, 2016. ‘Towards optimal cluster power spectrum analysis’, *Mon. Not. Roy. Astron. Soc.* **457** (3), 2968 [1406.1800].
- [135] (BOSS Collaboration) AJ Ross, WJ Percival, AG Sánchez, L Samushia, S Ho, E Kazin et al., 2012. ‘The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: analysis of potential systematics’, *Mon. Not. Roy. Astron. Soc.* **424** (1), 564 [1203.6499].
- [136] (BOSS Collaboration) L Anderson, E Aubourg, S Bailey, F Beutler, V Bhardwaj, M Blanton et al., 2014. ‘The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: baryon acoustic oscillations in the Data Releases 10 and 11 Galaxy samples’, *Mon. Not. Roy. Astron. Soc.* **441** (1), 24 [1312.4877].
- [137] (BOSS Collaboration) L Samushia, BA Reid, M White, WJ Percival, AJ Cuesta, G-B Zhao et al., 2014. ‘The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: measuring growth rate and geometry with anisotropic clustering’, *Mon. Not. Roy. Astron. Soc.* **439** (4), 3504 [1312.4899].
- [138] (eBOSS Collaboration) J Hou, AG Sánchez, AJ Ross, A Smith, R Neveux, J Bautista et al., 2020. ‘The Completed SDSS-IV extended Baryon Oscillation Spectroscopic Survey: BAO and RSD measurements from anisotropic clustering analysis of the Quasar Sample in configuration space between redshift 0.8 and 2.2’, *Mon. Not. Roy. Astron. Soc.* **500** (1), 1201 [2007.08998].
- [139] K Yamamoto, M Nakamichi, A Kamino, BA Bassett and H Nishioka, 2006. ‘A Measurement of the Quadrupole Power Spectrum in the Clustering of the 2dF QSO Survey’, *Pub. Astron. Soc. Jap.* **58** (1), 93 [astro-ph/0505115].
- [140] R Scoccimarro, 2015. ‘Fast estimators for redshift-space clustering’, *Phys. Rev. D* **92**, 083532 [1506.02729].
- [141] P Reimberg, F Bernardeau and C Pitrou, 2016. ‘Redshift-space distortions with wide angular separations’, *J. Cosmol. Astropart. Phys.* **2016** (01), 048 [1506.06596].
- [142] F Beutler, E Castorina and P Zhang, 2019. ‘Interpreting measurements of the anisotropic galaxy power spectrum’, *J. Cosmol. Astropart. Phys.* **2019** (03), 040 [1810.05051].
- [143] I Szapudi, 2004. ‘Wide-Angle Redshift Distortions Revisited’, *Astrophys. J.* **614** (1), 51 [astro-ph/0404477].
- [144] P Pápai and I Szapudi, 2008. ‘Non-perturbative effects of geometry in wide-angle redshift distortions’, *Mon. Not. Roy. Astron. Soc.* **389** (1), 292 [0802.2940].

- [145] D Bianchi, H Gil-Marín, R Ruggeri and WJ Percival, 2015. ‘Measuring line-of-sight-dependent Fourier-space clustering using FFTs’, *Mon. Not. Roy. Astron. Soc.* **453** (1), L11 [1505.05341].
- [146] N Hand, Y Li, Z Slepian and U Seljak, 2017. ‘An optimal FFT-based anisotropic power spectrum estimator’, *J. Cosmol. Astropart. Phys.* **2017** (07), 002 [1704.02357].
- [147] (eBOSS Collaboration) A de Mattia, V Ruhlmann-Kleider, A Raichoor, AJ Ross, A Tamone, C Zhao et al., 2021. ‘The completed SDSS-IV extended baryon oscillation spectroscopic survey: measurement of the baryon acoustic oscillation and growth rate of structure of the emission line galaxy sample from the anisotropic power spectrum between redshift 0.6 and 1.1’, *Mon. Not. Roy. Astron. Soc.* **501** (4), 5616 [2007.09008].
- [148] (BOSS Collaboration) JN Grieb, AG Sánchez, S Salazar-Albornoz, R Scoccimarro, M Crocce, C Dalla Vecchia et al., 2017. ‘The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: cosmological implications of the Fourier space wedges of the final sample’, *Mon. Not. Roy. Astron. Soc.* **467** (2), 2085 [1607.03143].
- [149] MJ Wilson, JA Peacock, AN Taylor and S de la Torre, 2016. ‘Rapid modelling of the redshift-space power spectrum multipoles for a masked density field’, *Mon. Not. Roy. Astron. Soc.* **464** (3), 3121 [1511.07799].
- [150] JC Adams, 1878. ‘III. On the Expression of the Product of any two Legendre’s Coefficients by means of a Series of Legendre’s Coefficients’, *Proc. R. Soc. Lond.* **27**, 185.
- [151] SD Landy and AS Szalay, 1993. ‘BIAS AND VARIANCE OF ANGULAR CORRELATION FUNCTIONS’, *Astrophys. J.* **412**, 64.
- [152] JD Talman, 1978. ‘Numerical Fourier and Bessel Transforms in Logarithmic Variables’, *J. Comput. Phys.* **29** (1), 35.
- [153] AJS Hamilton, 2000. ‘Uncorrelated modes of the non-linear power spectrum’, *Mon. Not. Roy. Astron. Soc.* **312** (2), 257 [astro-ph/9905191].
- [154] JA Peacock and D Nicholson, 1991. ‘The large-scale clustering of radio galaxies’, *Mon. Not. Roy. Astron. Soc.* **253** (2), 307.
- [155] (BOSS Collaboration) F Beutler, S Saito, H-J Seo, J Brinkmann, KS Dawson, DJ Eisenstein et al., 2014. ‘The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: testing gravity with redshift space distortions using the power spectrum multipoles’, *Mon. Not. Roy. Astron. Soc.* **443** (2), 1065 [1312.4611].
- [156] A de Mattia and V Ruhlmann-Kleider, 2019. ‘Integral constraints in spectroscopic surveys’, *J. Cosmol. Astropart. Phys.* **2019** (08), 036 [1904.08851].
- [157] (BOSS Collaboration) M Manera, R Scoccimarro, WJ Percival, L Samushia, CK McBride, AJ Ross et al., 2013. ‘The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: a large sample of mock galaxy catalogues’, *Mon. Not. Roy. Astron. Soc.* **428** (2), 1036 [1203.6609].
- [158] (BOSS Collaboration) F-S Kitaura, S Rodríguez-Torres, C-H Chuang, C Zhao, F Prada, H Gil-Marín et al., 2016. ‘The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: mock galaxy catalogues for the BOSS Final Data Release’, *Mon. Not. Roy. Astron. Soc.* **456** (4), 4156 [1509.06400].

- [159] (DES Collaboration) S Avila, M Crocce, AJ Ross, J García-Bellido, WJ Percival, N Banik et al., 2018. ‘Dark Energy Survey Year-1 results: galaxy mock catalogues for BAO’, *Mon. Not. Roy. Astron. Soc.* **479** (1), 94 [1712.06232].
- [160] M Lippich, AG Sánchez, M Colavincenzo, E Sefusatti, P Monaco, L Blot et al., 2019. ‘Comparing approximate methods for mock catalogues and covariance matrices – I. Correlation function’, *Mon. Not. Roy. Astron. Soc.* **482** (2), 1786 [1806.09477].
- [161] L Blot, M Crocce, E Sefusatti, M Lippich, AG Sánchez, M Colavincenzo et al., 2019. ‘Comparing approximate methods for mock catalogues and covariance matrices II: power spectrum multipoles’, *Mon. Not. Roy. Astron. Soc.* **485** (2), 2806 [1806.09497].
- [162] M Colavincenzo, E Sefusatti, P Monaco, L Blot, M Crocce, M Lippich et al., 2019. ‘Comparing approximate methods for mock catalogues and covariance matrices – III: bispectrum’, *Mon. Not. Roy. Astron. Soc.* **482** (4), 4883 [1806.09499].
- [163] AC Pope and I Szapudi, 2008. ‘Shrinkage estimation of the power spectrum covariance matrix’, *Mon. Not. Roy. Astron. Soc.* **389** (2), 766 [0711.2509].
- [164] CG Kaufman, MJ Schervish and DW Nychka, 2008. ‘Covariance Tapering for Likelihood-Based Estimation in Large Spatial Data Sets’, *J. Am. Stat. Assoc.* **103** (484), 1545.
- [165] DJ Paz and AG Sánchez, 2015. ‘Improving the precision matrix for precision cosmology’, *Mon. Not. Roy. Astron. Soc.* **454** (4), 4326 [1508.03162].
- [166] E Sellentin, AH Jaffe and AF Heavens, 2017. ‘On the use of the Edgeworth expansion in cosmology I: how to foresee and evade its pitfalls’, preprint [1709.03452].
- [167] T Eifler, P Schneider and J Hartlap, 2009. ‘Dependence of cosmic shear covariances on cosmology - Impact on parameter estimation’, *Astron. Astrophys.* **502** (3), 721 [0810.4254].
- [168] A Hall and A Mead, 2017. ‘Perturbative Gaussianizing transforms for cosmological fields’, *Mon. Not. Roy. Astron. Soc.* **473** (3), 3190 [1709.03924].
- [169] P Schneider and J Hartlap, 2009. ‘Constrained correlation functions’, *Astron. Astrophys.* **504** (3), 705 [0905.0577].
- [170] D Keitel and P Schneider, 2011. ‘Constrained probability distributions of correlation functions’, *Astron. Astrophys.* **534**, A76 [1105.3672].
- [171] P Wilking and P Schneider, 2013. ‘A quasi-Gaussian approximation for the probability distribution of correlation functions’, *Astron. Astrophys.* **556**, A70 [1304.4781].
- [172] L Sun, Q Wang and H Zhan, 2013. ‘LIKELIHOOD OF THE POWER SPECTRUM IN COSMOLOGICAL PARAMETER ESTIMATION’, *Astrophys. J.* **777** (1), 75 [1306.5064].
- [173] B Kalus, WJ Percival and L Samushia, 2016. ‘Cosmological parameter inference from galaxy clustering: the effect of the posterior distribution of the power spectrum’, *Mon. Not. Roy. Astron. Soc.* **455** (3), 2573 [1504.03979].
- [174] (WMAP Science Team) L Verde, HV Peiris, DN Spergel, MR Nolta, CL Bennett, M Halpern et al., 2003. ‘First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Parameter Estimation Methodology’, *Astrophys. J. Suppl.* **148** (1), 195 [astro-ph/0302218].

- [175] WJ Percival and ML Brown, 2006. ‘Likelihood techniques for the combined analysis of cmb temperature and polarization power spectra’, *Mon. Not. Roy. Astron. Soc.* **372** (3), 1104 [astro-ph/0604547].
- [176] S Smith, A Challinor and G Rocha, 2006. ‘What can be learned from the lensed cosmic microwave background B-mode polarization power spectrum?’, *Phys. Rev. D* **73**, 023517 [astro-ph/0511703].
- [177] S Hamimeche and A Lewis, 2008. ‘Likelihood analysis of CMB temperature and polarization power spectra’, *Phys. Rev. D* **77**, [0801.0554].
- [178] MF Neuts, *Matrix-Geometric Solutions in Stochastic Models: an Algorithmic Approach*. Dover Publications Inc., 1981.
- [179] A Golubev, 2016. ‘Applications and implications of the exponentially modified gamma distribution as a model for time variabilities related to cell proliferation and gene expression’, *J. Theor. Biol.* **393**, 203.
- [180] V Laparra, G Camps-Valls and J Malo, 2011. ‘Iterative Gaussianization: From ICA to Random Rotations’, *IEEE Trans. Neural Netw.* **22** (4), 537 [1602.00229].
- [181] B Joachimi and AN Taylor, 2011. ‘Forecasts of non-Gaussian parameter spaces using Box–Cox transformations’, *Mon. Not. Roy. Astron. Soc.* **416** (2), 1010 [1103.3370].
- [182] RL Schuhmann, B Joachimi and HV Peiris, 2016. ‘Gaussianization for fast and accurate inference from cosmological data’, *Mon. Not. Roy. Astron. Soc.* **459** (2), 1916 [1510.00019].
- [183] GEP Box and DR Cox, 1964. ‘An Analysis of Transformations’, *J. Royal Stat. Soc.* **26** (2), 211.
- [184] T Burić and N Elezović, 2012. ‘New asymptotic expansions of the quotient of gamma functions’, *Integral Transforms Spec. Funct.* **23** (5), 355.
- [185] JR Bond, AH Jaffe and L Knox, 1998. ‘Estimating the power spectrum of the cosmic microwave background’, *Phys. Rev. D* **57**, 2117 [astro-ph/9708203].
- [186] DJ Eisenstein and W Hu, 1998. ‘Baryonic Features in the Matter Transfer Function’, *Astrophys. J.* **496** (2), 605 [astro-ph/9709112].
- [187] AR Duffy, 2014. ‘Probing the nature of dark energy through galaxy redshift surveys with radio telescopes’, *Ann. Phys.* **526** (7–8), 283 [1405.7465].
- [188] Y Shao and M Zhou, 2010. ‘A characterization of multivariate normality through univariate projections’, *J. Multivar. Anal.* **101**, 2637.
- [189] R D’Agostino and ES Pearson, 1973. ‘Tests for Departure from Normality. Empirical Results for the Distributions of b_2 and $\sqrt{b_1}$ ’, *Biometrika* **60** (3), 613.
- [190] DW Hogg and D Foreman-Mackey, 2018. ‘Data Analysis Recipes: Using Markov Chain Monte Carlo’, *Astrophys. J. Suppl.* **236** (1), 11.
- [191] JO Berger, *Statistical Decision Theory and Bayesian Analysis*. Springer-Verlag, 1985.
- [192] MB Wilk and R Gnanadesikan, 1968. ‘Probability Plotting Methods for the Analysis of Data’, *Biometrika* **55** (1), 1.
- [193] S Kullback and RA Leibler, 1951. ‘On Information and Sufficiency’, *Ann. Math. Stat.* **22** (1), 79.

- [194] D Bertacca, R Maartens, A Raccanelli and C Clarkson, 2012. ‘Beyond the plane-parallel and Newtonian approach: wide-angle redshift distortions and convergence in general relativity’, *J. Cosmol. Astropart. Phys.* **2012** (10), 025 [1205.5221].
- [195] A Raccanelli, F Montanari, D Bertacca, O Doré and R Durrer, 2016. ‘Cosmological measurements with general relativistic galaxy correlations’, *J. Cosmol. Astropart. Phys.* **2016** (05), 009 [1505.06179].
- [196] A Raccanelli, D Bertacca, D Jeong, MC Neyrinck and AS Szalay, 2018. ‘Doppler term in the galaxy two-point correlation function: wide-angle, velocity, Doppler lensing and cosmic acceleration effects’, *Phys. Dark Universe* **19**, 109 [1602.03186].
- [197] CS Lorenz, D Alonso and PG Ferreira, 2018. ‘Impact of relativistic effects on cosmological parameter estimation’, *Phys. Rev. D* **97**, 023537 [1710.02477].
- [198] L Lombriser, J Yoo and K Koyama, 2013. ‘Relativistic effects in galaxy clustering in a parametrized post-Friedmann universe’, *Phys. Rev. D* **87**, 104019 [1301.3132].
- [199] C Bonvin, 2014. ‘Isolating relativistic effects in large-scale structure’, *Class. Quantum Gravity* **31** (23), 234002 [1409.2224].
- [200] C Bonvin, F Oliveira Franco and P Fleury, 2020. ‘A null test of the equivalence principle using relativistic effects in galaxy surveys’, *J. Cosmol. Astropart. Phys.* **2020** (08), 004 [2004.06457].
- [201] F Beutler and E Di Dio, 2020. ‘Modeling relativistic contributions to the halo power spectrum dipole’, *J. Cosmol. Astropart. Phys.* **2020** (07), 048 [2004.08014].
- [202] (eBOSS Collaboration) N Palanque-Delabrouille, Ch Magneville, Ch Yèche, I Pâris, P Petitjean, E Burtin et al., 2016. ‘The extended Baryon Oscillation Spectroscopic Survey: Variability selection and quasar luminosity function’, *Astron. Astrophys.* **587**, A41 [1509.05607] [PD2016].
- [203] T Namikawa, T Okamura and A Taruya, 2011. ‘Magnification effect on the detection of primordial non-Gaussianity from photometric surveys’, *Phys. Rev. D* **83**, 123514 [1103.1118].
- [204] A Raccanelli, D Bertacca, R Maartens, C Clarkson and O Doré, 2016. ‘Lensing and time-delay contributions to galaxy correlations’, *Gen. Relativ. Gravit.* **48** (7), 84 [1311.6813].
- [205] G Jelic-Cizmek, F Lepori, C Bonvin and R Durrer, 2020. ‘On the importance of lensing for galaxy clustering in photometric and spectroscopic surveys’, *J. Cosmol. Astropart. Phys.* **2021** (04), 055 [2004.12981].
- [206] L Pozzetti, CM Hirata, JE Geach, A Cimatti, C Baugh, O Cucciati et al., 2016. ‘Modelling the number density of $h\alpha$ emitters for future spectroscopic near-IR space missions’, *Astron. Astrophys.* **590**, A3 [1603.01453].
- [207] BJ Boyle, T Shanks, SM Croom, RJ Smith, L Miller, N Loaring et al., 2000. ‘The 2dF QSO Redshift Survey – I. the optical luminosity function of quasi-stellar objects’, *Mon. Not. Roy. Astron. Soc.* **317** (4), 1014 [astro-ph/0005368].
- [208] (SDSS Collaboration) GT Richards, MA Strauss, X Fan, PB Hall, S Jester, DP Schneider et al., 2006. ‘The Sloan Digital Sky Survey Quasar Survey: Quasar Luminosity Function from Data Release 3’, *Astrophys. J.* **131** (6), 2766 [astro-ph/0601434].

- [209] M Karamanis and F Beutler, 2020. ‘Ensemble Slice Sampling’, preprint [2002.06212].
- [210] DM Caditz, 2017. ‘Recalculating the quasar luminosity function of the extended Baryon Oscillation Spectroscopic Survey’, *Astron. Astrophys.* **608**, A64.
- [211] (*Planck* Collaboration) PAR Ade, N Aghanim, M Arnaud, M Ashdown, J Aumont, C Baccigalupi et al., 2016. ‘*Planck* 2015 results - XIII. Cosmological parameters’, *Astron. Astrophys.* **594**, A13 [1502.01589].
- [212] DM Caditz, 2018. ‘Non-parametric Estimation of Evolution of the Quasar Luminosity Function’, *Astrophys. J.* **869** (2), 96.
- [213] (DESI Collaboration) A Aghamousa, J Aguilar, S Ahlen, S Alam, LE Allen, C Allende Prieto et al., 2016. ‘The DESI Experiment Part I: Science, Targeting, and Survey Design’, arXiv eprint [1611.00036].
- [214] M White, AD Myers, NP Ross, DJ Schlegel, JF Hennawi, Y Shen et al., 2012. ‘The clustering of intermediate-redshift quasars as measured by the Baryon Oscillation Spectroscopic Survey’, *Mon. Not. Roy. Astron. Soc.* **424** (2), 933 [1203.5306].
- [215] Y Shen, CK McBride, M White, Z Zheng, AD Myers, H Guo et al., 2013. ‘CROSS-CORRELATION OF SDSS DR₇ QUASARS AND DR₁₀ BOSS GALAXIES: THE WEAK LUMINOSITY DEPENDENCE OF QUASAR CLUSTERING AT $z \sim 0.5$ ’, *Astrophys. J.* **778** (2), 98 [1212.4526].
- [216] AG Krolewski and DJ Eisenstein, 2015. ‘MEASURING THE LUMINOSITY AND VIRIAL BLACK HOLE MASS DEPENDENCE OF QUASAR–GALAXY CLUSTERING AT $z \sim 0.8$ ’, *Astrophys. J.* **803** (1), 4 [1501.03898].
- [217] Y Shen, 2009. ‘SUPERMASSIVE BLACK HOLES IN THE HIERARCHICAL UNIVERSE: A GENERAL FRAMEWORK AND OBSERVATIONAL TESTS’, *Astrophys. J.* **704** (1), 89 [0903.4492].
- [218] DJ Croton, 2009. ‘A simple model to link the properties of quasars to the properties of dark matter haloes out to high redshift’, *Mon. Not. Roy. Astron. Soc.* **394** (2), 1109 [0901.4104].
- [219] C Conroy and M White, 2012. ‘A SIMPLE MODEL FOR QUASAR DEMOGRAPHICS’, *Astrophys. J.* **762** (2), 70 [1208.3198].
- [220] JD Timlin, NP Ross, GT Richards, AD Myers, A Pellegrino, FE Bauer et al., 2018. ‘The Clustering of High-redshift ($2.9 \leq z \leq 5.1$) Quasars in SDSS Stripe 82’, *Astrophys. J.* **859** (1), 20 [1712.03128].
- [221] G Kulkarni, G Worseck and JF Hennawi, 2019. ‘Evolution of the AGN UV luminosity function from redshift 7.5’, *Mon. Not. Roy. Astron. Soc.* **488** (1), 1035 [1807.09774].
- [222] (*Euclid* Consortium) R Laureijs, J Amiaux, S Arduini, JL Auguères, J Brinchmann, R Cole et al., 2011. ‘Euclid Definition Study Report’, arXiv eprint [1110.3193].
- [223] U Seljak, 2009. ‘Extracting primordial non-Gaussianity without cosmic variance’, *Phys. Rev. Lett.* **102** (2), 021302 [0807.1770].
- [224] P McDonald and U Seljak, 2009. ‘How to evade the sample variance limit on measurements of redshift-space distortions’, *J. Cosmol. Astropart. Phys.* **2009** (10), 007 [0810.0323].

- [225] N Grimm, F Scaccabarozzi, J Yoo, SG Biern and J-O Gong, 2020. ‘Galaxy power spectrum in general relativity’, *J. Cosmol. Astropart. Phys.* **2020** (11), 064 [2005.06484].
- [226] AS Szalay, T Matsubara and SD Landy, 1998. ‘Redshift-Space Distortions of the Correlation Function in Wide-Angle Galaxy Surveys’, *Astrophys. J.* **498** (1), L1 [astro-ph/9712007].
- [227] J Yoo and U Seljak, 2015. ‘Wide angle effects in future galaxy surveys’, *Mon. Not. Roy. Astron. Soc.* **447** (2), 1789 [1308.1093].
- [228] K Fisher, O Lahav, Y Hoffman, D Lynden-Bell and S Zaroubi, 1995. ‘Wiener reconstruction of density, velocity, and potential fields from all-sky galaxy redshift surveys’, *Mon. Not. Roy. Astron. Soc.* **272** (4), 885 [astro-ph/9406009].
- [229] J Yoo and V Desjacques, 2013. ‘All-sky analysis of the general relativistic galaxy power spectrum’, *Phys. Rev. D* **88**, 023502 [1301.4501].
- [230] A Raccanelli, L Samushia and WJ Percival, 2010. ‘Simulating redshift-space distortions for galaxy pairs with wide angular separation’, *Mon. Not. Roy. Astron. Soc.* **409** (4), 1525 [1006.1652].
- [231] (eBOSS Collaboration) G-B Zhao, Y Wang, S Saito, H Gil-Marín, WJ Percival, D Wang et al., 2018. ‘The clustering of the SDSS-IV extended Baryon Oscillation Spectroscopic Survey DR14 quasar sample: a tomographic measurement of cosmic structure growth and expansion rate based on optimal redshift weights’, *Mon. Not. Roy. Astron. Soc.* **482** (3), 3497 [1801.0304].
- [232] (eBOSS Collaboration) J Zheng, G-B Zhao, J Li, Y Wang, C-H Chuang, F-S Kitaura et al., 2019. ‘The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: a tomographic measurement of structure growth and expansion rate from anisotropic galaxy clustering in Fourier space’, *Mon. Not. Roy. Astron. Soc.* **484** (1), 442 [1806.01920].
- [233] J Binney and T Quinn, 1991. ‘Gaussian random fields in spherical coordinates’, *Mon. Not. Roy. Astron. Soc.* **249**, 678.
- [234] H Tadros, WE Ballinger, AN Taylor, AF Heavens, G Efstathiou, W Saunders et al., 1999. ‘Spherical harmonic analysis of the PSCz galaxy catalogue: redshift distortions and the real-space power spectrum’, *Mon. Not. Roy. Astron. Soc.* **305** (3), 527 [astro-ph/9901351].
- [235] IM Schmoldt, V Saar, P Saha, E Branchini, GP Efstathiou, CS Frenk et al., 1999. ‘On Density and Velocity Fields and β from the *IRAS* PSCz Survey’, *Astrophys. J.* **118** (3), 1146 [astro-ph/9906035].
- [236] AN Taylor, WE Ballinger, AF Heavens and H Tadros, 2001. ‘Redshift-space distortions in the PSCz galaxy catalogue’, *Mon. Not. Roy. Astron. Soc.* **327** (3), 689.
- [237] P Erdoğdu, JP Huchra, O Lahav, M Colless, RM Cutri, E Falco et al., 2006. ‘The dipole anisotropy of the 2 Micron All-Sky Redshift Survey’, *Mon. Not. Roy. Astron. Soc.* **368** (4), 1515 [astro-ph/0507166].
- [238] P Erdoğdu, O Lahav, JP Huchra, M Colless, RM Cutri, E Falco et al., 2006. ‘Reconstructed density and velocity fields from the 2MASS Redshift Survey’, *Mon. Not. Roy. Astron. Soc.* **373** (1), 45 [astro-ph/0610005].

- [239] (2dFGRS Team) WJ Percival, D Burkey, A Heavens, A Taylor, S Cole, JA Peacock et al., 2004. ‘The 2dF Galaxy Redshift Survey: spherical harmonics analysis of fluctuations in the final catalogue’, *Mon. Not. Roy. Astron. Soc.* **353** (4), 1201 [astro-ph/0406513].
- [240] B Leistedt, A Rassat, A Réfrégier and J-L Starck, 2012. ‘3DEX: a code for fast spherical Fourier-Bessel decomposition of 3D surveys’, *Astron. Astrophys.* **540**, A60 [1111.3591].
- [241] A Rassat and A Refregier, 2012. ‘3D spherical analysis of baryon acoustic oscillations’, *Astron. Astrophys.* **540**, A115 [1112.3100].
- [242] C Shapiro, RG Crittenden and WJ Percival, 2012. ‘The complementarity of redshift-space distortions and the integrated Sachs–Wolfe effect: a 3D spherical analysis’, *Mon. Not. Roy. Astron. Soc.* **422** (3), 2341 [1109.1981].
- [243] G Pratten and D Munshi, 2013. ‘Effects of linear redshift space distortions and perturbation theory on BAOs: a 3D spherical analysis’, *Mon. Not. Roy. Astron. Soc.* **436** (4), 3792 [1301.3673].
- [244] A Nicola, A Refregier, A Amara and A Paranjape, 2014. ‘Three-dimensional spherical analyses of cosmological spectroscopic surveys’, *Phys. Rev. D* **90**, 063515 [1405.3660].
- [245] F Lanusse, A Rassat and J-L Starck, 2015. ‘3D galaxy clustering with future wide-field surveys: Advantages of a spherical Fourier-Bessel analysis’, *Astron. Astrophys.* **578**, A10 [1406.5989].
- [246] M Tegmark, 1997. ‘How to measure CMB power spectra without losing information’, *Phys. Rev. D* **55**, 5895 [astro-ph/9611174].
- [247] JR Bond, AH Jaffe and L Knox, 2000. ‘Radical Compression of Cosmic Microwave Background Data’, *Astrophys. J.* **533** (1), 19 [astro-ph/9808264].
- [248] BD Wandelt, E Hivon and KM Górski, 2001. ‘Cosmic microwave background anisotropy power spectrum statistics for high precision cosmology’, *Phys. Rev. D* **64**, [astro-ph/0008111].
- [249] G Efstathiou, 2004. ‘Myths and truths concerning estimation of power spectra: the case for a hybrid estimator’, *Mon. Not. Roy. Astron. Soc.* **349** (2), 603 [astro-ph/0307515].
- [250] (WMAP Science Team) G Hinshaw, MR Nolta, CL Bennett, R Bean, O Dore, MR Greason et al., 2007. ‘Three-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Temperature Analysis’, *Astrophys. J. Suppl.* **170** (2), 288 [astro-ph/0603451].
- [251] (Planck Collaboration) PAR Ade, N Aghanim, C Armitage-Caplan, M Arnaud, M Ashdown, F Atrio-Barandela et al., 2014. ‘Planck 2013 results. XV. CMB power spectra and likelihood’, *Astron. Astrophys.* **571**, A15 [1303.5075].
- [252] (Planck Collaboration) N Aghanim, M Arnaud, M Ashdown, J Aumont, C Baccigalupi, AJ Banday et al., 2016. ‘Planck 2015 results. XI. CMB power spectra, likelihoods, and robustness of parameters’, *Astron. Astrophys.* **594**, A11 [1507.02704].
- [253] (Planck Collaboration) N Aghanim, Y Akrami, M Ashdown, J Aumont, C Baccigalupi, M Ballardini et al., 2020. ‘Planck 2018 results. V. CMB power spectra and likelihoods’, *Astron. Astrophys.* **641**, A5 [1907.12875].

- [254] AM Webster, SL Bridle, MP Hobson, AN Lasenby, O Lahav and G Rocha, 1998. ‘Joint Estimation of Cosmological Parameters from Cosmic Microwave Background and *IRAS* Data’, *Astrophys. J.* **509** (2), L65 [astro-ph/9802109].
- [255] A Heavens, 2003. ‘3D weak lensing’, *Mon. Not. Roy. Astron. Soc.* **343** (4), 1327 [astro-ph/0304151].
- [256] PG Castro, AF Heavens and TD Kitching, 2005. ‘Weak lensing analysis in three dimensions’, *Phys. Rev. D* **72**, 023516 [astro-ph/0503479].
- [257] TD Kitching, AF Heavens, J Alsing, T Erben, C Heymans, H Hildebrandt et al., 2014. ‘3D cosmic shear: cosmology from CFHTLenS’, *Mon. Not. Roy. Astron. Soc.* **442** (2), 1326 [1401.6842].
- [258] J McMahon, 1894. ‘On the Roots of the Bessel and Certain Related Functions’, *Ann. Math.* **9** (1/6), 23.
- [259] J Jewell, S Levin and CH Anderson, 2004. ‘Application of Monte Carlo Algorithms to the Bayesian Analysis of the Cosmic Microwave Background’, *Astrophys. J.* **609** (1), 1 [astro-ph/0209560].
- [260] BD Wandelt, DL Larson and A Lakshminarayanan, 2004. ‘Global, exact cosmic microwave background data analysis using Gibbs sampling’, *Phys. Rev. D* **70**, 1 [astro-ph/0310080].
- [261] HK Eriksen, IJ O’Dwyer, JB Jewell, BD Wandelt, DL Larson, KM Gorski et al., 2004. ‘Power Spectrum Estimation from High-Resolution Maps by Gibbs Sampling’, *Astrophys. J. Suppl.* **155** (2), 227 [astro-ph/0407028].
- [262] RA Wooding, 1956. ‘The Multivariate Distribution of Complex Normal Variables’, *Biometrika* **43** (1/2), 212.
- [263] NR Goodman, 1963. ‘Statistical Analysis Based on a Certain Multivariate Complex Gaussian Distribution (An Introduction)’, *Ann. Math. Stat.* **34** (1), 152.
- [264] O Lahav, SL Bridle, MP Hobson, AN Lasenby and L Sodre, 2000. ‘Bayesian ‘hyper-parameters’ approach to joint estimation: the Hubble constant from CMB measurements’, *Mon. Not. Roy. Astron. Soc.* **315** (4), L45 [astro-ph/9912105].
- [265] MP Hobson, SL Bridle and O Lahav, 2002. ‘Combining cosmological data sets: hyperparameters and Bayesian evidence’, *Mon. Not. Roy. Astron. Soc.* **335** (2), 377 [astro-ph/0203259].
- [266] Y-Z Ma and A Berendsen, 2014. ‘How to combine correlated data sets—A Bayesian hyperparameter matrix method’, *Astron. Comput.* **5**, 45 [1309.3271].
- [267] M Tegmark, AN Taylor and AF Heavens, 1997. ‘Karhunen-Loève Eigenvalue Problems in Cosmology: How Should We Tackle Large Data Sets?’, *Astrophys. J.* **480** (1), 22 [astro-ph/9603021].
- [268] S Yahia-Cherif, A Blanchard, S Camera, S Ilić, K Markovič, A Pourtsidou et al., 2020. ‘Validating the Fisher approach for stage IV spectroscopic surveys’, *Astron. Astrophys.* [2007.01812].
- [269] GH Golub and CF Van Loan, *Matrix Computations*. The John Hopkins University Press, 1996.

- [270] A Lewis and S Bridle, 2002. ‘Cosmological parameters from CMB and other data: A Monte Carlo approach’, *Phys. Rev. D* **66**, 103511 [astro-ph/0205436].
- [271] M Crocce, S Pueblas and R Scoccimarro, 2006. ‘Transients from initial conditions in cosmological simulations’, *Mon. Not. Roy. Astron. Soc.* **373** (1), 369 [astro-ph/0606505].
- [272] R Scoccimarro, L Hui, M Manera and KC Chan, 2012. ‘Large-scale bias and efficient generation of initial conditions for nonlocal primordial non-Gaussianity’, *Phys. Rev. D* **85**, 083002 [1108.5512].
- [273] V Springel, 2005. ‘The cosmological simulation code GADGET-2’, *Mon. Not. Roy. Astron. Soc.* **364** (4), 1105 [astro-ph/0505010].
- [274] SR Knollmann and A Knebe, 2009. ‘AHF: AMIGA’S HALO FINDER’, *Astrophys. J. Suppl.* **182** (2), 608 [0904.3662].
- [275] KM Gorski, E Hivon, AJ Banday, BD Wandelt, FK Hansen, M Reinecke et al., 2005. ‘HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere’, *Astrophys. J.* **622** (2), 759 [astro-ph/0409513].
- [276] L Samushia, 2019. ‘Proper Fourier decomposition formalism for cosmological fields in spherical shells’, preprint [1906.05866].
- [277] EP Hubble, 1926. ‘EXTRA-GALACTIC NEBULAE’, *Astrophys. J.* **64**, 321.
- [278] EP Hubble, 1929. ‘A RELATION BETWEEN DISTANCE AND RADIAL VELOCITY AMONG EXTRA-GALACTIC NEBULAE’, *Proc. Natl. Acad. Sci. U.S.A.* **15** (3), 168.
- [279] G Lemaître, 1927. ‘Un univers homogène de masse constante et de rayon croissant rendant compte de la vitesse radiale des nébuleuses extra-galactiques’, *Ann. Soc. Sci. Bruxelles* **47**, 49.
- [280] H Nussbaumer and L Bieri, *Discovering the Expanding Universe*. Cambridge University Press, 2009.
- [281] CD Shane and CA Wirtanen, *Pub. Lick Obs.*, vol. 22 (Part 1). University of California, 1967.
- [282] F Zwicky and collaborators, *Catalogue of Galaxies and of Clusters of Galaxies*. California Institute of Technology, 1968.
- [283] PJE Peebles and MG Hauser, 1974. ‘Statistical Analysis of Catalogs of Extragalactic Objects. III. The Shane–Wirtanen and Zwicky Catalogs’, *Astrophys. J. Suppl.* **28**, 19.
- [284] M Davis, J Huchra, DW Latham and J Tonry, 1982. ‘A SURVEY OF GALAXY REDSHIFTS. II. THE LARGE SCALE SPACE DISTRIBUTION’, *Astrophys. J.* **253**, 423.
- [285] G Efstathiou, N Kaiser, W Saunders, A Lawrence, M Rowan-Robinson, RS Ellis et al., 1990. ‘Large-scale clustering of *IRAS* galaxies’, *Mon. Not. Roy. Astron. Soc.* **247**, 10P.
- [286] (6dFGRS Team) DH Jones, MA Read, W Saunders, M Colless, T Jarrett, QA Parker et al., 2009. ‘The 6dF Galaxy Survey: final redshift release (DR3) and southern large-scale structures’, *Mon. Not. Roy. Astron. Soc.* **399** (2), 683 [0903.5451].
- [287] (DESI Collaboration) M Levi, LE Allen, A Raichoor, C Baltay, S BenZvi, F Beutler et al., 2019. ‘The Dark Energy Spectroscopic Instrument (DESI)’, *Bull. Am. Astron. Soc.* **51**, 57 [1907.10688].

- [288] E Di Valentino, A Melchiorri and J Silk, 2015. ‘Beyond six parameters: Extending Λ cdm’, *Phys. Rev. D* **92** (12), 121302 [1507.06646].
- [289] S Vagnozzi, 2020. ‘New physics in light of the h_0 tension: An alternative view’, *Phys. Rev. D* **102** (2), 023518 [1907.07569].
- [290] L Knox and M Millea, 2020. ‘Hubble constant hunter’s guide’, *Phys. Rev. D* **101** (4), 043533 [1908.03663].
- [291] T Tröster, M Asgari, C Blake, M Cataneo, C Heymans, H Hildebrandt et al., 2020. ‘KiDS-1000 Cosmology: constraints beyond flat Λ cdm’, *A&A* **649**, A88 [2010.16416].
- [292] M Masjedi, DW Hogg, RJ Cool, DJ Eisenstein, MR Blanton, I Zehavi et al., 2006. ‘Very small scale clustering and merger rate of luminous red galaxies’, *Astrophys. J.* **644** (1), 54 [astro-ph/0512166].
- [293] C Hahn, R Scoccimarro, MR Blanton, JL Tinker and S Rodríguez-Torres, 2017. ‘The effect of fibre collisions on the galaxy power spectrum multipoles’, *Mon. Not. Roy. Astron. Soc.* **467** (2), 1940 [1609.01714].
- [294] D Bianchi and WJ Percival, 2017. ‘Unbiased clustering estimation in the presence of missing observations’, *Mon. Not. Roy. Astron. Soc.* **472** (1), 1106 [1703.02070].
- [295] P McDonald, 2009. ‘Gravitational redshift and other redshift-space distortions of the imaginary part of the power spectrum’, *J. Cosmol. Astropart. Phys.* **2009** (11), 026 [0907.5220].
- [296] LR Abramo and KE Leonard, 2013. ‘Why multitracer surveys beat cosmic variance’, *Mon. Not. Roy. Astron. Soc.* **432** (1), 318 [1302.5444].
- [297] Y Wang and G-B Zhao, 2020. ‘A brief review on cosmological analysis of galaxy surveys with multiple tracers’, *Res. Astron. Astrophys.* **20** (10), 158 [2009.03862].
- [298] G-B Zhao, Y Wang, A Taruya, W Zhang, H Gil-Marín, A de Mattia et al., 2020. ‘The completed SDSS-IV extended Baryon Oscillation Spectroscopic Survey: a multi-tracer analysis in Fourier space for measuring the cosmic structure growth and expansion rate’, *Mon. Not. Roy. Astron. Soc.* **504** (1), 33 [2007.09011].
- [299] O Doré, J Bock, M Ashby, P Capak, A Cooray, R de Putter et al., 2014. ‘Cosmology with the SPHEREX All-Sky Spectral Survey’, arXiv eprint [1412.4872].
- [300] J Wishart, 1928. ‘The Generalised Product Moment Distribution in Samples from a Normal Multivariate Population’, *Biometrika* **20A** (1–2), 32.
- [301] EP Wigner, 1955. ‘Characteristic Vectors of Bordered Matrices With Infinite Dimensions’, *Ann. Math.* **62** (3), 548.
- [302] EP Wigner, 1955. ‘Characteristics Vectors of Bordered Matrices with Infinite Dimensions II’, *Ann. Math.* **65** (2), 548.
- [303] EP Wigner, 1958. ‘On the Distribution of the Roots of Certain Symmetric Matrices’, *Ann. Math.* **67** (2), 325.
- [304] A Aazami and R Easther, 2006. ‘Cosmology from random multifield potentials’, *J. Cosmol. Astropart. Phys.* **2006** (03), 013 [hep-th/0512050].

- [305] R Easther and L McAllister, 2006. ‘Random matrices and the spectrum of N-flation’, *J. Cosmol. Astropart. Phys.* 2006 (05), 018 [hep-th/0512102].
- [306] MD Marsh, L McAllister, E Pajer and T Wrase, 2013. ‘Charting an inflationary landscape with random matrix theory’, *J. Cosmol. Astropart. Phys.* 2013 (11), 040 [1307.3559].
- [307] GB Rybicki and WH Press, 1992. ‘Interpolation, Realization, and Reconstruction of Noisy, Irregularly Sampled Data’, *Astrophys. J.* 398, 169.
- [308] B Kalus, WJ Percival, DJ Bacon and L Samushia, 2016. ‘Unbiased contaminant removal for 3D galaxy power spectrum measurements’, *Mon. Not. Roy. Astron. Soc.* 463 (1), 467 [1607.02417].
- [309] B Kalus, WJ Percival, DJ Bacon, EM Mueller, L Samushia, L Verde et al., 2018. ‘A map-based method for eliminating systematic modes from galaxy clustering power spectra with application to boss’, *Mon. Not. Roy. Astron. Soc.* 482 (1), 453 [1806.02789].
- [310] JP Bouchaud and M Potters, ‘Financial applications of random matrix theory: a short review’, in *The Oxford Handbook of Random Matrix Theory*, G Akemann, J Baik and P Di Francesco (eds.), Oxford University Press 2015, DOI:10.1093/oxfordhb/9780198744191.013.40.
- [311] The Nobel Foundation, ‘Press release: The Nobel Prize in Physics 2019’, 2019.
[🔗 nobelprize.org/prizes/physics/2019/press-release/](https://nobelprize.org/prizes/physics/2019/press-release/), retrieved 2020-11-30.
- [312] J Fergusson, ‘Part III Cosmology’. University of Cambridge Mathematical Tripos, 2016.
[🔗 damtp.cam.ac.uk/user/examples/3R2La.pdf](https://damtp.cam.ac.uk/user/examples/3R2La.pdf), retrieved 2020-10-04.
- [313] H Reall, ‘Part 3 General Relativity’. University of Cambridge Mathematical Tripos, 2012.
[🔗 damtp.cam.ac.uk/user/hsr1000/lecturenotes_2012.pdf](https://damtp.cam.ac.uk/user/hsr1000/lecturenotes_2012.pdf), retrieved 2020-10-04.
- [314] CW Misner, KS Thorne and JA Wheeler, *Gravitation*. W.H. Freeman, 1973.
- [315] JM Bardeen, 1980. ‘Gauge-invariant cosmological perturbations’, *Phys. Rev. D* 22, 1882.
- [316] JM Stewart, 1990. ‘Perturbations of Friedmann–Robertson–Walker cosmological models’, *Class. Quantum Gravity* 7 (7), 1169.
- [317] VF Mukhanov, HA Feldman and RH Brandenberger, 1992. ‘Theory of cosmological perturbations’, *Phys. Rep.* 215, 203.
- [318] DJ Heath, 1977. ‘The growth of density perturbations in zero pressure Friedmann–Lemaître universes’, *Mon. Not. Roy. Astron. Soc.* 179 (3), 351.
- [319] U Seljak and M Zaldarriaga, 1996. ‘A LINE-OF-SIGHT APPROACH TO COSMIC MICROWAVE BACKGROUND ANISOTROPIES’, *Astrophys. J.* 469, 437 [astro-ph/9603033].
- [320] A Lewis, A Challinor and A Lasenby, 2000. ‘Efficient computation of CMB anisotropies in closed FRW models’, *Astrophys. J.* 538, 473 [astro-ph/9911177].
- [321] J Lesgourgues, 2011. ‘The Cosmic Linear Anisotropy Solving System (CLASS) I: Overview’, arXiv eprint [1104.2932].
- [322] JE Gunn and JR Gott, III, 1972. ‘ON THE INFALL OF MATTER INTO CLUSTERS OF GALAXIES AND SOME EFFECTS ON THEIR EVOLUTION’, *Astrophys. J.* 176, 1.
- [323] AR Zentner, 2007. ‘The Excursion Set Theory of Halo Mass Functions, Halo Clustering, and Halo Growth’, *Int. J. Mod. Phys. D* 16 (05), 763 [astro-ph/0611454].

- [324] WJ Percival, 2005. 'Cosmological structure formation in a homogeneous dark energy background', *Astron. Astrophys.* **443** (3), 819 [astro-ph/0508156].
- [325] WH Press and P Schechter, 1974. 'FORMATION OF GALAXIES AND CLUSTERS OF GALAXIES BY SELF-SIMILAR GRAVITATIONAL CONDENSATION', *Astrophys. J.* **187**, 425.
- [326] GR Grimmett and DR Stirzaker, *Probability and Random Processes*. Oxford University Press, 2001.
- [327] JR Bond, S Cole, G Efstathiou and N Kaiser, 1991. 'EXCURSION SET MASS FUNCTIONS FOR HIERARCHICAL GAUSSIAN FLUCTUATIONS', *Astrophys. J.* **379**, 440.

“ *I have sometimes been wildly, despairingly, acutely miserable, racked with sorrow; but through it all I still know quite certainly that just to be alive is a grand thing.* ”

— Dame Agatha Christie, Lady Mallowan, *An Autobiography*

FURTHER ACKNOWLEDGEMENTS

There is a choice between long closing credits typically seen in a film and total omission; it is clear which choice has been made here. The reader may be surprised to find some names here and others missing—apologies to those in the latter group, and to those in the former group, let it be known that there is at least one kind thing you have done for which I am grateful.

Fellow students (2017). Natalie H, Michael K-A, Sam L, Andrius T, Maria V, Sam Y.

LSS group. Many past and present members including but not limited to Julian B, Jamie D-Mc, Minas K, Eva(-Maria) M, Sesh N, Obinna U, Hans W...

Office 3.14. Many past and present members including but not limited to Steve C, Max F-M, Lewis H, Connor Mc, Chris P, Sam P, Bill W...

Lockdown 'locals.' Guilleherme B, Jacob E, Laura I, Chakkrit K, Robyn M, Philipp S, Jinglan Zh...

Squash players. Steve B (special thanks as my coach), Gareth B, Chris F, Max F, Mike F, Toby F-H, Evan G, Henry L, Jon L.

Friday pub 'regulars.' This group also includes any missing names.



Certificate of Ethics Review

Project Title:	Theoretical Considerations of Large-Scale Structure Analyses
User ID:	875922
Name:	Shengbo Wang
Application Date:	06/10/2017 17:43:40

You must download your certificate, print a copy and keep it as a record of this review.

It is your responsibility to adhere to the University Ethics Policy and any Department/School or professional guidelines in the conduct of your study including relevant guidelines regarding health and safety of researchers and University Health and Safety Policy.

It is also your responsibility to follow University guidance on Data Protection Policy:

- General guidance for all data protection issues
- University Data Protection Policy

You are reminded that as a University of Portsmouth Researcher you are bound by the UKRIO Code of Practice for Research; any breach of this code could lead to action being taken following the University's Procedure for the Investigation of Allegations of Misconduct in Research.

Any changes in the answers to the questions reflecting the design, management or conduct of the research over the course of the project must be notified to the Faculty Ethics Committee. Any changes that affect the answers given in the questionnaire, not reported to the Faculty Ethics Committee, will invalidate this certificate.

This ethical review should not be used to infer any comment on the academic merits or methodology of the project. If you have not already done so, you are advised to develop a clear protocol/proposal and ensure that it is independently reviewed by peers or others of appropriate standing. A favourable ethical opinion should not be perceived as permission to proceed with the research; there might be other matters of governance which require further consideration including the agreement of any organisation hosting the research.

Governance Checklist

A1-BriefDescriptionOfProject: The Project will be working on theoretical aspects related to clustering measurements to be made with galaxy surveys. Two areas that I will consider are methods to ameliorate problems caused by missing galaxies within these surveys, and the likelihood analysis that turns the measurements into cosmological model constraints. These methods will be applied to both current and future galaxy surveys.

A2-Faculty: Technology

A3-VoluntarilyReferToFEC: No

A5-AlreadyExternallyReviewed: No
B1-HumanParticipants: No
HumanParticipantsDefinition
B2-HumanParticipantsConfirmation: Yes
C6-SafetyRisksBeyondAssessment: No
SafetyRisksBeyondAssessmentWarning
D2-PhysicalEcologicalDamage: No
D4-HistoricalOrCulturalDamage: No
E1-ContentiousOrIllegal: No
E2-SociallySensitiveIssues: No
F1-InvolvesAnimals: No
F2-HarmfulToThirdParties: No
G1-ConfirmReadEthicsPolicy: Confirmed
G2-ConfirmReadUKRIOCodeOfPractice: Confirmed
G3-ConfirmReadConcordatToSupportResearchIntegrity: Confirmed
G4-ConfirmedCorrectInformation: Confirmed

FORM UPR16

Research Ethics Review Checklist



Please include this completed form as an appendix to your thesis (see the Research Degrees Operational Handbook for more information)

Postgraduate Research Student (PGRS) Information		Student ID:	875922
PGRS Name:	Shengbo (Mike) Wang		
Department:	ICG	First Supervisor:	Prof Robert Crittenden
Start Date: (or progression date for Prof Doc students)	2 Oct 2017		
Study Mode and Route:	Part-time <input type="checkbox"/>	MPhil <input type="checkbox"/>	MD <input type="checkbox"/>
	Full-time <input checked="" type="checkbox"/>	PhD <input checked="" type="checkbox"/>	Professional Doctorate <input type="checkbox"/>

Title of Thesis:	Theoretical Considerations in Large-Scale Structure Analyses
Thesis Word Count: (excluding ancillary data)	42852

If you are unsure about any of the following, please contact the local representative on your Faculty Ethics Committee for advice. Please note that it is your responsibility to follow the University's Ethics Policy and any relevant University, academic or professional guidelines in the conduct of your study

Although the Ethics Committee may have given your study a favourable opinion, the final responsibility for the ethical conduct of this work lies with the researcher(s).

UKRIO Finished Research Checklist:
(If you would like to know more about the checklist, please see your Faculty or Departmental Ethics Committee rep or see the online version of the full checklist at: <http://www.ukrio.org/what-we-do/code-of-practice-for-research/>)

a) Have all of your research and findings been reported accurately, honestly and within a reasonable time frame?	YES <input checked="" type="checkbox"/> NO <input type="checkbox"/>
b) Have all contributions to knowledge been acknowledged?	YES <input checked="" type="checkbox"/> NO <input type="checkbox"/>
c) Have you complied with all agreements relating to intellectual property, publication and authorship?	YES <input checked="" type="checkbox"/> NO <input type="checkbox"/>
d) Has your research data been retained in a secure and accessible form and will it remain so for the required duration?	YES <input checked="" type="checkbox"/> NO <input type="checkbox"/>
e) Does your research comply with all legal, ethical, and contractual requirements?	YES <input checked="" type="checkbox"/> NO <input type="checkbox"/>

Candidate Statement:

I have considered the ethical dimensions of the above named research project, and have successfully obtained the necessary ethical approval(s)

Ethical review number(s) from Faculty Ethics Committee (or from NRES/SCREC): A416-53F4-4B64-8B71-CB54-EB17-4A93-4D3D

If you have *not* submitted your work for ethical review, and/or you have answered 'No' to one or more of questions a) to e), please explain below why this is so:

Signed (PGRS):	<i>Mike Shengbo Wang</i>	Date:	20 May 2021
-----------------------	--------------------------	--------------	-------------