

UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

Non-Gaussian Signatures of Primordial Interactions During Inflation in the Large-Scale Structure

Permalink

<https://escholarship.org/uc/item/4gc9z869>

ISBN

9798171610159

Author

Han, Jiashu

Publication Date

2026-10-23

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA SAN DIEGO

Non-Gaussian Signatures of Primordial Interactions During Inflation in the Large-Scale
Structure

A Dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy

in

Physics

by

Jiashu Han

Committee in charge:

Professor Daniel Green, Chair
Professor Alexey Arefiev
Professor Kam Arnold
Professor Raphael Flauger

Copyright

Jiashu Han, 2026

All rights reserved.

The Dissertation of Jiashu Han is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2026

DEDICATION

To my partner, for bringing light to life's most challenging days.

To my parents, for their unwavering love and support throughout this journey.

To my late grandparents, whose kindness continues to inspire me.

To my advisor, for his invaluable guidance, patience, and mentorship.

To my friends, for filling these seven years with laughter, joy, and lasting memories.

EPIGRAPH

The past is never dead. It's not even past.

William Faulkner

The book of nature is written in the language of mathematics.

Galileo Galilei

Statistical laws are not necessarily used as a result of our ignorance. Statistical laws can reflect how things really are. There are matters that can only be treated statistically.

James Clerk Maxwell

TABLE OF CONTENTS

Dissertation Approval Page	iii
Dedication	iv
Epigraph	v
Table of Contents	vi
List of Figures	viii
List of Tables	xii
Acknowledgements	xiii
Vita	xv
Abstract of the Dissertation	xvi
Introduction	1
Chapter 1 From Inflationary Interactions to Galaxy Clustering	6
1.1 Inflationary Origin of Primordial Fluctuations	7
1.1.1 Primordial Perturbations in Single-Field Inflation	7
1.1.2 Interactions and Extra Fields	10
1.2 Structure Formation	14
1.2.1 Linear Growth and Transfer Functions	14
1.2.2 Galaxy Biasing	16
1.2.3 Observational and Astrophysical Systematics	19
Chapter 2 Light Fields during Inflation from BOSS and Future Galaxy Surveys	22
2.1 Synopsis	22
2.2 Introduction	23
2.3 Light Fields and Galaxies	27
2.3.1 Galaxy Power Spectrum	27
2.3.2 Scale-Dependent Bias	33
2.4 Information Content of the Galaxy Power Spectrum	37
2.4.1 Fisher Information Matrix	38
2.4.2 Information and Survey Design	40
2.5 Constraints and Data Analysis for General Scaling Exponents	53
2.5.1 Search Strategy	53
2.5.2 BOSS DR12 Analysis	56
2.6 Multiple Tracers and the Dependence on Astrophysics	61
2.6.1 Optimizing Spectroscopic Multi-Tracer Analyses	62
2.6.2 Multi-Tracer LSST and Comparison to SPHEREx	66

2.7	Conclusions and Outlook	71
Chapter 3	Extending the Cosmological Collider: New Scaling Regimes and Constraints from BOSS	74
3.1	Synopsis	74
3.2	Introduction	75
3.3	Cosmological-Collider Models	82
3.3.1	Frequency Range of Two-Field Models	82
3.3.2	The Amplitude of Primordial Fluctuations	85
3.3.3	Non-Gaussianity and the Squeezed Limit	92
3.4	Oscillations in the Galaxy Power Spectrum	96
3.4.1	Oscillations in the Scale-Dependent Bias	97
3.4.2	Modeling the Galaxy Power Spectrum	103
3.4.3	Forecasts for Current and Future Surveys	110
3.5	Constraints from BOSS Galaxy Clustering	122
3.5.1	Analysis Pipeline for BOSS DR12 Data	123
3.5.2	Bounding Non-Gaussian Oscillations	126
3.6	Conclusions and Outlook	132
Chapter 4	Conclusions	136
Appendix A	Appendix for Chapter 2	140
A.1	Forecasting Details	140
A.1.1	Galaxy Power Spectrum Model	140
A.1.2	Survey Specifications	142
A.2	Measuring the Scaling Behavior	147
Appendix B	Appendix for Chapter 3	152
B.1	Details for the Fisher Forecasts	152
B.2	Details for the BOSS Analysis	161
Bibliography		165

LIST OF FIGURES

Figure 1.1.	The matter power spectrum $P_{\text{mm}}(k)$ at $z = 0$	15
Figure 1.2.	Modulation of the small-scale variance in matter density fluctuations by a long-wavelength mode.	18
Figure 2.1.	Illustration of the effect of the scale-dependent bias b_{NG}^{Δ} from a non-zero non-Gaussian amplitude f_{NL}^{Δ} on the linear galaxy power spectrum.	24
Figure 2.2.	Contributions to the galaxy power spectrum $P_{gg}^{(i)}(k)$ at $z = 0$ from the nonlinear, bias and non-Gaussian terms.	31
Figure 2.3.	Forecasted constraints on f_{NL}^{Δ} for general scaling $\Delta \in [0, 2]$ of the scale-dependent bias, for current and future surveys, compared to the Planck bispectrum constraints for local and equilateral non-Gaussianity.	41
Figure 2.4.	Dependence of $\sigma(f_{\text{NL}}^{\Delta})$ on the shot noise for the billion-object survey for four representative values of Δ	43
Figure 2.5.	The information density with respect to $\log k_{\text{min}}$ in the billion-object survey, with fixed Λ CDM cosmology, and the same k_{min} , $\Delta k = 0.003 h \text{Mpc}^{-1}$ and $k_{\text{max}} = 0.2 h \text{Mpc}^{-1}$ in each redshift bin.	45
Figure 2.6.	Dependence of $\sigma(f_{\text{NL}}^{\Delta})$ on the maximum wavenumber k_{max}	46
Figure 2.7.	Dependence of $\sigma(f_{\text{NL}}^{\Delta})$ on the fiducial value of the linear bias for both the single-tracer (solid) and double-tracer (dashed) versions of the billion-object survey for $\Delta = 0$ and $\Delta = 1$	48
Figure 2.8.	Dependence of $\sigma(f_{\text{NL}}^{\Delta=1})$ on the sky fraction f_{sky} and the maximum redshift z_{max} at fixed observational effort for the double-tracer billion-object survey.	50
Figure 2.9.	Comparison of $\sigma(f_{\text{NL}}^{\Delta})$ for a range of photometric redshift errors $\sigma_{z0} \in [0.001, 0.2]$ for the billion-object survey.	52
Figure 2.10.	Correlation matrix for measurements of galaxy power spectra with different values of the non-Gaussian exponent Δ	55
Figure 2.11.	Comparison of the measured and derived values for $2\sigma(f_{\text{NL}}^{\Delta})$ for the BOSS DR12 dataset.	59
Figure 2.12.	Double-tracer forecast for $\sigma(f_{\text{NL}}^{\Delta=0})$ for the billion-object survey and relative comparison to a single-tracer forecast, $\sigma(f_{\text{NL}}^{\Delta})_{\text{double}}/\sigma(f_{\text{NL}}^{\Delta})_{\text{single}}$, with fixed Λ CDM parameters.	64

Figure 2.13.	Forecasted constraints on f_{NL}^{Δ} for different values of the parameter $p^{(i)}$ in the universality relation of b_{ϕ} for the double-tracer configuration of the billion-object survey and fixed Λ CDM parameters.	65
Figure 2.14.	Forecast for $\sigma(f_{\text{NL}}^{\Delta})$ for different Gaussian priors $\sigma(p)$ on the parameter p in the universality relation for the billion-object survey (at fixed Λ CDM parameters and with fiducial value $\bar{f}_{\text{NL}}^{\Delta} = 1$).	67
Figure 2.15.	Comparison of the constraining power of single- and double-tracer versions of LSST and the billion-object survey for $\Delta = 0$ and 2.	67
Figure 2.16.	Constraining power of SPHEREx with its fiducial biases $b_1^{(i)}$ and number densities $\bar{n}_g^{(i)}$ compared to a survey with either of these observational design factors replaced by LSST-like numbers for fixed and marginalized Gaussian bias parameters.	70
Figure 3.1.	Illustration of the effect of oscillatory primordial non-Gaussianity on the linear galaxy power spectrum through the scale-dependent bias for different scaling dimensions $\Delta = \alpha + i\nu$	79
Figure 3.2.	Comparison of the numerical and analytical solutions for the dimensionless mode function $\tilde{\sigma}(u)$ and its respective power spectrum $H^{-2}k^3 P_{\sigma}(u)$ as a function of $u = k\tau$ for the case of no mixing and equal masses of the two fields.	89
Figure 3.3.	Numerical solutions for the mode functions $\tilde{\sigma}^{(\pm)}(u)$ for six different cases of non-zero mixing.	91
Figure 3.4.	Suppression of the overall amplitude of the oscillatory PNG signal in the galaxy power spectrum as captured by $ \Sigma_M^2(\alpha, \nu) $	104
Figure 3.5.	Illustration of the effect of the oscillatory scale-dependent bias on the linear BOSS galaxy power spectrum at $z = 0.35$, with $A_{\text{sin}} = 0$	107
Figure 3.6.	Forecasted constraints on the amplitude of primordial non-Gaussianity in terms of its 95% exclusion limit as a function of the oscillation frequency ν for four values of the scaling exponent α	114
Figure 3.7.	Ratio of the 95% exclusion limits for the oscillatory PNG amplitude $A f_{\text{NL}}^{\alpha, \nu}$ with $\nu \neq 0$ and the (absolute value of the) power-law PNG amplitude $ f_{\text{NL}}^{\alpha, \nu} $ with $\nu = 0$ as a function of frequency ν for a range of scaling exponents α	115
Figure 3.8.	Forecasted posterior distributions of $(A_{\text{cos}}, A_{\text{sin}})$ for BOSS, a scaling exponent of $\alpha = 0$ and different frequencies ν	118

Figure 3.9.	Dependence of the 95% exclusion limit of the PNG amplitude $f_{\text{NL}}^{\alpha, \nu}$ on the maximum wavenumber k_{max} of the billion-object survey for different values of the scaling exponent α	120
Figure 3.10.	Constraints on the non-Gaussian amplitude(s) inferred from BOSS DR12 for a scaling exponent of $\alpha = 0$ as a function of frequency ν	127
Figure 3.11.	Exclusion limits at 95% c.l. on the amplitude of the oscillatory primordial non-Gaussianity, $Af_{\text{NL}}^{\alpha, \nu}$, inferred from the BOSS DR12 dataset as a function of the oscillatory frequency ν of the non-Gaussian signature.	128
Figure 3.12.	Exclusion limits at 95% c.l. on the amplitude of the oscillatory primordial non-Gaussianity, $Af_{\text{NL}}^{\alpha, \nu}$, from BOSS DR12 relative to the 95% limit on the absolute value of the amplitude of the non-oscillatory, simple power-law non-Gaussianity, $ f_{\text{NL}}^{\alpha, \nu} $	130
Figure A.1.	Correlation matrix for measurements of galaxy power spectra with different values of the non-Gaussian scaling exponent Δ for the billion-object survey and four biasing models.	148
Figure A.2.	Forecasts for $\sigma(f_{\text{NL}}^{\Delta})$ and $\sigma(\Delta)$ as a function of Δ if the scaling exponent is either fixed or varied for the billion-object survey.	149
Figure B.1.	Separation of zeroes δk of the isolated oscillatory signal (or, alternatively, its effective linear frequency ω_{lin}) for five values of the logarithmic frequency ν as a function of wavenumber k	153
Figure B.2.	Forecasted posterior distributions for $(A_{\cos}, A_{\sin})$ and the derived limits on $Af_{\text{NL}}^{\alpha, \nu}$ for different fiducial phases ϕ_{fid} if the phase parametrization had been used to model the logarithmic oscillations.	155
Figure B.3.	Effect of marginalizing over the cosmological parameters and imposing a CMB prior on the forecasted 95% exclusion limits on the non-Gaussian amplitude $Af_{\text{NL}}^{\alpha, \nu}$ for four values of α as a function of ν	156
Figure B.4.	Relative improvement of the 95% upper limit on $Af_{\text{NL}}^{\alpha, \nu}$ for the billion-object survey when increasing the maximum wavenumber from a fiducial and conservative choice of k_{max} to a more optimistic choice for four scaling exponents.	158
Figure B.5.	Forecasted constraints on the amplitude of primordial non-Gaussianity in terms of its 95% exclusion limit as a function of the oscillation frequency ν for four additional values of the scaling exponent α	159

Figure B.6.	Posterior distributions of $(A_{\cos}, A_{\sin})$ for a scaling exponent of $\alpha = 0$ and the 19 frequencies ν sampled in the analysis of BOSS DR12 data.	162
Figure B.7.	Exclusion limits at 95% c.l. on the amplitude of the oscillatory primordial non-Gaussianity, $A f_{\text{NL}}^{\alpha, \nu}$, inferred from the BOSS DR12 dataset, shown as a function of both the scaling exponent α and oscillatory frequency ν of the non-Gaussian signature.	163
Figure B.8.	Exclusion limit at 95% c.l. on the absolute value of the non-Gaussian non-oscillatory amplitude $ f_{\text{NL}}^{\alpha, \psi} $ inferred from the BOSS DR12 dataset as a function of the scaling exponent α	163

LIST OF TABLES

Table 2.1.	Parameters of the fiducial Λ CDM model with the sum of neutrino masses $\Sigma m_\nu = 0.06 \text{ eV}$ and the biasing model employed in the forecasts.	32
Table A.1.	Basic specifications for a BOSS-like survey, covering a total sky area of 10252 deg^2 with a total of about 1.2×10^6 objects.	142
Table A.2.	Basic specifications for DESI, covering a sky area of 14000 deg^2 with a total of about 2.7×10^7 objects.	144
Table A.3.	Basic specifications for the Euclid survey covering a sky area of 15000 deg^2 with a total of about 4.9×10^7 objects.	144
Table A.4.	Basic specifications for SPHEREx covering a sky fraction $f_{\text{sky}} = 0.75$ with a total of about 7.0×10^8 objects.	145
Table A.5.	Basic specifications for MegaMapper covering a sky area of 14000 deg^2 with a total of about 6.6×10^7 objects.	146
Table A.6.	Basic specifications for a futuristic spectroscopic billion-object survey, covering a sky fraction $f_{\text{sky}} = 0.5$ with a total of 10^9 objects.	146
Table A.7.	Basic specifications for Vera Rubin Observatory's LSST covering a sky area of 20000 deg^2 with a total of about 4×10^9 objects.	146

ACKNOWLEDGEMENTS

The seven years of my Ph.D. journey have been nothing short of life-changing. Through countless moments of struggle, excitement, self-doubt, and frustration, I have been extremely fortunate to meet and work with so many incredible people. Their presence not only helped me reach the finish line of this long and challenging journey, but also helped me rediscover my passion for physics through teaching.

First and foremost, I would like to thank my advisor and the chair of my committee, Professor Daniel Green, whose passion, guidance, and patience allowed me to develop my own identity and follow a unique path toward what I truly enjoy. His curiosity, inquisitiveness, and remarkable ability to dissect and simplify complex physical problems shaped the direction of this work and inspired many of its findings. I am also deeply grateful for the countless insightful discussions I shared with my collaborators, Yi Guo and Benjamin Wallisch, whose intellectual rigor and dedication to detail helped resolve many of the mysteries in our studies.

I am deeply grateful to the many professors I worked with as a teaching assistant, especially Professor Phil Tsai and Professor Gurleen Bal, whose kindness and meticulous work ethic contributed tremendously to both my academic and personal growth. Their mentorship helped me discover and embrace my passion for teaching over the past few years.

I would like to express my appreciation to Professor Kam Arnold, Professor Raphael Flauger, and Professor Alexey Arefiev for serving on my thesis committee and for their valuable feedback and suggestions. I also cannot overstate the support I received from the department staff, especially Sharmila Poddar and Catherine McConney.

Outside of academics, I am incredibly grateful for my enduring friendships with Nijaid Arredondo, Stephen Ebert, Yifan Hong, and Dhruv Muley, and the many friends I met during my time in San Diego, including Zipei Zhang, Lingfeng Wei, Varun Ramaprasad, Biswajit Sahoo, Tianhai Wu, Yuan Zhang, Ying Yu, and many others. Their friendship has been a constant source of joy, laughter, and encouragement throughout these years.

I want to express my deepest gratitude to my partner, whose radiance inspired confidence

and courage in me, and whose presence helped me persevere when the finish line felt impossibly distant. Her love and support continue to give me strength.

Finally, I want to thank my parents for their love, belief, unwavering support, and countless sacrifices over the years. Their faith in me carried me through the most difficult moments of my life, and I would not be here without them.

Chapter 2, in full, is a reprint of the material as it appears in *Light Fields during Inflation from BOSS and Future Galaxy Surveys*, Green, Daniel; Guo, Yi; Han, Jiashu; Wallisch, Benjamin, JCAP, 2024. The dissertation author was one of the primary investigators and authors of this paper.

Chapter 3, in full, is a reprint of the material as it appears in *Extending the Cosmological Collider: New Scaling Regimes and Constraints from BOSS*, Green, Daniel; Han, Jiashu; Wallisch, Benjamin, arXiv:2602.12232. The dissertation author was one of the primary investigators and authors of this paper.

VITA

2019 Bachelor of Arts, University of California Berkeley
2026 Doctor of Philosophy, University of California San Diego

PUBLICATIONS

J. Han, S. Ferraro, E. Giusarma and S. Ho, “Probing Gravitational Lensing of the CMB with SDSS-IV Quasars,” *MNRAS* **485** (2019) 1720-1726, arXiv:1809.04196 [astro-ph.CO].

D. Green, Y. Guo, J. Han and B. Wallisch, “Light Fields during Inflation from BOSS and Future Galaxy Surveys,” *JCAP* **05** (2024) 090, arXiv:2311.04882 [astro-ph.CO].

D. Green, J. Han and B. Wallisch, “Extending the Cosmological Collider: New Scaling Regimes and Constraints from BOSS,” arXiv:2602.12232 [astro-ph.CO].

ABSTRACT OF THE DISSERTATION

Non-Gaussian Signatures of Primordial Interactions During Inflation in the Large-Scale Structure

by

Jiashu Han

Doctor of Philosophy in Physics

University of California San Diego, 2026

Professor Daniel Green, Chair

Cosmic inflation is the leading theory for the generation of fluctuations in the early universe. Proposed interactions during inflation can lead to the generation of novel signatures on top of the minimal Gaussian, nearly scale-invariant fluctuations predicted by the standard model of cosmology. While the fields generating these signatures have long since decayed and are no longer relevant for the dynamics of the late-time universe, their interactions with the inflaton leave perturbations in the spacetime curvature that seed the matter distribution of the present-day universe through gravitational interactions, allowing the large-scale structure (LSS) to be a laboratory for probing these primordial interactions at the present day. In this dissertation,

we study the non-Gaussian signals generated by field interactions during inflation by analyzing the scale-dependent features they imprint in the power spectrum of galaxies, and the degeneracies with nonlinear evolution and other late-time effects. In Chapter 1, we discuss the theoretical background and motivation of this work. In Chapter 2, we study the scale-dependent bias that results from interactions involving light fields during inflation, and explore the feasibility of detecting such signatures from present and future galaxy surveys. We show that the survey sensitivity can be significantly improved by optimizing the survey design. In Chapter 3, we expand the search further by considering a wider range of scale-dependent signatures that are generated through the cosmological collider. We make measurements and constraints on the non-Gaussian amplitude in the expanded parameter space. While no signal has been detected, our analysis offers insights into future galaxy surveys' ability to constrain inflationary models through the two-point statistics. We further discuss the effects of incorporating additional physical effects in our modeling on the constraints.

Introduction

The study of modern physical cosmology began with the development of the theory of general relativity in the early twentieth century, and primarily concerns the origin, evolution and structure of the Universe through models based on the laws of physics. Backed by major discoveries that followed in the last century, the most promising cosmological models support the Big Bang as the origin of the Universe. Since then, the standard Λ CDM model of cosmology has seen remarkable success in tying physical processes of the very small, such as Big Bang nucleosynthesis of the light nuclei, to the very big, such as the statistical properties of the spatial distribution of galaxies and the cosmic microwave background (CMB), as well as connecting the dynamics in the early seconds of the Universe to its structure and properties at the present time [1]. This allows us to study the early-universe dynamics in the extremely far past with incredible precision.

Despite these successes, some problems have persisted within the framework of the standard cosmological model. In particular, the apparent fine-tuning required to address the spatial flatness and large-scale homogeneity of the Universe, and the origin of the small-scale fluctuations need a convincing explanation. These issues have led to the development of many theories, among which the most widely accepted is cosmic inflation, which solves the aforementioned problems with a relatively simple argument [2–5].

Cosmic inflation postulates that a period of accelerated, approximately exponential expansion occurred during the early universe. In the simplest realization, this expansion is driven by a scalar field called inflaton, that slowly rolls down a sufficiently flat potential so that its potential energy dominates over its kinetic energy. The period of inflation ended when the

slow-roll conditions were violated, followed by a reheating phase during which the inflaton transferred its energy to Standard Model degrees of freedom [6]. While the theory requires the introduction of a new field, it simultaneously solves the fine-tuning problems and provides a mechanism for generating the primordial fluctuations that seed the cosmic structure we observe today [7].

The information about the field content and interactions during inflation is encoded in the correlation functions of the primordial curvature perturbation ζ . In the simplest model, inflation is driven by a single scalar field with minimal interactions. This results in perturbations that are nearly Gaussian, adiabatic, and approximately scale-invariant, which can be captured by the two-point correlation function or power spectrum

$$\langle \zeta(\vec{k}) \zeta(\vec{k}') \rangle = (2\pi)^3 \delta_D(\vec{k} + \vec{k}') P_\zeta(k). \quad (1)$$

More general models include additional fields that are coupled to the inflaton, leading to distinct signatures in the three-point and higher-order correlation functions of the primordial fluctuations [8, 9].

Depending on the fields at play during inflation, these signatures exhibit different qualitative features. Interactions involving additional light fields can generate enhanced squeezed-limit correlations that induce a scale-dependent bias in the galaxy power spectrum [10]

$$\Delta b(k) \propto k^{\Delta-2}. \quad (2)$$

On the other hand, interactions involving additional massive fields could lead to oscillatory or non-analytic features in the correlation functions, which encode the mass spectrum of inflationary-era particles [11–13]. These insights have led to the idea of a “cosmological collider,” in which the early universe acts as a high-energy environment capable of producing new particles and imprinting their signatures on cosmological observables [14]. Observing these signatures

would allow us to explore fundamental physics at scales far beyond the reach of any Earth-based experiment, but measuring them requires making connections between the early-universe processes and late-time observables.

During inflation, modes of perturbations are stretched beyond the Hubble horizon. For adiabatic modes, the curvature perturbation ζ becomes conserved on super-horizon scales, allowing primordial information to be preserved until the modes re-enter the horizon at a later time. Gravitational instability then causes the primordial fluctuations to be transferred into matter density fluctuations, which in turn evolve into halos and galaxies through linear and nonlinear processes, giving rise to the large-scale structure (LSS) of the Universe we observe today [15]. This opens a pathway for us to access primordial physics.

Primordial non-Gaussianity (PNG) is most directly encoded in the three-point (i.e. bispectrum) and higher- N -point correlation functions of cosmological observables. Existing observations place stringent constraints on the amplitudes of several well-studied bispectrum templates [16]. This implies that any potential signal beyond the two-point statistics must be small and therefore suffers from low signal-to-noise, making the detection of PNG contributions to higher- N -point statistics observationally challenging. In contrast, two-point statistics, such as the galaxy power spectrum, can be measured with relatively high precision in current and near-future galaxy surveys. Compared to other observables such as the CMB, large-scale structure measurements of PNG have an additional observational advantage due to its three-dimensional nature, because the number of independent Fourier modes in a three-dimensional LSS survey scales as k_{max}^3 , whereas the number of modes in a two-dimensional CMB map scales as ℓ_{max}^2 [17, 18]. Therefore, the scale-dependent features in the galaxy power spectrum offer a robust probe of primordial physics.

Nevertheless, the quest to extract information about the primordial universe from the large-scale structure faces many challenges. Theoretical considerations such as consistency conditions, symmetry arguments and Boltzmann suppression of heavy particle production for masses $m \gg H$ limit the observability of certain classes of primordial contributions [19–24].

Furthermore, nonlinear gravitational evolution, galaxy bias, and observational effects can result in features similar to the predicted primordial signal in the galaxy power spectrum, thereby obscuring the true signal of interest [25–28]. To disentangle the primordial signal from these various non-primordial effects, it is important to have a detailed model of the galaxy power spectrum that incorporates the underlying theory [29].

In this dissertation, our goal is to study how interactions involving both light and heavy fields during inflation leave imprints in the two-point statistics of the large-scale structure. We build upon theoretical frameworks of inflationary physics and the standard cosmological perturbation theory to identify and characterize these relic signatures in the distribution of galaxies, and place constraints on these signatures using existing data.

First, we investigate how primordial interactions between light fields and the inflaton modify the clustering of galaxies and introduce a scale-dependent bias on the galaxy two-point correlation. We quantify this scale-dependent contribution to the galaxy power spectrum in a minimally model-dependent framework, and determine conditions under which such signals can be distinguished from contributions of non-primordial origins, such as nonlinear gravitational evolution. We study the observational consequences of the scale-dependent bias in current and future galaxy surveys by forecasting the constraint on the signal amplitude f_{NL}^{Δ} , and provide insights into how these surveys can be optimized to search for these signals.

We then extend this analysis to a wider range of candidates including heavy fields during inflation, which give rise to oscillatory features in the scale-dependent bias in the galaxy power spectrum. These features are in principle distinguishable from biasing effects due to their characteristic oscillatory shape. We assess the feasibility of detecting such features in realistic survey settings, and illustrate how observations of the large-scale structure can allow us to probe the mass spectrum of particles during inflation.

Our results demonstrate that through the scale-dependent bias, the two-point correlation of the large-scale structure contains valuable information about interactions during cosmic inflation. With sufficiently precise theoretical modeling, the large-scale structure can therefore

provide a competitive and complementary probe of high energy physics beyond the reach of Earth-based experiments. The potential findings have implications not only for the field of cosmology and the history of the Universe, but also for our understanding of the fundamental interactions in nature.

This dissertation is organized as follows: In Chapter 1, we motivate this work by reviewing the connection between the late-time observables in the large-scale structure and the origin of the perturbations during cosmic inflation, and discuss treatments of the galaxy bias and other effects in the modeling of the galaxy power spectrum. In Chapter 2, we present the analysis of the impact of light-field interactions with the inflaton on the galaxy two-point statistics, and discuss strategies to improve survey sensitivity on this signal. In Chapter 3, we focus on oscillatory features in the galaxy power spectrum originating from heavy field interactions. Finally, we summarize the results in Chapter 4, and discuss future directions for probing primordial physics with the large-scale structure. A set of appendices provides additional details on the modeling, forecast and analysis of the light-field signals (Appendix A) and the oscillatory features (Appendix B).

Chapter 1

From Inflationary Interactions to Galaxy Clustering

In this chapter, we build a conceptual and mathematical bridge from cosmic inflation to the large-scale structure observables used throughout this dissertation, while providing context for the discussions and analyses in Chapters 2 and 3. This chapter is organized as follows: In §1.1.1, we review the origin of primordial scalar curvature perturbation ζ in the simplest inflationary scenario. Then in §1.1.2, we introduce additional fields and interactions responsible for generating the primordial non-Gaussianity (PNG), which is the main focus of this dissertation. Next, we describe how these primordial fluctuations are transferred into the late-time distribution of matter. We cover the linear growth of matter and the transfer function that relates ζ to the matter power spectrum in §1.2.1. We then discuss the biasing of galaxies with respect to the underlying matter density field, and the resulting scale-dependent signature of PNG in the galaxy power spectrum in §1.2.2. Finally, in §1.2.3, we examine additional observational and astrophysical systematics, such as redshift-space distortions, the degeneracy between the non-Gaussian bias and nonlinear gravitational evolution, and the damping of oscillatory features, that must be considered in the extraction of PNG from real galaxy survey data. Throughout this chapter, we focus on the statistical quantities and physical mechanisms most directly relevant to the analyses in the following chapters, and we emphasize that this chapter is not intended as a full review of inflation or structure formation. For more comprehensive treatments of these

topics, we refer the reader to refs. [17, 30] for inflation, refs. [8, 9] for the mechanisms generating primordial non-Gaussianity, ref. [25] for galaxy biasing, and ref. [31] for observational effects that complicate the measurement of the galaxy power spectrum.

1.1 Inflationary Origin of Primordial Fluctuations

1.1.1 Primordial Perturbations in Single-Field Inflation

Among the many realizations of cosmic inflation that have emerged since inflation was first proposed [2–5], the simplest and most influential is the single-field, slow-roll inflation, described by the action

$$S = \int d^4x \sqrt{-g} \left[\frac{M_{\text{pl}}^2}{2} R - \frac{1}{2} (\partial\phi)^2 - V(\phi) \right], \quad (1.1)$$

where R is the Ricci scalar, M_{pl} is the reduced Planck mass, and $V(\phi)$ is the potential for the inflaton ϕ . The signature of the metric is $(-1, 1, 1, 1)$.

The inflaton potential $V(\phi)$ takes a specific shape such that the inflaton field slowly rolls down the flat part of the potential, resulting in much larger potential energy than kinetic energy. This effectively behaves like a cosmological constant, and leads to a prolonged period of accelerated expansion. A set of potential slow-roll parameters are defined to conveniently characterize the slow roll,

$$\epsilon_v \equiv \frac{M_{\text{pl}}^2}{2} \left(\frac{V'}{V} \right)^2, \quad \eta_v \equiv M_{\text{pl}}^2 \frac{V''}{V}, \quad (1.2)$$

with the smallness of the parameters indicating the slow roll. Alternatively, slow-roll inflation can be characterized by the Hubble slow-roll parameters

$$\epsilon \equiv -\frac{\dot{H}}{H^2}, \quad \eta \equiv -\frac{\ddot{\phi}}{H\dot{\phi}}, \quad (1.3)$$

where H is the Hubble parameter. In the slow-roll regime, the two sets of slow-roll parameters are related by $\varepsilon \approx \varepsilon_v$, $\eta \approx \eta_v - \varepsilon_v$. An accelerated expansion requires $\varepsilon < 1$, and slow roll requires $\varepsilon, |\eta| \ll 1$, which implies that H evolves slowly, i.e. $|\dot{H}| \ll H^2$. The inflaton ϕ therefore acts as a “clock” and its behavior determines how long inflation lasts. A sufficiently long period of accelerated expansion satisfying the slow-roll conditions would significantly increase the radius of spatial curvature of the Universe, and allow perturbation modes to exit the Hubble horizon during the expansion and re-enter the horizon at a later time. The latter allows seemingly causally disconnected points in the Universe to be in causal contact with each other in the far past. As a result, inflation addresses the apparent fine-tuning required to explain the observed flatness and large-scale homogeneity of the Universe.

On the other hand, quantum fluctuations of the inflaton about its classical background trajectory, $\delta\phi(x, t)$, will result in perturbations in the spacetime curvature, described by the power spectrum of the curvature perturbations ζ ,

$$\langle \zeta(\vec{k}) \zeta(\vec{k}') \rangle = (2\pi)^3 \delta_D(\vec{k} + \vec{k}') P_\zeta(k). \quad (1.4)$$

The inflaton fluctuation $\delta\phi$, evaluated on spatially flat slices, is related at linear order to ζ , at the time of horizon crossing ($k = aH$), by

$$\zeta = -H \frac{\delta\phi}{\dot{\phi}}. \quad (1.5)$$

At leading order in the slow-roll parameters, this relationship fixes both the amplitude and the tilt of the power spectrum in (1.4), and it is conventional to define a dimensionless power spectrum

$$\Delta_\zeta^2 \equiv \frac{k^3 P_\zeta(k)}{2\pi^2} = A_s \left(\frac{k}{k_*} \right)^{n_s-1}, \quad (1.6)$$

where A_s is the amplitude and $n_s = 1 + 2\eta_v - 6\varepsilon_v$ is the spectral index, both evaluated at the time of horizon crossing ($k = aH$). k_* is an arbitrarily chosen pivot scale. CMB observations constrain

these parameters to $\ln(10^{10}A_s) = 3.044 \pm 0.014$ and $n_s = 0.9649 \pm 0.0042$ at $k_* = 0.05 \text{ Mpc}^{-1}$ [1], which are consistent with the slow-roll prediction of a nearly scale-invariant power spectrum.

The power spectrum fully captures the statistical information of Gaussian primordial fluctuations. Any deviations from a Gaussian distribution, i.e. non-Gaussianity, would result in correlations beyond the two-point function/power spectrum, with the leading contribution coming from the bispectrum B_ζ , defined via

$$\langle \zeta(\vec{k}_1) \zeta(\vec{k}_2) \zeta(\vec{k}_3) \rangle = (2\pi)^3 \delta_D(\vec{k}_1 + \vec{k}_2 + \vec{k}_3) B_\zeta(k_1, k_2, k_3), \quad (1.7)$$

where the delta function imposes conservation of momentum on the three input momenta. This condition implies that configurations of the wavevectors contributing to the bispectrum must form closed triangles. It is often convenient to consider some special configurations of the bispectrum, which correspond to specific model predictions:

1. Squeezed triangle ($k_1 \rightarrow 0, k_2 \approx k_3$): particularly sensitive to additional light or heavy fields during inflation [8],
2. Folded triangle ($k_1 = k_2 = k_3/2$): enhanced in models with non-standard initial states [32],
3. Equilateral triangle ($k_1 = k_2 = k_3$): enhanced in models with higher-derivative interactions [33].

The simplest and most widely studied way to parameterize non-Gaussianity is *local* non-Gaussianity, defined through a nonlinear correction to a Gaussian perturbation ζ_g that is local in real space

$$\zeta(\vec{x}) = \zeta_g(\vec{x}) + \frac{3}{5} f_{\text{NL}}^{\text{loc}} (\zeta_g(\vec{x})^2 - \langle \zeta_g(\vec{x})^2 \rangle) + \dots \quad (1.8)$$

Signals of local non-Gaussianity peak in the squeezed-triangle configuration of the bispectrum, with the parameter $f_{\text{NL}}^{\text{loc}}$ describing the amplitude of the non-Gaussian signal.

For single-field inflation, the squeezed limit of the bispectrum, which correlates a long-

wavelength mode ($k_L \equiv k_1 \rightarrow 0$) and two short-wavelength modes ($k_2 \approx k_3 \equiv k_S$), is fixed entirely by the tilt of the power spectrum, i.e. $n_s - 1$, through the single-field consistency relation [19, 34],

$$\lim_{k_L \rightarrow 0} B_\zeta(k_L, k_S, k_S) = -[(n_s - 1) + O(k_L^2/k_S^2)] P_\zeta(k_L) P_\zeta(k_S). \quad (1.9)$$

Using (1.7) and (1.8), this result translates to an effective local non-Gaussianity of

$$f_{\text{NL}}^{\text{loc}} = \frac{5}{12}(1 - n_s). \quad (1.10)$$

Observationally, $n_s - 1$ is constrained to be $O(10^{-2})$, so single-field inflation predicts a very small $f_{\text{NL}}^{\text{loc}}$. Moreover, the non-Gaussian contribution in (1.9) is the only type of non-Gaussian signature permitted by the single-field model in the squeezed limit [20–22]. Therefore, any observable primordial non-Gaussianity in violation of the single-field consistency condition (1.9) would point to the existence of additional degrees of freedom beyond the inflaton, or a breakdown of the assumptions underlying the single-clock attractor inflaton scenario, which we will discuss next.

1.1.2 Interactions and Extra Fields

As discussed in the previous section, single-field inflation can only generate a specific type of non-Gaussianity with very small amplitude. In contrast, additional fields present during inflation can shift the value of the curvature perturbation ζ through their mixing with the inflaton or through nonlinear super-horizon evolution. In the δN formalism, the curvature perturbation on uniform-density slices is given by the change in the number of e -folds due to fluctuations of all light fields ϕ_i present at horizon crossing [35, 36],

$$\zeta(\vec{x}) = \delta N = N_i \delta \phi_i + \frac{1}{2} N_{ij} \delta \phi_i \delta \phi_j + \dots, \quad (1.11)$$

where $N_i \equiv \partial N / \partial \phi_i$ and its derivatives are evaluated on the unperturbed inflaton trajectory. When $N_{ij} \neq 0$, the nonlinear terms involving the light field fluctuations $\delta \phi_i$ in (1.11) directly source non-Gaussianity in ζ .

We will again consider the local non-Gaussianity introduced in §1.1.1. Local-type PNG can be generated when fluctuations of an additional massless isocurvature field are nonlinearly converted into the adiabatic curvature perturbation ζ after horizon exit. According to (1.8), this produces a bispectrum that peaks in the squeezed limit,

$$B_\zeta(k_L, k_S, k_S) \simeq \frac{12}{5} f_{\text{NL}}^{\text{loc}} P_\zeta(k_L) P_\zeta(k_S), \quad k_L \ll k_S. \quad (1.12)$$

In contrast to the single-field prediction in (1.10), this squeezed bispectrum is not suppressed by the slow-roll conditions.

The squeezed bispectrum measures the response of small-scale power (fluctuations in ζ) to the amplitude of a long-wavelength mode, and the scaling of the squeezed bispectrum with k_L/k_S directly probes features of the fields present during inflation, through a scaling exponent Δ that we define explicitly below,

$$B_\zeta(k_L, k_S, k_S) \propto f_{\text{NL}}^\Delta \left(\frac{k_L}{k_S} \right)^\Delta P_\zeta(k_L) P_\zeta(k_S), \quad (1.13)$$

with f_{NL}^Δ the amplitude of the squeezed bispectrum of general scaling Δ . A massless field would produce the scale-invariant ($\Delta = 0$) scaling in (1.12), i.e. local non-Gaussianity, while a massive field would leave its own characteristic scale dependence in the squeezed bispectrum.

To understand how the mass of an additional scalar controls the scaling exponent in the squeezed limit, consider the quasi-single-field inflation (QSFI) model, in which a massive spectator field σ of mass m propagates in an approximately de Sitter background, described by the action

$$S_\sigma = \int d^4x \sqrt{-g} \left[-\frac{1}{2} (\partial \sigma)^2 - \frac{1}{2} m^2 \sigma^2 + \rho \delta \phi \sigma \right], \quad (1.14)$$

in which σ is directly coupled to the inflaton through a linear mixing term with coupling ρ , which generically arises whenever the inflationary trajectory is not exactly straight in field space [11, 12].

The dominant super-horizon (i.e. $k \ll aH$) behavior of σ is captured in the squeezed bispectrum by a real scaling exponent $\Delta = \frac{3}{2} - \nu$ where

$$\nu = \sqrt{\frac{9}{4} - \frac{m^2}{H^2}}. \quad (1.15)$$

For light fields with $0 \leq m \leq 3H/2$, this translates to $\Delta \in [0, 3/2]$. This result, along with (1.13), generalizes local non-Gaussianity to a squeezed bispectrum with a general power-law scaling. In §1.2.2, we show that this general scaling dimension Δ controls the scale-dependent bias in the galaxy power spectrum.

Although light fields alone only produce $\Delta \in [0, 3/2]$, our analysis in Chapter 2 considers a wider range of $\Delta \in [0, 2]$, with $\Delta = 2$ corresponding to the squeezed-limit scaling characteristic of the equilateral non-Gaussianity. This might seem surprising at first: $\Delta = 2$ not only lies outside the light-field range derived above, but it also does not match the $m \rightarrow \infty$ limit of the heavy-field case below, where the scaling dimension instead approaches the complex value $\Delta \approx \frac{3}{2} + im/H$, set by the wavefunction of a single massive particle in an expanding universe [14]. As we will explain in more detail in §2.3.2, the $\Delta = 2$ contribution arises from a distinct, subleading effect: the virtual exchange of a massive field leaves behind a purely local self-interaction of the inflaton, which has the same scaling as the one produced by the derivative self-interactions of single-field inflation [37, 38]. Because this interaction is local (i.e. k -independent) at the level of the primordial bispectrum, it was long thought to be indistinguishable from the contributions described by an ordinary galaxy bias, and therefore unobservable through the scale-dependent bias [39]. However, because matter continues to evolve gravitationally after horizon re-entry, a term that is local *during* inflation is not necessarily degenerate with a term that is local at *late times*, which allows equilateral non-Gaussianity to be measured through the scale-dependent

bias after all [40].

Following the same prescription in (1.14), the scaling dimension Δ analytically continues to a complex value for heavier fields with $m > 3H/2$, and (1.15) is replaced by

$$\nu = \sqrt{\frac{m^2}{H^2} - \frac{9}{4}}, \quad (1.16)$$

with the scaling dimension being $\Delta = \frac{3}{2} + i\nu$. Instead of following a smooth power law, the squeezed bispectrum now oscillates logarithmically in k_L/k_S , with a frequency set by ν ,

$$B_\zeta(k_L, k_S, k_S) \propto P_\zeta(k_L)P_\zeta(k_S) \left(\frac{k_L}{k_S}\right)^{3/2} \cos\left[\nu \ln\left(\frac{k_L}{k_S}\right) + \varphi\right], \quad (1.17)$$

where the phase φ and the overall amplitude of the bispectrum are model-dependent. The amplitude is in general exponentially suppressed by the Boltzmann factor for producing a particle of mass $m \gg H$.

Equations (1.16) and (1.17) form the basis of the so-called “cosmological collider” [14], in which the oscillation frequency ν directly encodes the mass of the additional particle interacting with the inflaton. A measurement of the squeezed bispectrum therefore acts as a probe of the particle spectrum during inflation, much like what a collider experiment does.

Note that the value $\alpha \equiv \text{Re}\Delta = 3/2$ above is not a coincidence for heavy fields in the QSFI example. It is the minimum value allowed by unitarity for any principal-series (oscillatory) field propagating in pure de Sitter space. For these fields, the optical theorem forces $\alpha \geq 3/2$, regardless of the details or strength of the field’s interactions with the inflaton [23], and the QSFI construction above saturates this bound.

In Chapter 3, we generalize this picture further by allowing the scaling dimension to take a fully complex value $\Delta = \alpha + i\nu$ that is not fixed to the principal-series value of $\alpha = 3/2$. This generalization is possible because inflation only approximates de Sitter space. As discussed in §1.1.1, the time evolution of the inflaton provides a physical clock, so any interaction involving

additional fields during inflation that explicitly depends on this clock, such as a direct mixing between two fields written in terms of physical rather than conformal time, breaks the de Sitter isometries. As a result, the optical-theorem bound on α no longer applies. It has been shown that models containing explicit Lorentz-violating or time-dependent mixing terms can allow $\alpha < 3/2$ together with a nonzero oscillation frequency ν , and in some cases it is possible to achieve $\alpha \rightarrow 0$ [23, 24]. Since the power-law suppression, $(k_L/k_S)^\alpha$, of the squeezed bispectrum in (1.17) is controlled by α , these models with smaller values of α can produce oscillatory signals that are less suppressed and therefore considerably easier to observe than the fixed $\alpha = 3/2$ case. We refer to this generalization as the *extended cosmological collider* and study its phenomenology in detail in Chapter 3.

1.2 Structure Formation

In the previous section, we discussed how interactions during inflation can generate non-Gaussian correlations in the primordial curvature perturbation ζ . We now shift our focus to how these primordial fluctuations are transferred to the late-time distribution of matter, and ultimately galaxies, which are the main observables we consider in Chapters 2 and 3.

1.2.1 Linear Growth and Transfer Functions

The rapid expansion of the Universe during inflation causes modes of the primordial fluctuations to exit the horizon, and become frozen at super-horizon scales. After the end of inflation, these modes re-enter the horizon, with smaller-scale modes generally re-entering earlier. After a mode re-enters the horizon, gravitational instability converts the primordial curvature perturbation ζ into a matter density fluctuation (*overdensity*),

$$\delta_m(\vec{x}, a) \equiv \frac{\rho_m(\vec{x}, a) - \bar{\rho}_m(a)}{\bar{\rho}_m(a)}, \quad (1.18)$$

where $\rho_m(\vec{x}, a)$ is the matter density field, and $\bar{\rho}_m(a)$ is the average matter density.

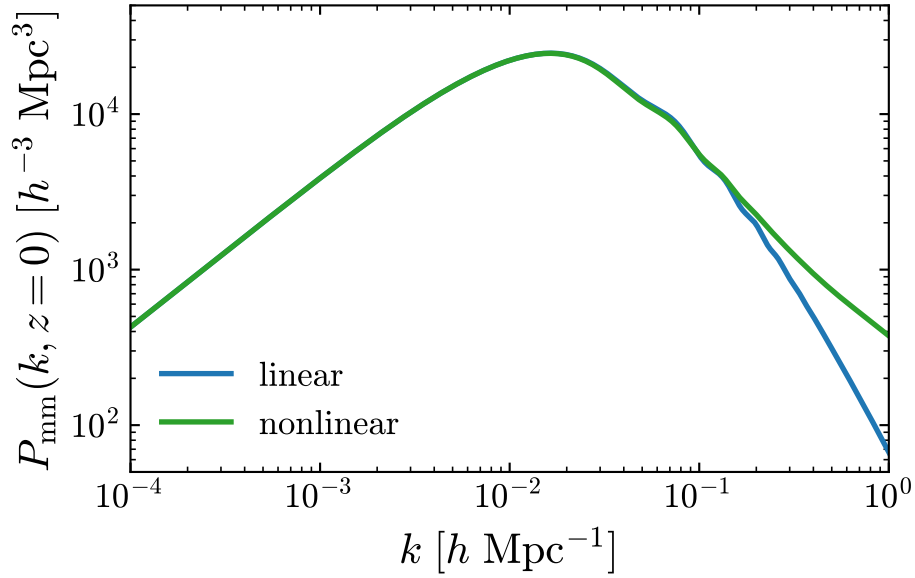


Figure 1.1. The matter power spectrum $P_{\text{mm}}(k)$ at $z = 0$. The growth of structure is linear at the largest scales, but the linear approximation breaks down at small scales as matter collapses under gravity to form structures like halos and galaxies. The transition between the linear and nonlinear regimes is marked by the nonlinear scale k_{NL} , which is roughly $0.1 \, h \, \text{Mpc}^{-1}$ at $z = 0$.

When δ_m is small, its evolution in Fourier-space factorizes into a transfer function $T(k)$, which encodes sub-horizon physics such as the suppression of growth for modes that entered the horizon during radiation domination, and a linear growth factor $D(a)$, normalized as $D(a) = a$ during matter domination,

$$\delta_m(\vec{k}, a) \equiv k^2 \mathcal{T}(k, a) \Phi(\vec{k}) = \frac{2k^2 T(k) D(a)}{3\Omega_m H_0^2} \Phi(\vec{k}), \quad (1.19)$$

where $\Phi = -\frac{3}{5}\zeta$ is the primordial Newtonian potential, Ω_m is the present-day matter density parameter and H_0 is the Hubble constant. In the linear regime ($\delta_m \ll 1$), the fluctuations in δ_m are captured by the linear matter power spectrum,

$$P_{\text{mm}}^{\text{lin}}(k, a) \propto k^4 D^2(a) T^2(k) P_\zeta(k). \quad (1.20)$$

Figure 1.1 shows the matter power spectrum at $z = 0$, with a transition at the nonlinear scale

$k_{\text{NL}}(z=0) \approx 0.1 h \text{Mpc}^{-1}$. At the largest scales ($k \ll k_{\text{NL}}$), structures grow linearly with $\delta_m \ll 1$, and the fluctuations in δ_m are well-described by the linear matter power spectrum. On the other hand, gravitational nonlinearities become important at smaller scales ($k \gtrsim k_{\text{NL}}$) [15] as δ_m becomes $O(1)$. In order to accurately capture the growth of structure near the nonlinear scale, it is necessary to model the nonlinearities via a bias expansion, which we will discuss in §1.2.2.

As mentioned in §1.1.1, single-field inflation generates a nearly scale-invariant power spectrum given by (1.6), which translates to $P_{\text{mm}}(k) \propto k^{n_s}$ at the largest scales. The existence of light and heavy fields during inflation will lead to the generation of PNG which is encoded in the matter bispectrum through (1.7). On the other hand, the linear matter power spectrum is completely insensitive to PNG at tree level. Modifications of $P_{\text{mm}}(k)$ due to PNG only enter through loop corrections and are suppressed. This appears to limit the prospect of detecting PNG through the two-point statistics of late-time observables. However, the squeezed bispectrum in (1.13) couples long-wavelength modes with small-scale variance. In the next section, we will discuss how this coupling induces a scale-dependent bias in the galaxy power spectrum that can become significant at specific scales.

1.2.2 Galaxy Biasing

The formation of galaxies from matter density fluctuations is a complicated process that depends on the local environment, and the matter density fluctuations do not map perfectly to fluctuations in the number density of galaxies. As a result, galaxies are biased tracers of the underlying matter density field. In order to extract information about the matter overdensity δ (previously denoted by δ_m) from the distribution of galaxies, the nonlinear galaxy formation process must be modeled empirically. On sufficiently large scales, this can be done through the bias expansion, which treats the galaxy overdensity δ_g as an expansion in the matter overdensity

δ and its derivatives [25],

$$\delta_g(\vec{x}, z) = b_1(z)\delta(\vec{x}, z) + \frac{b_2(z)}{2}\delta^2(\vec{x}, z) + b_{s^2}(z)s^2(\vec{x}, z) + b_{\nabla^2\delta}(z)\nabla^2\delta(\vec{x}, z) + \varepsilon(\vec{x}, z) + \dots, \quad (1.21)$$

where s^2 is an operator constructed from the tidal tensor $\nabla_i\nabla_j\Phi$. The term containing $\nabla^2\delta$ is a gradient bias term capturing nonlocal effects on δ , and $\varepsilon(\vec{x}, z)$ is a stochastic term uncorrelated with δ on large scales. Chapters 2 and 3 extend this expansion to one-loop order with additional operators to model the full galaxy power spectrum in the mildly nonlinear regime ($k \gtrsim k_{\text{NL}}$).

Following the expansion in (1.21), in the absence of primordial non-Gaussianity, the leading order galaxy auto- and galaxy-matter cross-power spectra are related to the underlying matter power spectrum directly through the linear bias $b_1(z)$,

$$P_{gg}(k, z) = b_1^2(z)P_{\text{mm}}(k, z) + P_{\varepsilon\varepsilon}(k, z), \quad P_{gm}(k, z) = b_1(z)P_{\text{mm}}(k, z). \quad (1.22)$$

The presence of PNG, however, changes the behavior of the galaxy power spectrum by introducing a scale-dependent effect in the linear bias. Consider again the example of local non-Gaussianity introduced in §1.1.2: As a direct result of equation (1.8), the late-time Newtonian potential Φ is a nonlinear function of a Gaussian field ϕ_G ,

$$\Phi(\vec{x}) = \phi_G(\vec{x}) - f_{\text{NL}}^{\text{loc}} [\phi_G^2(\vec{x}) - \langle \phi_G^2 \rangle]. \quad (1.23)$$

Because Φ directly sources the matter density fluctuations δ , the presence of local PNG ($f_{\text{NL}}^{\text{loc}} \neq 0$) introduces a modulation of $\langle \delta(k_S)\delta(k_S) \rangle$, i.e. the small-scale variance of the matter density fluctuations, by a long-wavelength mode of ϕ_G . This modulation results in locally enhanced/suppressed variance in $\delta(k_S)$. Since galaxies form in overdense regions, an increase in the local variance of $\delta(k_S)$ will lead to an increase in the local abundance of galaxies, as depicted in Figure 1.2. Therefore, the galaxy number density fluctuation δ_g acquires a scale dependence that becomes

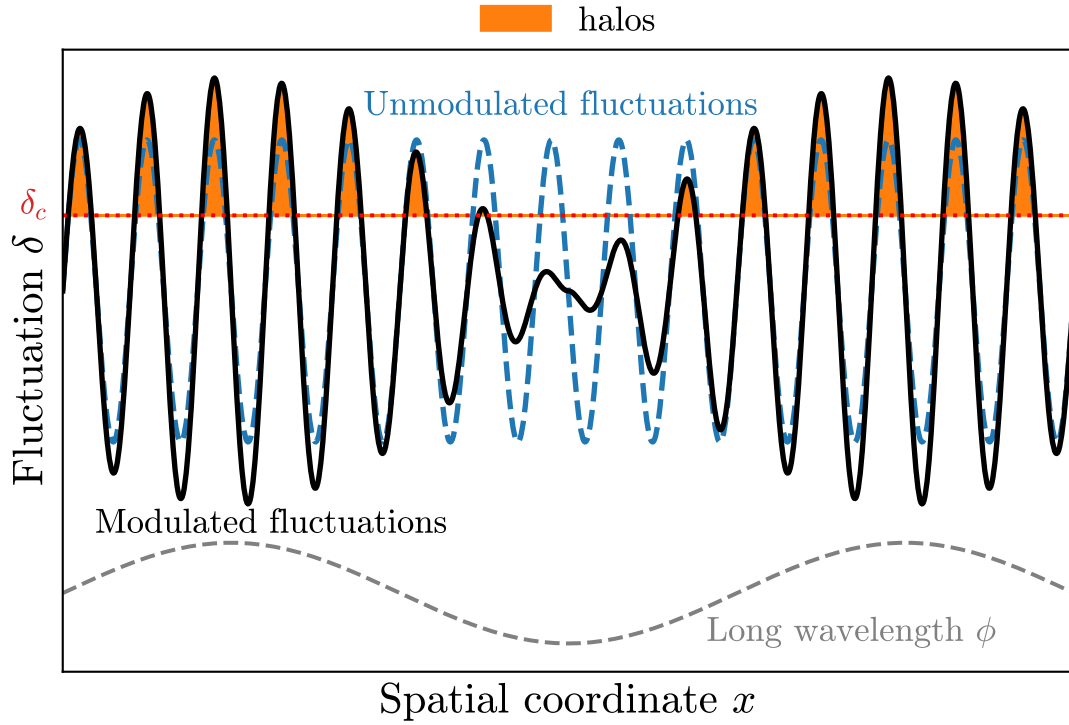


Figure 1.2. Modulation of the small-scale variance in matter density fluctuations, $\langle \delta(k_S) \delta(k_S) \rangle$, by a long-wavelength mode $\phi(\vec{x})$. In this simplistic model, halos/galaxies form in places where the local matter density is above some threshold δ_c . The large-scale modulation of the variance in $\delta(k_S)$ effectively changes the local formation threshold, leading to boosted (suppressed) halo/galaxy formation in places with larger (smaller) variance.

significant at long wavelengths. At leading order, this effect appears in the galaxy power spectrum as a correction to the linear galaxy bias [10, 41, 42],

$$\Delta b(k, z) \simeq \frac{f_{\text{NL}}^{\text{loc}} b_{\phi}(z)}{k^2 \mathcal{T}(k, z)}, \quad (1.24)$$

where $\mathcal{T}(k, z)$ is the Poisson-equation kernel in (1.19), and $b_{\phi}(z)$ is a non-Gaussian bias parameter whose form depends on the modeling of halo formation. In the limit $k \rightarrow 0$, $\mathcal{T}(k, z)$ approaches a constant, leading to $\Delta b(k, z) \propto k^{-2}$ on the largest scales.

Equation (1.24) is motivated using the example of local non-Gaussianity, but this result can be extended by considering the generalized squeezed bispectrum in (1.13), for which local non-Gaussianity is a special case with $\Delta = 0$. This leads to the general scale-dependent bias

$$\Delta b(k, z) \propto \frac{f_{\text{NL}}^{\Delta} b_{\phi}(z)}{k^2 \mathcal{T}(k, z)} (k R_*)^{\Delta} \propto k^{\Delta-2}, \quad (1.25)$$

where R_* corresponds to the locality scale for halo formation. In Chapters 2 and 3, we derive and use the general scale-dependent bias expression extensively.

Note that in (1.25), the non-Gaussian amplitude f_{NL}^{Δ} is degenerate with the non-Gaussian response coefficient $b_{\phi}(z)$, which is often related to the linear bias b_1 through a universality relation for halo formation. In §1.2.3, we will briefly discuss how this degeneracy can be treated in practice, along with other observational effects that must be considered in order to place constraints on the non-Gaussian parameter f_{NL}^{Δ} .

1.2.3 Observational and Astrophysical Systematics

Actual measurement of primordial non-Gaussianity through the scale-dependent bias in general requires careful consideration of additional effects related to the survey observation, model assumptions, and large-scale gravitational nonlinearities. We briefly summarize these effects in this section, and discuss how they are treated in the forecasts and data analyses in

Chapters 2 and 3.

The first effect is the anisotropic distortion of the observed clustering of galaxies along the line of sight, known as redshift-space distortions. In an actual galaxy survey, the positions of galaxies are indirectly measured through the measurement of redshifts. The radial positions inferred from redshift measurements receive contributions from the galaxies' peculiar velocities, producing distortions in the observed radial distribution of galaxies. At linear order, the modification of the galaxy bias due to the motion of galaxies falling into overdense regions is given by

$$b(k, \mu, z) = b_1(z) + f(z)\mu^2 + \dots, \quad (1.26)$$

where μ is the cosine of the angle between the wavevector $\vec{k}$ and the line of sight $\hat{n}$, and $f \equiv d \log D / d \log a$ is the linear growth rate [43].

A second observational distortion, the Alcock–Paczynski effect, arises when the fiducial cosmology used to convert observed angles and redshifts into comoving wavenumbers differs from the true cosmology. This effect rescales both the amplitude and wavenumbers of the measured power spectrum [44]. Because both the redshift-space distortions and the Alcock–Paczynski effect depend on the scale and angle, they must be modeled jointly with the scale-dependent bias.

Another complication in the measurement of the scale-dependent bias comes from the nonlinear bias operators in §1.2.2 that describe gravitational clustering and nonlinearities. These operators generate their own scale-dependent contributions to the galaxy power spectrum that become increasingly important toward smaller scales, and can become degenerate with the primordial signal of interest, particularly for squeezed-limit scalings with larger Δ . These late-time contributions to the galaxy power spectrum must be marginalized over in order to obtain a meaningful measurement of the non-Gaussian amplitude f_{NL}^Δ . In practice, this is done by fitting for all of the parameters in the bias expansion in (1.21) together with the non-Gaussian bias $b_\phi(z)$, which without further theoretical input is completely degenerate with f_{NL}^Δ in (1.25). We break

this degeneracy in the analyses in Chapters 2 and 3 using a commonly assumed universality relation motivated by the physics of halo formation,

$$b_\phi(z) = 2\delta_c [b_1(z) - p], \quad (1.27)$$

where $\delta_c \approx 1.686$ is the critical overdensity for spherical collapse, and $p = 1$ for a universal halo mass function [25], which we adopt in the forecasts. Nevertheless, the validity of this universality assumption for specific galaxy populations remains an active area of study.

Finally, the oscillatory non-Gaussian signatures due to heavy fields in §1.1.2 can be degenerate with features of the baryon acoustic oscillations (BAO). Both leave sharp, oscillatory features in the galaxy power spectrum, and both are subject to the same physical damping mechanism. Large-scale bulk flows displace galaxies from their initial positions by an amount that is only known statistically, which smears out sharp features in the linear power spectrum,

$$P(k, \mu) \simeq P_{\text{nw}}(k) + e^{-k^2 \Sigma^2(\mu)} [P_{\text{lin}}(k) - P_{\text{nw}}(k)], \quad (1.28)$$

where $P_{\text{nw}}(k)$ is the smooth, “no-wiggle” power spectrum, and the damping scale $\Sigma^2(\mu) = \Sigma_\perp^2 + \mu^2(\Sigma_\parallel^2 - \Sigma_\perp^2)$ is set by the (anisotropic) variance of the large-scale linear displacement field [45–47]. This damping mechanism suppresses the amplitude of the oscillatory scale-dependent bias of Chapter 3 at high frequency ν , and in general must be accounted for at scales where the effective linear oscillation period is comparable to the damping scale $\Sigma(\mu)$ [18].

In summary, this chapter lays out the theoretical motivation for the signatures studied in the rest of this work. In Section 1.1, the effect of additional fields during inflation is characterized in the squeezed bispectrum (1.13) using a scaling dimension Δ that captures the information about the field content. In Section 1.2, the mapping of this scaling into a scale-dependent bias in the galaxy power spectrum is presented, along with the late-time and observational effects that should be considered in the measurement of this scale-dependent signature.

Chapter 2

Light Fields during Inflation from BOSS and Future Galaxy Surveys

2.1 Synopsis

Primordial non-Gaussianity generated by additional fields present during inflation offers a compelling observational target for galaxy surveys. These fields are of significant theoretical interest since they offer a window into particle physics in the inflaton sector. They also violate the single-field consistency conditions and induce a scale-dependent bias in the galaxy power spectrum. In this paper, we explore this particular signal for light scalar fields and study the prospects for measuring it with galaxy surveys. We find that the sensitivities of current and future surveys are remarkably stable for different configurations, including between spectroscopic and photometric redshift measurements. This is even the case at non-zero masses where the signal is not obviously localized on large scales. For realistic galaxy number densities, we demonstrate that the redshift range and galaxy bias of the sample have the largest impact on the sensitivity in the power spectrum. These results additionally motivated us to explore the potentially enhanced sensitivity of Vera Rubin Observatory’s LSST through multi-tracer analyses. Finally, we apply this understanding to current data from the last data release of the Baryon Oscillation Spectroscopic Survey (BOSS DR12) and place new constraints on light fields coupled to the inflaton.

2.2 Introduction

The statistics of the primordial density perturbations offer a window into the dynamics of the very early universe [29, 48, 49], in particular the inflationary epoch [17, 50]. While current maps of the universe are consistent with purely Gaussian fluctuations [16], non-Gaussian correlation functions encode the particles and interactions relevant to the origin of structure [9, 17, 51]. Inflation predicts a lower bound on the amount of primordial non-Gaussianity (PNG) due to gravity alone that is approximately two orders of magnitude below current constraints [52]. On the other hand, signals that are large enough to be detected in the next generation of surveys arise in many models, including examples with only Planck-suppressed interactions [53].

Due to the limited number of modes remaining to be measured to the cosmic variance limit in the cosmic microwave background (CMB), surveys of the large-scale structure (LSS) of the universe present the best opportunity to improve our understanding of the primordial statistics [17, 48, 54–56]. With the benefit of three-dimensional information, even current surveys have the raw statistical power to compete with the CMB [18]. Unfortunately, for generic non-Gaussian correlation functions, late-time nonlinearities make many modes inaccessible and present a serious obstacle to our progress [26–28, 57]. In some circumstances, however, the manifestation of the primordial signal is robust to nonlinear physics and can be measured reliably [29].

The fluctuations of extra fields beyond the inflaton present such an opportunity. In the absence of these fields, correlation functions of matter and galaxies are subject to the constraints of the single-field consistency conditions [19–21]. In practice, these conditions are similar to the equivalence principle [22] and ensure that correlations are determined by the curvature of spacetime (i.e. derivatives of the metric fluctuations). Light fields break these conditions and can lead to enhanced long-distance correlations [11–13, 58–60]. These apparent violations of the equivalence principle cannot arise from nonlinear gravity, and can therefore be distinguished from astrophysical and gravitational sources of non-Gaussianity.

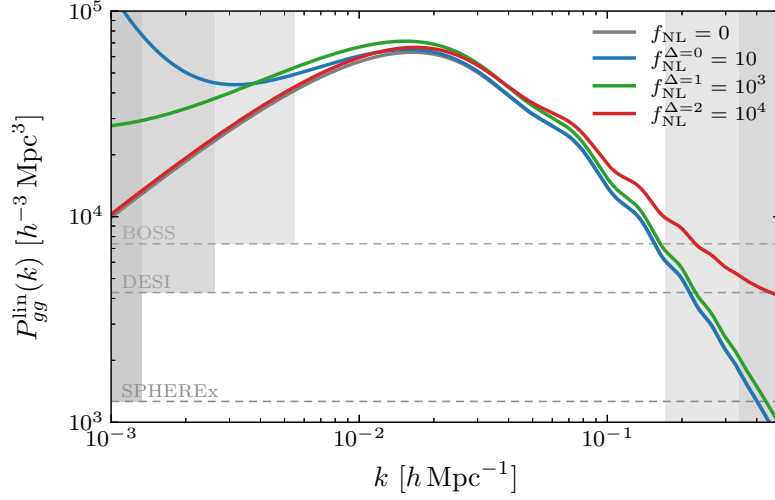


Figure 2.1. Illustration of the effect of the scale-dependent bias b_{NG}^{Δ} from a non-zero non-Gaussian amplitude f_{NL}^{Δ} on the linear galaxy power spectrum. We see that the enhancement of power shifts from large to small scales as we increase the exponent Δ (see §2.3.2 for its definition and details). The horizontal dashed lines indicate the effective shot noise level for BOSS, DESI and SPHEREx after scaling them to the displayed redshift $z = 0$ and linear bias $b_1 = 1.6$. The gray-shaded regions on large scales indicate the wavenumbers below the minimum wavenumber of these surveys, $k < k_{\text{min}}$, computed based on their entire spherical volume. The gray-shaded regions for large wavenumbers indicate the regimes where the scales exceed the nonlinear scale, $k > k_{\text{NL}}$, for the maximum redshift of BOSS and DESI, respectively.

From the perspective of both data analysis and survey design, the unique advantage of these non-Gaussian signals is that they can be observed in the two-point statistics of galaxies. Violations of the single-field consistency conditions introduce couplings between long- and short-wavelength modes. The short-wavelength modes control the number density of galaxies, which are then non-trivially correlated on large scales in the presence of this type of non-Gaussianity. The resulting (non-local) change to the shape of the galaxy power spectrum, shown in Figure 2.1, is known as scale-dependent bias [10] and has been searched for in many existing datasets (see e.g. [42, 61–69]).

The case of local primordial non-Gaussianity [70], for which the Gaussian and non-Gaussian Newtonian potentials φ and Φ are related by

$$\Phi(\vec{x}) = \varphi(\vec{x}) - f_{\text{NL}}^{\text{loc}} \varphi^2(\vec{x}) + \dots, \quad (2.1)$$

has been very well studied in the literature, both theoretically and observationally. Local PNG arises in the presence of multiple massless fields, such that φ is an isocurvature mode during inflation, but is converted to an adiabatic mode at later times [58–60]. Future measurements are expected to reach particularly interesting thresholds for the physics of multi-field inflation [17, 57, 71, 72].

The contribution to primordial non-Gaussianity from massive particles is much less studied observationally. It is however a classic signature of quasi-single-field inflation [11–13], which is also known as cosmological collider physics [14]. Importantly, these models are compelling targets for current and future surveys. In spite of this, little has been known about the optimal survey strategy to search for these relics from inflation up to now.

In this paper, we will examine how future surveys can best constrain these non-Gaussian signals using galaxy power spectra. For local PNG (massless fields), the signal is dominated by the largest scales and most observational strategies are designed accordingly [73, 74]. On the other hand, the signals of massive fields may arise at large or small scales (see Fig. 2.1), which means that the characteristics which best constrain these models are less clear a priori [40, 75, 76]. We therefore study this signal in detail, and investigate how redshift coverage, biasing, multi-tracer techniques and target selection impact forecasts over the full range of scaling laws induced by these particles.

We find that target selection, specifically finding highly biased objects, is the largest factor in driving current and future measurements of this effect when assuming a fixed universality relation. While large volumes help to increase the number of modes, especially in spectroscopic surveys, we observe that the benefits of going to higher redshifts are more driven by the large biases of high-redshift objects and only secondarily by the increased volume. In photometric surveys, the larger number density of galaxies also enables a significant increase in sensitivity through sample-variance cancellation. This requires splitting the sample according to the bias, which we explore in the context of Vera Rubin Observatory’s Legacy Survey of Space and

Time (LSST).

Using these insights, we explore current constraints on the scale-dependent bias of galaxies using the BOSS DR12 dataset [77] for the full range of scaling behaviors of light inflationary fields. We measure the contributions to the power spectrum for one scaling exponent (PNG shape) at a time and compute a correlation matrix to extrapolate between the discrete choices of scaling dimensions (masses). In terms of the amplitude f_{NL}^{Δ} of primordial non-Gaussianity (and, therefore, of the scale-dependent bias), we infer the following discrete results at 2σ :

$$f_{\text{NL}}^{\Delta} = \left\{ 1_{-39}^{+35}, 0_{-300}^{+260}, 400_{-2500}^{+2900}, 250_{-3900}^{+4000}, 200_{-5300}^{+5400} \right\} \quad \text{for} \quad \Delta = \{0.0, 0.5, 1.0, 1.5, 2.0\}, \quad (2.2)$$

where the parameter Δ is related to the mass of the fields, with definitions and details provided below. This means that our analysis is consistent with no scale-dependent bias and zero primordial non-Gaussianity for all light inflationary fields. While our inferred constraints are less sensitive to the PNG amplitude than those from Planck, the CMB constraints for non-zero masses are driven by bispectra in equilateral configurations rather than the scaling behavior in the squeezed limit probed by the galaxy power spectrum. The only previous constraint from LSS data in this regime was made in [78]. At the same time, we forecast that galaxy power spectrum measurements in future surveys will exceed the sensitivity of Planck to these light fields for a sizable range of their masses.

This paper is organized as follows: In Section 2.3, we review the relevant theoretical background for our non-Gaussian signal. In Section 2.4, we present forecasts for a wide range of model parameters and experimental configurations. Our goal is to identify the choices that most directly impact the sensitivity to PNG beyond the local type. In Section 2.5, we apply our understanding of the signal to BOSS DR12 galaxy clustering data and present new constraints on non-Gaussianity from light fields. In Section 2.6, we discuss the role of multi-tracer analyses and astrophysical effects on the forecasts. We conclude in Section 2.7. A set

of appendices contains technical details on the modeling of the galaxy power spectrum and the survey specifications (Appendix A.1), and a discussion on our ability to measure the scaling behavior of the non-Gaussian signature from the galaxy power spectrum (Appendix A.2).

2.3 Light Fields and Galaxies

Our goal in this paper is to explore the signal of additional fields that manifest themselves in the galaxy power spectrum via the scale-dependent bias. This section reviews the necessary background to understand the signal. Readers who are familiar with the scale-dependent bias due to general forms of primordial non-Gaussianity (i.e. beyond the local type) may proceed to Section 2.4.

2.3.1 Galaxy Power Spectrum

The formation of structure in the universe is driven by the growth of density fluctuations in the dark matter [15]. The evolution of the density contrast Fourier mode in redshift space, $\delta_m(\vec{k}, z) \equiv \delta\rho_m(\vec{k}, z)/\bar{\rho}_m(z)$, can be solved at linear order to give

$$\delta_m(\vec{k}, z) = \frac{2k^2 T(k) D(z)}{3\Omega_m H_0^2} \Phi(\vec{k}) \equiv k^2 \mathcal{T}(k, z) \Phi(\vec{k}), \quad (2.3)$$

where $k = |\vec{k}|$, $T(k)$ is the transfer function defined such that $T(k \rightarrow 0) \rightarrow 1$, and $D(z)$ is the linear growth factor normalized as $D(z) = 1/(1+z)$ during matter domination. The primordial Newtonian potential $\Phi(\vec{k})$ encodes the primordial density fluctuations generated during inflation which are concretely of the form

$$\langle \Phi(\vec{k}) \Phi(\vec{k}') \rangle = \frac{9}{25} \frac{A_s}{k^3} k^{n_s-1} (2\pi)^3 \delta_D(\vec{k} + \vec{k}'), \quad (2.4)$$

where δ_D is the Dirac delta function and we used $\Phi = -\frac{3}{5}\zeta$, with curvature fluctuation ζ , for modes that re-entered the horizon during matter domination. The linear matter power spectrum

is therefore given by

$$P_{\text{lin}}(k) = k^4 \mathcal{T}(k, z)^2 \frac{A_s}{k^{3-(n_s-1)}} = \frac{4}{25} \frac{A_s D(z)^2}{\Omega_m^2 H_0^4} T(k)^2 k^{n_s}. \quad (2.5)$$

The nonlinear matter power spectrum $P_m(k)$ follows the linear power spectrum on large scales, $P_m(k) \approx P_{\text{lin}}(k)$ for $k \ll k_{\text{NL}}$, for some scale $k_{\text{NL}} \approx 0.1 h \text{Mpc}^{-1}$ at $z = 0$. Gravitational evolution corrects this behavior on smaller scales, $k \gtrsim k_{\text{NL}}$.

We will generally work with galaxies in the regime $k < k_{\text{NL}}$. The nonlinear effects that are relevant in this regime can be described solely in terms of the bias expansion [25, 79]. In its most general form, the bias expansion is simply the assumption that galaxy formation is a local process:¹

$$\delta_g(\vec{x}) = \sum_i b_{\mathcal{O}_i} \mathcal{O}_i(\vec{x}), \quad (2.6)$$

where the operators $\mathcal{O}_i(\vec{x})$ are any locally measurable quantities. Since we typically assume that the matter controls the formation of galaxies, these are usually local products of $\delta_m(\vec{x})$ or the tidal tensor [80], $s_{ij} = \left(\frac{\nabla_i \nabla_j}{\nabla^2} - \frac{1}{3} \delta_{ij} \right) \delta_m(\vec{x})$,

$$\mathcal{O} \in \{ \delta_m, \nabla^2 \delta_m, \delta_m^2, s_{ij}^2, \delta_m^3, \delta s_{ij}^2, \text{Tr}[(\Pi^{[1]})^3]^{(3)}, \text{Tr}[\Pi^{[1]} \Pi^{[2]}]^{(3)}, \dots \}, \quad (2.7)$$

where the tensors $\Pi_{ij}^{[n]}$ are defined in [25].² In principle, there are an infinite number of operators to consider. At any fixed accuracy, we however understand this as an expansion in powers of the small density contrast, $\delta_m^n \ll 1$, and gradients $R^2 \nabla^2 \ll 1$, for some fixed scale R . To model the galaxy power spectrum at one-loop order, we consider the following set of operators (up to third

¹Technically, biasing is best understood as a local process in Lagrangian space. The non-locality in Eulerian space is captured by the Zel'dovich approximation [79].

²The superscript (3) indicates that the operator includes only terms up to third order in perturbation theory.

order in δ_m) in the bias expansion:

$$\begin{aligned}\delta_g(\vec{x}) = & b_1 \delta_m(\vec{x}) + b_{\nabla^2} R^2 \nabla^2 \delta_m(\vec{x}) + b_{\delta^2} (\delta_m^2(\vec{x}) - \sigma^2) + b_{s^2} (s^2(\vec{x}) - \langle s^2 \rangle) \\ & + b_{\Pi\Pi^{[2]}} \text{Tr}[\Pi^{[1]} \Pi^{[2]}]^{(3)}(\vec{x}) + \dots,\end{aligned}\tag{2.8}$$

with $\sigma^2 = \langle \delta_m^2 \rangle$. The coefficients $b_{\mathcal{O}_i} = b_1, b_{\nabla^2}, b_{\delta^2}, b_{s^2}, b_{\Pi\Pi^{[2]}}, \dots$, which we refer to as bias parameters, are constants that are determined by the details of galaxy formation and evolution. Put differently, we parameterize the complex physics of galaxies by these coefficients.

On large scales and for initial conditions set by single-field inflation, this is the complete list of operators for practical purposes on large scales. Because the fluctuations are necessarily adiabatic throughout cosmic history, the evolution of the universe is controlled by a single statistical quantity, which we have chosen to represent in terms of Φ and δ_m . On smaller scales, the evolution of different components of the universe can however lead to some new types of biasing terms.

The redshift evolution of the bias parameters, particularly $b_1(z)$, is important for projecting the sensitivity of future surveys. If (proto-)galaxies form at a high redshift and primarily evolve with the expansion of the universe, then we expect that the large-scale comoving galaxy power spectrum remains constant, $P_{gg}(k, z) \approx P_{gg}(k, z')$. In other words, the power spectrum is due to the inhomogeneities in the density field at the time they are formed, as one would find in Lagrangian biasing. This holds if

$$b_1(z) = D(z)^{-1} b_1(z=0),\tag{2.9}$$

which is a common simplifying assumption in many forecasts. Changes to the galaxy sample through mergers, for example, can however lead to behavior that differs significantly from Lagrangian biasing. When possible, we will therefore use the biases for surveys that were defined by target selection, but the assumption of this type of time evolution is also important in

the context of multi-tracer analyses discussed in Section 2.6.

Given the bias expansion for the galaxy overdensity, we can write down our model for the galaxy power spectrum after including a few additional effects. First, we work in Fourier space by transforming from real-space coordinates $\vec{x}$ to (k, μ) , where μ is the cosine between the wavevector $\vec{k}$ and the line-of-sight direction. (Note that the composite operators $\mathcal{O}_i(\vec{x})$ are simply related to $\tilde{\mathcal{O}}_i(k, \mu)$ by a single Fourier integral.) Second, there are redshift-space distortions which arise from the peculiar velocities of galaxies. We account for the Kaiser effect by adding $\Delta b(\mu, z) = f(z)\mu^2$, with the linear growth rate $f \equiv d\log D/d\log a$, to the linear bias b_1 [43]. Collecting all terms in the bias expansion at linear order in $\delta_m(\vec{k})$ therefore results in

$$b(k, \mu, z) = b_1(z) + f(z)\mu^2 + \sum_{n>0} b_{k^{2n}}(z)(kR_*)^{2n}, \quad (2.10)$$

where the last term are the gradient biases in Fourier space with the comoving Lagrangian radius R_* of the halos of interest. Third, we describe gravitational nonlinearities in the very mildly nonlinear regime via the one-loop terms P_{22} and P_{13} in standard Eulerian perturbation theory. Finally, the leading-order stochastic contribution is given by P_{ϵ_0} . Putting it all together, we therefore model the theoretical anisotropic power spectrum for two galaxy samples A and B as follows:

$$\begin{aligned} P_{AB}^{\text{th}}(k, \mu, z) = & b^A(k)b^B(k)[P_{\text{lin}}(k) + P_{22}(k) + P_{13}(k)] + P_{\epsilon_0}\delta_{AB} \\ & + \left[b^A(k)b_{\delta^2}^B + b_{\delta^2}^A b^B(k) \right] P_{\delta^2}(k) + \left[b^A(k)b_{s^2}^B + b_{s^2}^A b^B(k) \right] P_{s^2}(k) \\ & + b_{\delta^2}^A b_{\delta^2}^B (P_{\delta^2\delta^2}(k) - 2\sigma^4) + \left(b_{\delta^2}^A b_{s^2}^B + b_{s^2}^A b_{\delta^2}^B \right) (P_{\delta^2 s^2}(k) - \frac{4}{3}\sigma^4) \\ & + b_{s^2}^A b_{s^2}^B (P_{s^2 s^2}(k) - \frac{8}{9}\sigma^4) + \frac{3}{2} \left[b^A(k)b_{\Pi\Pi[2]}^B + b_{\Pi\Pi[2]}^A b^B(k) \right] P_{\Pi\Pi[2]}(k), \end{aligned} \quad (2.11)$$

where we have suppressed the dependence on μ and z on the right-hand side, explicitly define all loop and bias contributions in Appendix A.1, and δ_{AB} is the Kronecker delta. To illustrate these terms, we display the individual contributions in Fig. 2.2.

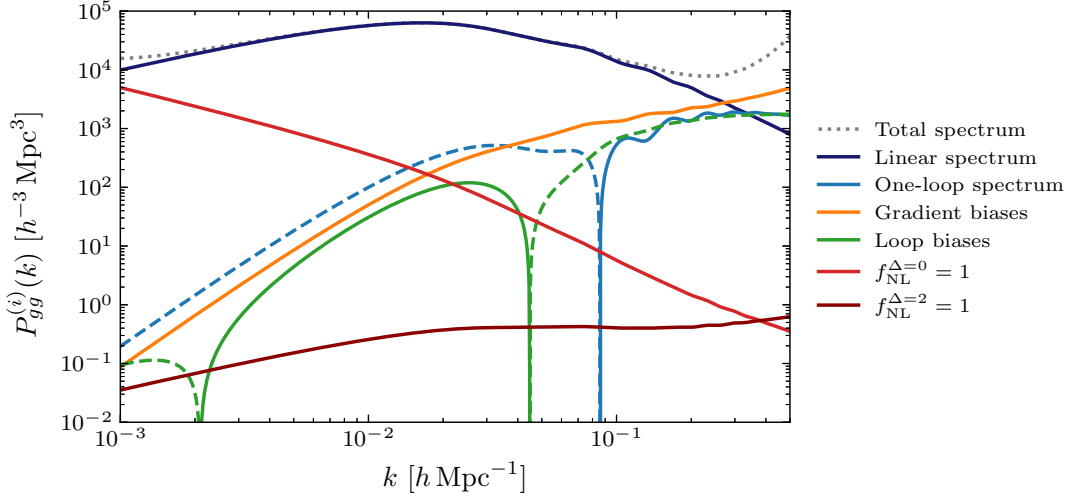


Figure 2.2. Contributions to the galaxy power spectrum $P_{gg}^{(i)}(k)$ at $z = 0$ from the nonlinear, bias and non-Gaussian terms. The linear bias is chosen to be $b_1 = 1.6$ and the gradient biases $b_{k^{2n}}$ (with $n \leq 2$) are taken to be unity. The bias parameters b_{s^2} and $b_{\text{III}[2]}$ are calculated in the Lagrangian local-in-matter-density model [25], and b_{δ^2} is calculated using the halo simulation fit of [81]. For comparison, we already include the scale-dependent bias from general PNG, which we introduce in §2.3.2. The non-Gaussian parameter f_{NL}^Δ is set to unity for both cases of $\Delta = 0$ and $\Delta = 2$. The dashed lines indicate negative contributions.

In practice, we compute the linear matter power spectrum $P_{\text{lin}}(k)$, which underlies all contributions of the theoretical galaxy power spectrum $P^{\text{th}}(k)$, with the Boltzmann solver CAMB [82] using a fiducial Λ CDM cosmological model based on the Planck 2018 TT, TE, EE + lowE + lensing + BAO best-fit cosmology [1]. In Table 2.1, we list the fiducial values of the Λ CDM, non-Gaussian and bias parameters. We set the fiducial values of the three loop biases b_{δ^2} , b_{s^2} and $b_{\text{III}[2]}$ to 0 instead of using the Lagrangian local-in-matter-density model [25] and fit results from halo simulations [81], as we did in Fig. 2.2, since we cover a larger redshift range than the fitted bias relations were derived from. This choice should have minimal impact on the results of this paper because the forecasted constraints on the loop biases are large enough for both choices of fiducial values to be consistent with each other.

Finally, to relate the theoretical galaxy power spectrum $P_{AB}^{\text{th}}(k, \mu, z)$ to the observed galaxy power spectrum $P_{AB}^{\text{obs}}(k, \mu, z)$, we account for two observational effects that are present in any real survey: redshift errors and the Alcock-Paczynski effect. Redshift errors are observational un-

Table 2.1. Parameters of the fiducial Λ CDM model, based on the Planck 2018 best-fit cosmology [1], with the sum of neutrino masses $\sum m_\nu = 0.06 \text{ eV}$, and the biasing model employed in our forecasts. Except for the optical depth, we vary these parameters in our forecasts unless stated otherwise.

Parameter	Description	Fiducial value
ω_b	Physical density of baryons $\omega_b \equiv \Omega_b h^2$	0.02242
ω_c	Physical density of cold dark matter $\omega_c \equiv \Omega_c h^2$	0.11933
$100\theta_s$	Angular size of the sound horizon at recombination	1.04119
τ	Optical depth due to reionization	0.0561
$\log(10^{10} A_s)$	Logarithm of the primordial scalar amplitude	3.047
n_s	Scalar spectral index	0.9665
f_{NL}^Δ	Non-Gaussian amplitude	0
b_1	Linear bias	$b_1(z)$
$b_{k^{2n}}$	Gradient biases ($n \leq 2$)	0
b_{δ^2}	Quadratic bias	0
b_{s^2}	Tidal bias	0
$b_{\text{III}[2]}$	Evolution bias	0

certainties in the measurement of galaxy redshifts, $\sigma_{z0}(1+z)$, which we model as an exponential suppression of the power spectrum along the line of sight,

$$f_{\sigma_z}^A(k, \mu, z) = \exp\left\{-k^2 \mu^2 [\sigma_{z0}^A(1+z)]^2 / H_{\text{fid}}^2(z)\right\}, \quad (2.12)$$

where σ_{z0}^A is the root-mean-square redshift error of sample A at $z = 0$ [83, 84]. (We neglect this effect for spectroscopic surveys by setting $\sigma_{z0}^A \equiv 0$, i.e. $f_{\sigma_z}^A = 1$.) The Alcock-Paczynski effect arises when the true cosmology differs from the fiducial cosmology that is used to convert the measured angular positions and redshifts of the galaxies to comoving wavenumbers in Fourier space. The effect of this mapping on the density contrast is captured by the following anisotropic rescaling of the power spectrum amplitude, the wavenumbers and the angles [44]:

$$f_{AP}(z) = \frac{1}{q_{\parallel} q_{\perp}^2}, \quad k'(k, \mu, z) = k \sqrt{\frac{\mu^2}{q_{\parallel}^2} + \frac{1 - \mu^2}{q_{\perp}^2}}, \quad \mu'(\mu, z) = \mu / \sqrt{\mu^2 + (1 - \mu^2) \frac{q_{\parallel}^2}{q_{\perp}^2}}, \quad (2.13)$$

where $q_{\parallel}(z) \equiv H_{\text{fid}}(z)/H(z)$ and $q_{\perp}(z) \equiv D_A(z)/D_A^{\text{fid}}(z)$, with the angular diameter distance $D_A(z)$ to redshift z . To summarize, our model for the observed galaxy power spectrum in redshift space therefore is

$$P_{AB}^{\text{obs}}(k, \mu, z) = f_{AP}(z) \sqrt{f_{\sigma_z}^A(k', \mu', z) f_{\sigma_z}^B(k', \mu', z)} P_{AB}^{\text{th}}(k', \mu', z), \quad (2.14)$$

where $P_{AB}^{\text{th}}(k, \mu, z)$ is given by (2.11).

2.3.2 Scale-Dependent Bias

The conventional description of biasing applies when the initial conditions are Gaussian. Gaussian initial conditions do not correlate modes of different scales, i.e. the collapse of small-scale density fluctuations to form halos in different regions is determined by locally observable quantities, such as the matter density, that vary over cosmological distances due to the long-wavelength fluctuations. On the other hand, mode coupling in the initial conditions due to primordial non-Gaussianity can introduce long-range correlations that are purely related to the statistics of the small-scale fluctuations themselves. When interpreted through the bias expansion, this mode coupling may appear to be non-local (scale-dependent) bias [10, 39, 41, 85] and/or long-range stochastic bias [86]. We will focus on the former since the latter is typically a subdominant contribution to the signal.

The coupling of long and short modes is described by the squeezed or collapsed limits of a non-Gaussian correlation function. In single-field inflation, the bispectrum is constrained by the single-field consistency conditions to take the following form in the squeezed limit [19, 34]:

$$\lim_{\vec{k}_1 \rightarrow 0} \langle \zeta(\vec{k}_1) \zeta(\vec{k}_2) \zeta(\vec{k}_3) \rangle = -[(n_s - 1) + O(k_1^2/k_3^2)] P_{\zeta}(k_1) P_{\zeta}(k_3) (2\pi)^3 \delta_D\left(\sum_i \vec{k}_i\right). \quad (2.15)$$

Furthermore, the leading term, which corresponds to $f_{\text{NL}}^{\text{loc}} = -\frac{5}{12}(n_s - 1)$, is unphysical [37, 87]. The leading physical term is suppressed by $O(k_1^2/k_3^2)$ and is due to the coupling of the short modes at horizon crossing to the curvature of the universe due to long-wavelength modes [88],

which is typically proportional to the amplitude of equilateral non-Gaussianity, $f_{\text{NL}}^{\text{eq}}$.

In the presence of additional fields, the short-scale power can depend on the long-wavelength values of these fields and not just derivatives of the metric fluctuations [89]. For an additional massless field, the late-time Newtonian potential Φ may be nonlinearly related to a light (isocurvature) field χ during inflation, $\Phi(\vec{x}) = \chi(\vec{x}) - f_{\text{NL}}^{\text{loc}} \chi^2(\vec{x})$. For a massive field, we have in the super-horizon limit

$$\lim_{\vec{k}_1 \rightarrow 0} \chi(\vec{k}_1, t) \approx c \frac{H}{k_1^{3/2}} \left(\frac{k_1}{aH} \right)^\Delta, \quad (2.16)$$

with a constant c . If χ alters the power spectrum of the short modes at horizon crossing, $k_2 = aH$, then we have a contribution to the bispectrum of the form

$$\lim_{\vec{k}_1 \rightarrow 0} \langle \zeta(\vec{k}_1) \zeta(\vec{k}_2) \zeta(\vec{k}_3) \rangle' = f_{\text{NL}}^\Delta \left(\frac{k_1}{k_2} \right)^\Delta P(k_1) P(k_2), \quad (2.17)$$

where we defined $\langle \dots \rangle = (2\pi)^3 \delta_D(\sum_i \vec{k}_i) \langle \dots \rangle'$. To have such a contribution to the bispectrum, we also need χ to mix with ζ . Even without a mixing term, this same effect however appears in the collapsed limit of the trispectrum. For light fields, $\Delta = 3/2 - \sqrt{9/4 - m^2/H^2}$ such that $m = 0$ leads to local PNG ($\Delta = 0$) as expected. For $m/H > 3/2$, $\Delta = \frac{3}{2} \pm i\nu$, with $\nu = \sqrt{m^2/H^2 - 9/4}$, is complex in which case the complex conjugate³ also contributes which results in the bispectrum being real valued, as required. More generally, it is possible to construct models with a wider range of real and complex values of Δ than those generated by a single massive field [23, 24, 90, 91].

In single-field inflation, physical mode coupling must involve derivatives of the metric fluctuations [37], whose leading behavior in the soft limit corresponds to $\Delta = 2$ [38]. One might naturally be surprised to find that the $m^2 \rightarrow \infty$ limit of Δ for a massive field does not

³For heavy fields, there are two solutions of the long-wavelength behavior which scale as Δ and its complex conjugate Δ^* . We note that there is also a second solution $\Delta' = (3 - \Delta)$ for light fields which is however suppressed relative to Δ .

yield this value of single-field inflation. It was nicely explained in [14] that $\Delta \approx 3/2 + im/H$ is precisely what is expected from the single-particle wavefunction of a massive particle in an expanding universe in the high-mass limit. This contribution to the squeezed limit arises from the physical production of these particles. This particle production is Boltzmann suppressed in some simple models [92], $f_{\text{NL}}^\Delta \propto e^{-\pi m/H}$, but can be enhanced depending on the nature of the interaction [93]. In addition, very massive fields will also produce a sub-leading contribution to the squeezed limit with $\Delta = 2$. This contribution arises from the virtual exchange of a massive field and is equivalent to a purely local interaction (i.e. to integrating out the massive field). It may be surprising that a local term will produce non-local scale-dependent bias. When $T(k) \rightarrow 1$, $\Delta = 2$ is indeed a local term and it was long thought that $f_{\text{NL}}^{\text{eq}}$ could not be measured via scale-dependent bias for this reason (see e.g. [39]). However, because of the evolution of matter after horizon entry ($T(k) \neq 1$), local interactions during inflation are distinguishable from local processes in structure formation at late-times [40], i.e. equilateral PNG can be measured in the power spectrum via scale-dependent bias.

In the conventional picture of galaxy biasing (see §2.3.1), the density contrast of galaxies is determined by a long list of composite operators in the initial density field. The origin of the unique signal of PNG in galaxies is that the behavior of these composite operators becomes a proxy for the light fields during inflation. For example, consider the impact of structure formation from the local amplitudes of fluctuations on a scale R ,

$$\sigma_R(\vec{x}) = \int d^3k e^{-i\vec{k}\cdot\vec{x}} \int \frac{d^3p}{(2\pi)^3} \zeta(\vec{p}) \zeta(\vec{k} - \vec{p}) F(pR), \quad (2.18)$$

for some filter function $F(x)$. This is the local variance of the primordial fluctuations which varies from place to place. Importantly, this variance is correlated with the long-wavelength metric perturbations as follows:

$$\langle \sigma_R(\vec{k}) \zeta(\vec{k}') \rangle' \propto f_{\text{NL}}^\Delta (kR)^\Delta P_\zeta(k) \quad \Rightarrow \quad \langle \sigma(\vec{k}) \delta_m(\vec{k}') \rangle' \propto f_{\text{NL}}^\Delta \frac{(kR)^\Delta}{k^2 T(k)} P_m(k). \quad (2.19)$$

Since the number density of galaxies at a location $\vec{x}$ is related to $\sigma_R(\vec{x})$, this introduces a long-distance correlation between $\delta_g(\vec{x})$ and δ_m that is non-local. The resulting non-local modification of the bias (2.10) is given by [86]⁴

$$b(k, \mu, z) = b_1(z) + f(z)\mu^2 + \sum_{n>0} b_{k^{2n}}(kR_*)^{2n} + Af_{\text{NL}} \frac{b_\phi(z)}{k^2 \mathcal{T}(k, z)} (kR_*)^\Delta, \quad (2.20)$$

where b_ϕ is the non-Gaussian bias parameter and we took $R = R_*$. The constant A normalizes the non-Gaussian parameter f_{NL}^Δ in the conventional way for PNG in the squeezed limit of the bispectrum,⁵ so that the last term in (2.20) is [41, 94–97]

$$b_{\text{NG}}^{\text{loc}}(k, z) = f_{\text{NL}}^{\text{loc}} \frac{b_\phi(z)}{k^2 \mathcal{T}(k, z)} \quad (\text{local}), \quad (2.21)$$

$$b_{\text{NG}}^{\text{eq}}(k, z) = 3f_{\text{NL}}^{\text{eq}} \frac{b_\phi(z)}{k^2 \mathcal{T}(k, z)} (kR_*)^2 \quad (\text{equilateral}), \quad (2.22)$$

$$b_{\text{NG}}^\Delta(k, z) = 3f_{\text{NL}}^\Delta \frac{b_\phi(z)}{k^2 \mathcal{T}(k, z)} (kR_*)^\Delta \quad (\text{general exponent } \Delta \in [0, 2]). \quad (2.23)$$

Note that we do not include the physical orthogonal shape, as it is produced in single-field inflation and, therefore, has the same squeezed limit as the equilateral shape, i.e. corresponds to $\Delta = 2$. Previous work on the scale-dependent bias used an orthogonal template which has $\Delta = 1$ (e.g. [76]), which is however unphysical and has not been used in recent LSS analyses.

Without any additional theoretical input, $b_\phi(z)$ and f_{NL}^Δ are both unknown quantities and are completely degenerate. If that was the case, we could only measure the combination $f_{\text{NL}}^\Delta b_\phi(z)$ in each redshift bin which would make it challenging to infer limits on f_{NL}^Δ alone. More optimistically, we can fix $b_\phi(z)$ in terms of the linear bias $b_1(z)$ using the universality relations

⁴It is important to note that, in principle, the scale-dependent bias is distinguishable from the expansion in (kR_*) even for $\Delta = 2n$ for any positive integer n . At $k \ll k_{\text{eq}}$, $\mathcal{T}(k) \rightarrow \text{const}$ and the scale-dependent bias scales as k^{2n-2} which is degenerate with b_1 and $b_{k^{2(n-1)}}$. However, since $k_{\text{eq}}R_* \ll 1$, the transfer function introduces a non-trivial scale dependence in the regime $k_{\text{eq}} \lesssim k \ll R_*^{-1}$ which is not captured in the local bias expansion.

⁵Note that for the special bispectrum configurations, when expressed in terms of the general exponent Δ , the equivalent f_{NL}^Δ definitions are related by $f_{\text{NL}}^{\text{loc}} = 3f_{\text{NL}}^{\Delta=0}$ and $f_{\text{NL}}^{\text{eq}} = f_{\text{NL}}^{\Delta=2}$.

found in phenomenological models [25],

$$b_\phi(z) = 2\delta_c(b_1(z) - p), \quad (2.24)$$

where $\delta_c = 1.686$ is the critical overdensity for spherical collapse and $p = 1$ if we assume a universal halo mass function. We will adopt this relation in our forecasts, but refer to e.g. [98–101] for a recent discussion of the validity and risks of this choice.

To conclude, we note that we incorporate the general scale-dependent bias (2.20) in the galaxy power spectrum as described in (2.11) and (2.14), i.e. $b^A = b_\Delta^A(k, \mu, z)$. In the rest of the paper, we follow [40] in assuming a minimum halo mass of $10^{13} M_\odot$, which corresponds to $R_* \approx 2.66 h^{-1} \text{ Mpc}$, and truncating the non-local gradient bias expansion at $n = 2$. We illustrate the effect of b_{NG}^A on the linear galaxy power spectrum in Fig. 2.1 and compare these non-Gaussian contributions to the nonlinear galaxy power spectrum in Fig. 2.2 at redshift $z = 0$ for linear bias $b_1 = 1.6$. We see from the latter figure that the $\Delta = 0$ contribution does not mimic the behavior of any other bias or nonlinear terms at low k . On the other hand, the $\Delta = 2$ contribution is both much smaller at $k < k_{\text{NL}}$ and similar in behavior to those Gaussian contributions.

2.4 Information Content of the Galaxy Power Spectrum

Our goal in this section is to understand how to design a galaxy survey to best measure f_{NL}^A for $\Delta \in [0, 2]$. To do so, we will explore the Fisher forecasts for f_{NL}^A as a function of Δ for a variety of survey configurations and biasing models.⁶ We will compare these results with analytic estimates to understand what aspects of the surveys drive major improvements in sensitivity. While our forecasts will be rooted in specific survey configurations, our goal is to isolate the features of these surveys that impact the forecasted constraining power the most.

⁶In Appendix A.2, we also discuss our ability to measure the scaling behavior of the non-Gaussian bias, where we also include Δ as an additional free parameter in the Fisher matrix.

2.4.1 Fisher Information Matrix

In the following, we describe the well-known Fisher matrix forecasting techniques for the galaxy power spectrum that we use. We review the basics here and refer to Appendix A.1 for the experimental specifications of the surveys.

The Fisher matrix for a set of galaxy samples, labeled by $\{A, B, \dots\}$, is defined by

$$F_{\alpha\beta} = \sum_{z_i} V(z_i) \int_{-1}^1 \frac{d\mu}{2} \int_{k_{\min}}^{k_{\max}} \frac{dk k^2}{2\pi^2} \frac{1}{2} \text{Tr} [\mathbf{C}_{,\alpha} \mathbf{C}^{-1} \mathbf{C}_{,\beta} \mathbf{C}^{-1}], \quad (2.25)$$

where $\theta_{\alpha,\beta}$ are the model parameters, the trace is over the samples and we defined $\mathbf{C}_{,\alpha} \equiv \partial \mathbf{C} / \partial \theta_{\alpha}$. The integrals are discretized, with the integration over wavenumbers k being limited by the minimum wavenumber $k_{\min}$ and the maximum wavenumber $k_{\max}$, and bins of width $\Delta k = k_{\min}$. We note that we vary all parameters listed in Table 2.1 throughout this paper, unless specified otherwise. The elements of the matrix $\mathbf{C}$ are given by

$$\mathbf{C}_{AB}(k, \mu, z) \equiv P_{AB}^{\text{obs}}(k, \mu, z) + \delta_{AB} N_A(k, \mu, z), \quad (2.26)$$

with the shot noise term $N_A(k, \mu, z) = \bar{n}_A(z)^{-1}$, where $\bar{n}_A(z)$ is the average comoving number density of tracers at redshift z . Shot noise arises because we observe a finite number of discrete objects in galaxy surveys. (Note that we absorbed the stochastic term $P_{\epsilon 0}$ of (2.11) into N_A .) We treat the redshift bins as spherical shells so that the volume of the i -th redshift bin is

$$V(z_i) = \frac{4\pi}{3} f_{\text{sky}} [d_c(z_i + \Delta z_i/2)^3 - d_c(z_i - \Delta z_i/2)^3], \quad (2.27)$$

where f_{sky} is the sky fraction observed by the survey, $d_c(z)$ is the comoving distance to redshift z and Δz_i is the width of the i^{th} redshift bin.

To ensure that only modes under perturbative control in Eulerian perturbation theory and the bias expansion enter our Fisher matrix calculation, and to exclude long-wavelength

modes inaccessible to a given survey due to its finite volume, we restrict the integral in (2.25) to a maximum wavenumber $k_{\max}$ and a minimum wavenumber $k_{\min}$, respectively.⁷ We make a conservative choice for $k_{\max}$ by taking the smaller of the two scales indicating the range of validity of perturbation theory and the bias expansion. To be precise, we take it to be the minimum of the nonlinear scale⁸ k_{NL} and the wavenumber $k_{\text{halo}} \approx 0.19 h \text{Mpc}^{-1}$ associated with the Lagrangian scale corresponding to the minimum halo mass of the surveyed population which we assume to be $R_* \approx 2.66 h^{-1} \text{Mpc}$ (cf. §2.3.2), i.e. $k_{\max}(z) = \min\{k_{\text{NL}}(z), k_{\text{halo}}\}$. For $k_{\min}$, we assume a spherical survey geometry per redshift bin which results in a conservative minimum wavenumber $k_{\min,i} = 2\pi[3V(z_i)/(4\pi)]^{-1/3}$.

For most of the forecasts in Sections 2.4 and 2.6, we use a futuristic spectroscopic survey with a billion objects divided into 10 redshift bins with $z \leq 5$, which we will hereafter refer to as the *billion-object survey*. We split the galaxy sample of this survey into two populations based on their linear bias at $z = 0$, chosen as $b_1^{(1)}(z = 0) = 2.0$ and $b_1^{(2)}(z = 0) = 1.2$. We model their redshift evolution according to (2.9) as previously discussed in §2.3.1. In addition, we assume a constant number density across all redshift bins and both galaxy samples, $\bar{n}_g^{(1)}(z) = \bar{n}_g^{(2)}(z) = \text{const}$. The details of this and other surveys used in our forecasts, i.e. the Baryon Oscillation Spectroscopic Survey (BOSS) [77], the Dark Energy Spectroscopic Instrument (DESI) [104], Euclid [105, 106], Vera Rubin Observatory’s LSST [107], the Spectro-Photometer for the History of the Universe, Epoch of Reionization and Ices Explorer (SPHEREx) [71], MegaMapper [55, 108] and the billion-object survey are provided in Appendix A.1. Except for SPHEREx and the billion-object survey, we treat all surveys as single-tracer surveys by combining different groups of tracers (if available) into a single effective number density and bias. The impact of this should not significantly impact our forecasts as we explain in Appendix A.1.

Unless mentioned otherwise, we marginalize over all bias parameters separately for

⁷Alternatively, we could have included the correlations between k bins for small wavenumbers and the theoretical uncertainties following [26, 102, 103] for large wavenumbers. Our approximate and conservative treatment using $k \in [k_{\min}, k_{\max}]$ follows the common procedure in Fisher forecasts.

⁸We use $k_{\text{NL}} = \frac{\pi}{2R_{\text{NL}}}$, with the radius R_{NL} at which the variance of linear fluctuations is $\sigma_R = 1/2$ at a given z .

each sample and redshift bin, and include CMB information on the Λ CDM parameters (but not on PNG) from Planck throughout the paper.⁹ We achieve this by computing the Λ CDM Fisher matrix directly following the Fisher methodology and the experimental specifications of [111] (see also [112] for the latter), marginalized over the optical depth τ . We then combine this marginalized Planck Fisher matrix with the respective LSS Fisher matrix for the five Λ CDM parameters, the non-Gaussian parameter f_{NL}^{Δ} and the galaxy bias parameters $b_{\mathcal{O}_i}$. We treat each redshift bin with mean redshift $\bar{z}_i$ as an independent survey and independently marginalize over the biases within each redshift bin.

2.4.2 Information and Survey Design

Using Fisher forecasts, our aim is to understand the possible strategies for optimizing the measurement of f_{NL}^{Δ} for shapes including but not limited to local PNG. The parameters that we most directly control in the design of a survey are the sky fraction (f_{sky}), the redshift range ($z \in [z_{\text{min}}, z_{\text{max}}]$) and the number density of tracers $[\bar{n}_g(z)]$. However, the actual statistical power of the survey is controlled by the smallest (k_{min}) and largest (k_{max}) wavenumbers that can be reliably measured, the linear bias of the sample $[b_1(z)]$, etc. These factors are influenced by the details of the survey, but can also be affected by systematics and astrophysics.

We show the overall landscape of future power spectrum measurements of f_{NL}^{Δ} for $\Delta \in [0, 2]$ in Fig. 2.3. Qualitatively, the forecasts show the overall behavior that we might expect: (i) constraints at $\Delta = 0$ are generally much stronger than at $\Delta = 2$, (ii) higher-order biasing affects larger Δ more than smaller Δ , (iii) larger surveys have more constraining power, and (iv) future LSS surveys will be able to improve over Planck bispectrum constraints from the galaxy power spectrum for $\Delta = 0$, but not for $\Delta = 2$ without additional information on bias parameters. In detail, the forecasts however have elements that are harder to understand

⁹When marginalizing over the full biasing model for the billion-object survey, for instance, the difference between using the Λ CDM covariance from Planck or CMB-S4 [109, 110] for $\Delta \lesssim 1.5$ is less than two percent. In fact, this CMB information decreases $\sigma(f_{\text{NL}}^{\Delta})$ only by a few percent for $\Delta \lesssim 1.3$, and by a maximum of 20% (15%) with Planck (CMB-S4) at $\Delta = 2$.

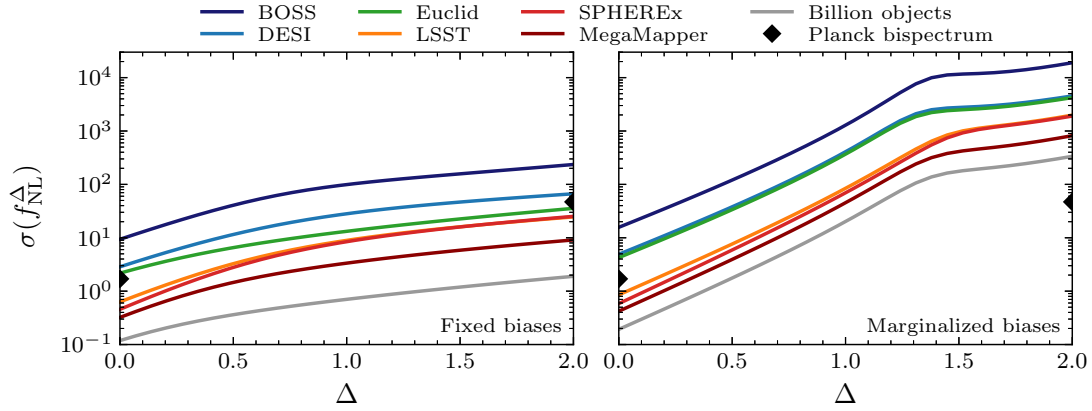


Figure 2.3. Forecasted constraints on f_{NL}^Δ for general scaling $\Delta \in [0, 2]$ of the scale-dependent bias, for current and future surveys, compared to the Planck bispectrum constraints for local and equilateral non-Gaussianity. We also marginalize over the five Λ CDM parameters, use Planck results as priors for the Λ CDM parameters, and either fix the bias parameters (*left*) or marginalize over them (*right*). Note that the SPHEREx forecast uses the five-tracer sample and the billion-object survey forecast uses the double-tracer sample as provided in Appendix A.1. For all the other surveys, we use their combined single-tracer sample.

without further investigation. First, the forecasts show clear features at values of $\Delta \sim 1.4$ that depend on the marginalization over biasing parameters (see also Fig. 2.6 below). This suggests a qualitative change in where the constraining power is coming from as we vary Δ . Second, we see that SPHEREx and LSST produce very similar forecasts despite having very different strengths and weaknesses. Furthermore, both significantly exceed the constraining power of Euclid and DESI which have the advantage of being three-dimensional (spectroscopic) surveys. In order to make sense of these forecasts, we will break them down according to the number density (noise), scales ($k_{\text{min}}/k_{\text{max}}$), the fiducial bias [$b_1(z)$], the survey geometry ($f_{\text{sky}}, z_{\text{max}}$) and redshift errors (σ_{z0}) in the following.

Dependence on Number Density

The most basic parameter in any experiment is the signal to noise. In a galaxy survey, shot noise is the dominant (irreducible) noise source. At fixed volume, increasing the number of objects increases the range of wavenumbers k where $P_{gg}(k)$ is measured with at least signal to noise of order one, $P_{gg}(k)/N(k) \gtrsim O(1)$.

For most cosmological parameters, sample variance presents an additional irreducible source of noise. However, for scale-dependent bias, the parameter f_{NL}^{Δ} can, in principle, be measured without sample variance when the signal to noise is large for multiple tracers [113]. Concretely, the relationship between $\delta_g(\vec{k})$ and $\delta_m(\vec{k})$ in (2.8) is completely deterministic and is therefore not limited by sample variance. However, this only works if we can measure the same Fourier mode with high signal to noise for tracers with different values of b_{ϕ} . The noise introduced by sample variance is proportional to $P(k)$, while shot noise is independent of k . As a result, sample-variance cancellation dramatically alters how different scales contribute to the measurement of f_{NL}^{Δ} and, therefore, the qualitative understanding of our forecasts. This is why it is important to distinguish from the outset to what degree multi-tracer sample-variance cancellation is relevant in our forecasts.

The impact of shot noise in both single- and multi-tracer scenarios is illustrated in Fig. 2.4. Two points should be very clear from these curves: (1) sample-variance cancellation has the potential to dramatically improve the sensitivity to f_{NL}^{Δ} for all values of Δ , and (2) realistic current and future surveys are very far from being in the regime where sample-variance cancellation has a large impact. In fact, most near-future surveys are, at best, at the boundary between single and multiple tracers.¹⁰ As a result, sample-variance cancellation is important for highly accurate forecasts, but the qualitative behavior is consistent with single-tracer forecasts. We therefore use a single-tracer emphasis for the rest of this section. We will revisit the advantages offered by multi-tracer techniques and how astrophysics affects the constraints on f_{NL}^{Δ} in Section 2.6.

Dependence on Scales

The defining feature of scale-dependent bias is that it is enhanced at small wavenumbers/large distances. For $\Delta < 2$ and $f_{\text{NL}}^{\Delta} \neq 0$, we have $P_{gg}(k) \gg b_1^2 P_m(k)$ as $k \rightarrow 0$. Since the

¹⁰Note that the curves in Fig. 2.4 are computed using the spectroscopic billion-object survey (cf. Table A.6). This means that the photometric LSST is not a direct comparison and the line labeled as LSST only actually indicates a spectroscopic follow-up of LSST. The potential sensitivity in a multi-tracer analysis of LSST itself will be discussed in §2.6.2.

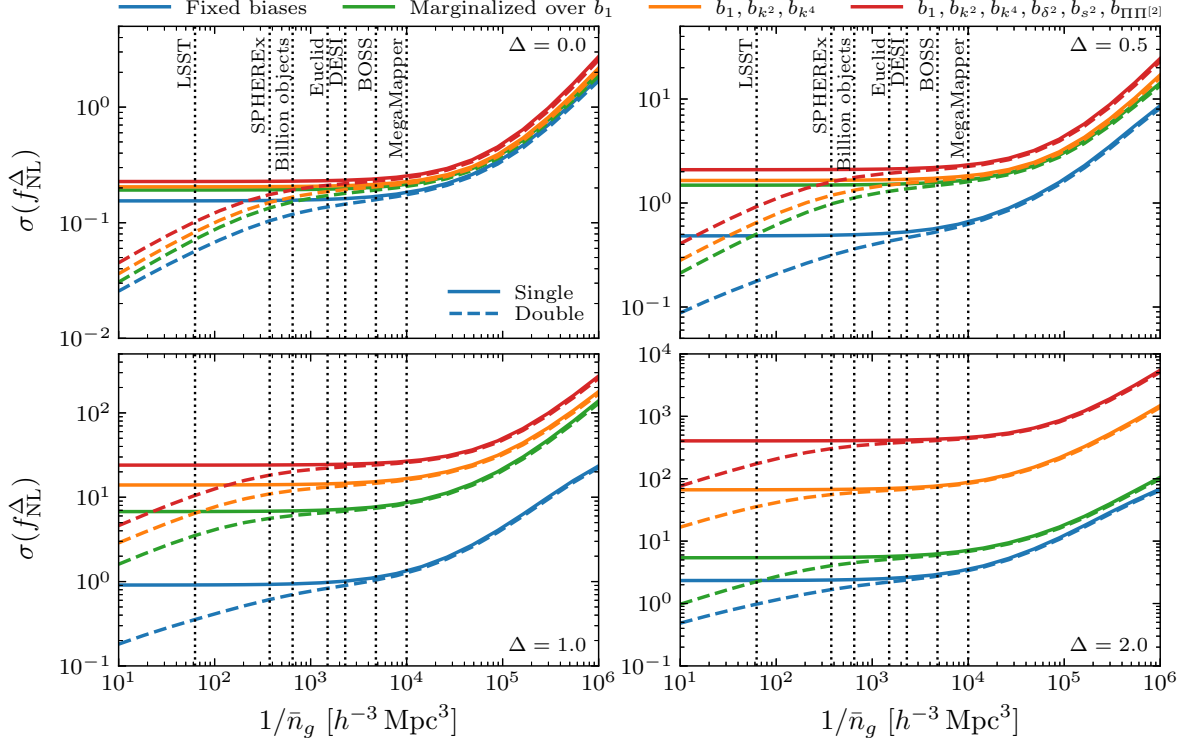


Figure 2.4. Dependence of $\sigma(f_{\text{NL}}^\Delta)$ on the shot noise, i.e. the inverse survey number density $\bar{n}_g^{-1}$, for the billion-object survey for four representative values of Δ . The solid lines show the single-tracer forecasts and the dashed lines display the double-tracer results. The different colors indicate which bias parameters are marginalized when calculating $\sigma(f_{\text{NL}}^\Delta)$. The vertical lines mark the averaged shot noise of various surveys which are calculated by taking the average of the number densities of the middle 50% of redshift bins of each survey, i.e. they should be taken as an approximate illustration of the noise level in these surveys.

noise is also a function of k , it is however not a given that the information resides at small k for all $\Delta < 2$. We now investigate this analytically in the Fisher matrix and numerically in our forecasts.

For a single-tracer analysis where we know the bias parameters exactly, i.e. $b_{\mathcal{O}_i}$ are held fixed, the Fisher information for a fiducial $f_{\text{NL}}^\Delta = 0$ is given by

$$F_{f_{\text{NL}}^\Delta f_{\text{NL}}^\Delta} = \sum_{z_i} V(z_i) \int \frac{d^3k}{(2\pi)^3} \frac{[6b_1 b_\phi(z_i) (kR_*)^\Delta]^2 P_m(k, \mu, z_i)^2}{2k^4 \mathcal{T}(k, z_i)^2 [b_1^2 P_m(k, \mu, z_i) + N(k, \mu, z_i)]^2}. \quad (2.28)$$

In the high signal-to-noise regime, $b_1^2 P_m(k) \gg N$, this becomes

$$F_{f_{\text{NL}}^\Delta f_{\text{NL}}^\Delta} \approx \frac{9}{\pi^2} \sum_{z_i} \frac{b_\phi^2(z_i)}{b_1^2(z_i)} R_*^{2\Delta} V(z_i) \int dk \frac{k^{2\Delta-2}}{\mathcal{T}(k, z_i)^2}. \quad (2.29)$$

For $2\Delta < 1$, the integral over k is dominated by $k_{\min}$ and the Fisher information arises from the largest scales. However, due to the limits placed by noise and volume, realistic surveys are far from the asymptotic $k \rightarrow 0$ regime. For modes with the largest signal to noise, $k \approx k_{\text{eq}} \approx 10^{-2} h \text{Mpc}^{-1}$, the matter power spectrum is flat, $P_m(k \sim k_{\text{eq}}) \propto k^0$ and, therefore, $\mathcal{T}(k, z)^2 \sim k^{-1}$. In this regime, repeating our analysis of (2.29) shows that the integral is dominated by large k for $\Delta \gtrsim 0$. In this regard, we expect that the transition from low to high k is gradual as we change Δ .

This behavior is seen in forecasts by varying $k_{\min}$ as illustrated in Fig. 2.5. Since increasing the minimum wavenumber $k_{\min}$ removes information, the constraints $\sigma(f_{\text{NL}}^\Delta)$ always decrease. We therefore show $d\sigma^{-2}(f_{\text{NL}}^\Delta)/d\log k_{\min}$ because this quantity provides a more quantitative illustration of how much information is being lost as we change $k_{\min}$.¹¹ For $\Delta = 0$, the largest impact is for the smallest values of $k_{\min}$, indicating that the lowest wavenumber has the most information. As we increase $\Delta > 0$, we see that this trend changes: the curves are mostly flat at small $k_{\min}$ for $\Delta \lesssim 0.5$. This is a reflection of the fact that there is no sharp transition in Δ

¹¹We consider the logarithmic derivative in order to numerically extract the relevant scaling that we analytically discussed in the previous paragraph. We additionally note that $\log k_{\min}$ is implicitly normalized by $1 h \text{Mpc}^{-1}$.

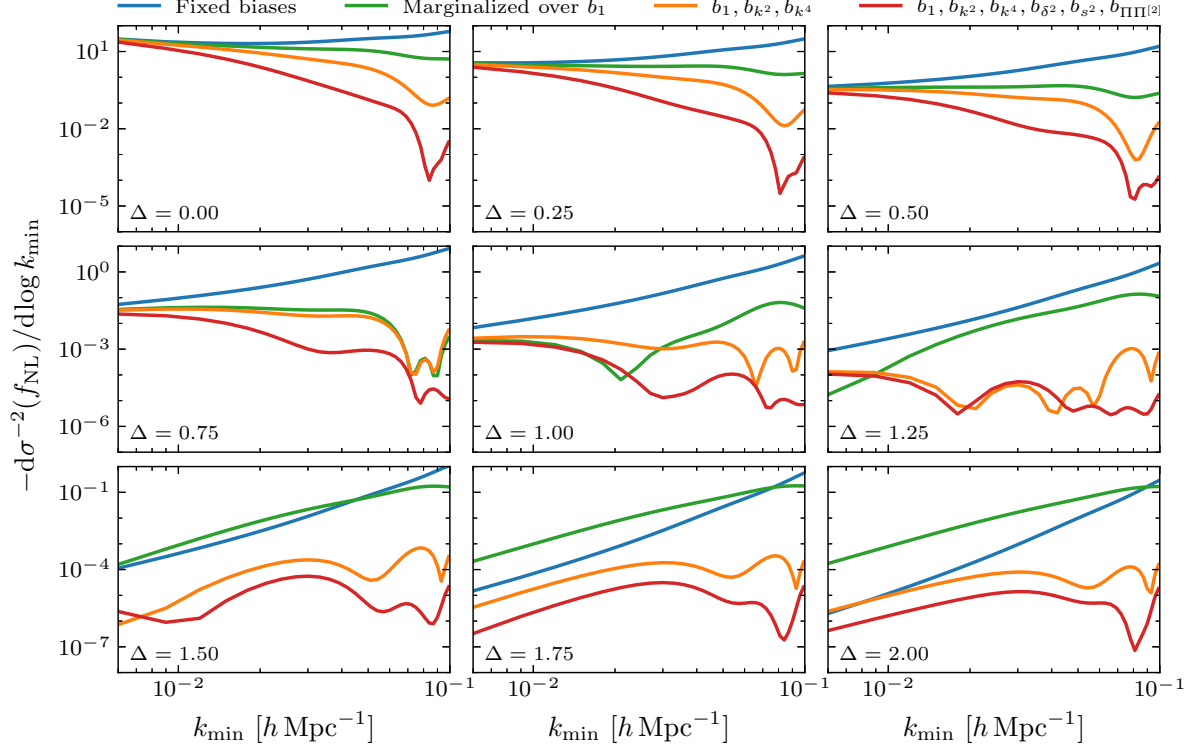


Figure 2.5. The information density with respect to $\log k_{\min}$ in the billion-object survey, with fixed Λ CDM cosmology, and the same $k_{\min}$, $\Delta k = 0.003 h \text{Mpc}^{-1}$ and $k_{\max} = 0.2 h \text{Mpc}^{-1}$ in each redshift bin. The different colors indicate which bias parameters are marginalized when calculating $\sigma(f_{\text{NL}}^{\text{A}})$ using the double-tracer version of the survey. A negative (positive) slope of the curves indicates that there is more (less) information on smaller wavenumbers k . We observe that the information density is dominated by small (large) wavenumbers for small (large) scaling exponents.

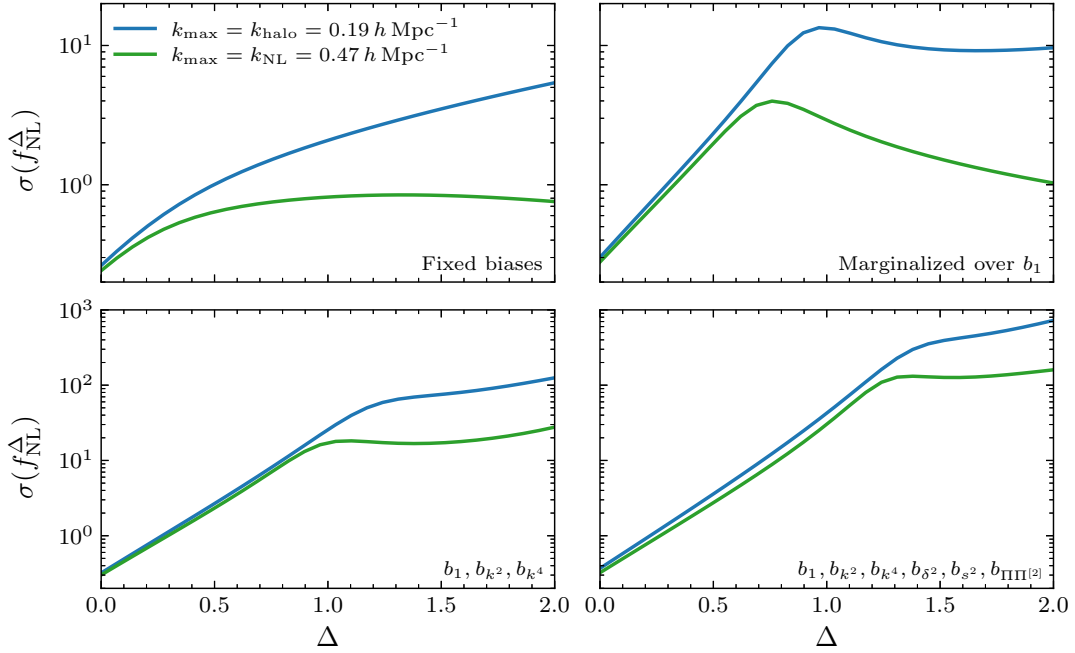


Figure 2.6. Dependence of $\sigma(f_{\text{NL}}^\Delta)$ on the maximum wavenumber where we take $k_{\text{max}} = k_{\text{halo}}$ and $k_{\text{max}} = k_{\text{NL}}(z)$, with the latter being less conservative at higher redshifts. Here, we combine the middle two redshift bins ($2 \leq z \leq 3$) of the double-tracer billion-object survey into a single comoving box with a total of about 1.3×10^8 objects. We show this dependence of the constraining power for four biasing models. The choice of k_{max} barely impacts $\sigma(f_{\text{NL}}^\Delta)$ at small Δ , but becomes significant for large Δ . We explain the origin of the feature around $\Delta = 1$ in the main text.

due to the impact of $\mathcal{T}(k, z)$ in (2.29).

As we move from large to small scales (or small to large k), we are increasingly sensitive to the marginalization over the bias parameters. The information density in Fig. 2.5 at large wavenumbers becomes highly suppressed as we marginalize so that the true constraining power remains at small wavenumbers. The oscillatory behavior in these figures highlights that the baryon acoustic oscillations contained in the transfer function are imprinted in the non-Gaussian signal and are not absorbed into the bias expansion.

At the same time, this short-distance information remains however crucial for breaking degeneracies between bias and cosmological parameters. This is shown in Fig. 2.6, which illustrates how the maximum wavenumber k_{max} impacts the constraining power. We see that the constraint on f_{NL}^Δ improves by roughly a factor of five to nine (depending on the biasing model)

for $\Delta = 2$ when we choose the less conservative $k_{\max} = k_{\text{NL}}$. This improvement is consistent with the number of additional modes available between k_{halo} and k_{NL} , and the fact that our signal depends on wavenumbers. This difference becomes even more pronounced at higher redshifts since $k_{\text{NL}}(z)$ increases significantly, while k_{halo} is independent of redshifts. For all of our other forecasts, we however use the more conservative choice of $k_{\max}(z) = \min\{k_{\text{halo}}, k_{\text{NL}}(z)\}$ as noted in §2.4.1.

When marginalizing over b_1 , we notice a feature around $\Delta \approx 1$ in Fig. 2.6 and not around $\Delta = 0.5$ as we might expect from (2.29). This is a consequence of the information on f_{NL}^Δ moving to larger k where $T(k) \neq 1$. Specifically, if we assume $P_m(k) \propto k^n$ for large k , then $k^2 T(k) \propto k^{n/2+3/2}$, which implies

$$P_{gg}(k) \approx \left(b_1 + 2b_\phi f_{\text{NL}}^\Delta C(kR_*)^{\Delta - \frac{n}{2} - \frac{3}{2}} + \dots \right)^2 P_m(k), \quad (2.30)$$

for some constant C . Using $n \approx -1$ for $k \sim 0.1 h\text{Mpc}^{-1}$ [114], we notice that $\Delta - \frac{3}{2} - \frac{n}{2} \approx \Delta - 1$. At large wavenumbers, the “scale-dependent bias” for $\Delta \approx 1$ is therefore degenerate with the linear bias b_1 , resulting in the observed feature.

Dependence on Fiducial Biases

All else being equal, a single-tracer survey with the largest possible (absolute) value of $b_\phi \propto (b_1 - 1)$ will yield the most sensitive measurement of f_{NL} . It may therefore seem self-evident that selecting highly biased targets is a central tool in the search for PNG. Having said that, in practice, target selection involves numerous factors that lie beyond the scope of these simple forecasts. Yet, when it comes to understanding the performance of a given survey, the biases of the objects in their sample will strongly influence the overall sensitivity. As a result, it is important to separately understand the role that the fiducial biases play from the aspects of the survey that we can control more directly.

For a single-tracer analysis, we only measure $b_\phi f_{\text{NL}}^\Delta$. At fixed redshift and assuming

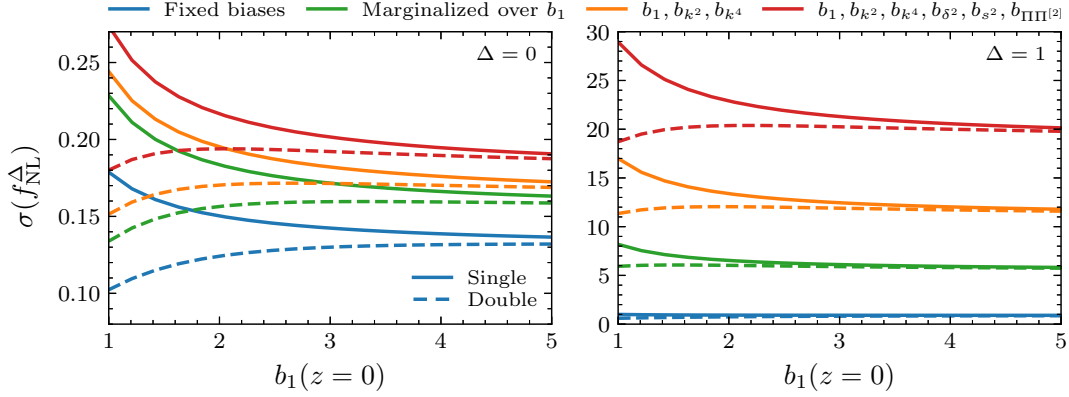


Figure 2.7. Dependence of $\sigma(f_{\text{NL}}^\Delta)$ on the fiducial value of the linear bias for both the single-tracer (solid) and double-tracer (dashed) versions of the billion-object survey for $\Delta = 0$ (left) and $\Delta = 1$ (right). We show this dependence as a function of the linear bias $b_1(z=0)$ of the single-tracer survey in which case we see the improvement in the constraining power of the survey for larger b_1 . In the double-tracer scenario, for which the galaxy sample is split into two populations of equal number density with $b_1^{(A,B)}(z=0) = b_1(z=0) \pm 0.4$, the effects of sample-variance cancellation can be observed. (Note that the y-axis is on a linear and not on a logarithmic scale like in the previous figures.)

$b_\phi \propto (b_1 - 1)$ from (2.24), we therefore have

$$\sigma(f_{\text{NL}}^\Delta)_{\text{single}}^{\text{fixed-}z} \propto \frac{1}{b_1 - 1}. \quad (2.31)$$

As a result, we can expect a large improvement in a single-tracer analysis from choosing a sample with $b_1 \gtrsim 2$. Figure 2.7 shows how the value of $b_1(z=0)$ influences the overall constraining power of our billion-object survey in both single- and multi-tracer analyses. We see that $\sigma(f_{\text{NL}})$ increases sharply when the bias approaches one. It does not diverge because of the redshift dependence of $b_1(z)$ which ensures that some of the redshift bins have $b_\phi \neq 0$. On the other hand, sample-variance cancellation in the multi-tracer configuration results in a much weaker dependence of $\sigma(f_{\text{NL}}^\Delta)$ on the linear bias and, in fact, benefits from having one of the tracers with $b_1^B(z=0) < 1$. At the same time, we however see that taking larger values of $b_1(z)$ in a single-tracer analysis closes the gap to the double-tracer case for $b_1(z=0) \gtrsim 3$.

Technically speaking, we cannot choose the linear bias $b_1(z)$ and the number densi-

ties $\bar{n}_g(z)$ independently, in particular since more highly biased objects are more rare. In practice, traditional spectroscopic surveys are typically limited by the number of spectra that can be measured and not the number of objects in the universe. However, other types of surveying techniques, such as line intensity mapping (see e.g. [115–120] in the context of PNG), may get closer to these fundamental limits. We refer to [40] for a discussion of the sensitivities of surveys in the regime where the galaxy power spectrum analysis is limited by the number of available objects.

Dependence on the Volume and Redshift

Given a single redshift bin with volume V and a fixed choice of targets, it is easy to see the impact of this choice on $k_{\min}$ and $k_{\max}$. However, when we design a survey, the choice of redshift range, volume and targets is not fixed, but is part of the survey design. The redshift range and volume affect $k_{\min}$ and $k_{\max}$, but they also change the biases of the targets and the number densities of objects in each redshift bin. For the purpose of designing a survey, we therefore want to understand the optimal choices for the sky fraction f_{sky} and the redshift range as given by $z_{\max}$.

It is important to compare survey strategies assuming “constant effort”. In practice, it is the time for acquiring spectra that limits the sensitivity of a spectroscopic survey. This is why we vary f_{sky} for a given redshift range while holding the total number of objects fixed. As a result, there is a trade-off between increasing the volume and increasing the shot noise, which needs to be optimized. When it comes to varying the redshift range, the time required to observe each object changes with redshift which means that the total number of objects in the sample decreases with larger $z_{\max}$. To hold the observational effort fixed, we use a simplified model where it takes twice the amount of time to obtain spectra for galaxies at $z > 3$ than for galaxies at $z < 3$ (cf. [54]). We further assume that it is desirable to maintain a constant level of shot noise over the survey volume, which is equivalent to $\bar{n}P_g = \text{const}$ given the assumption (2.9) for the redshift evolution of the linear bias, or that we obtain spectra evenly across the sky and along the

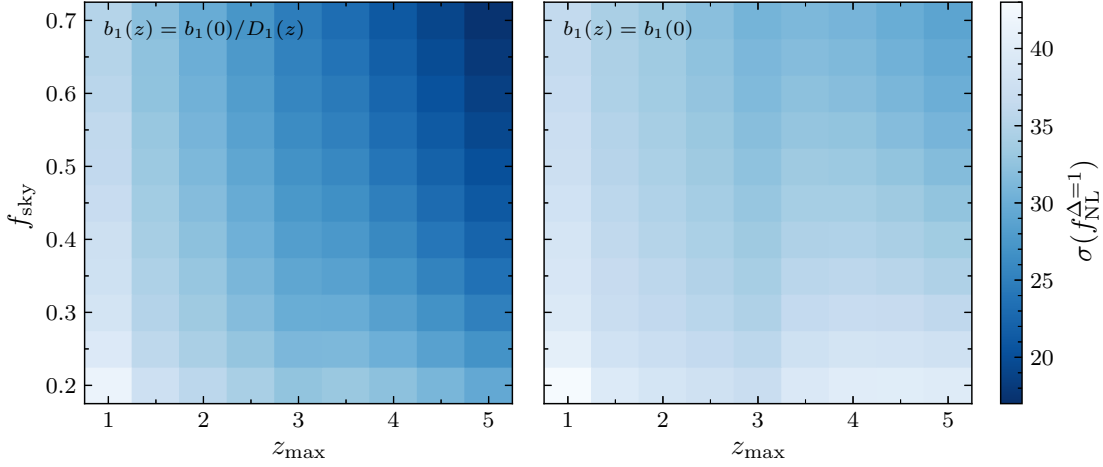


Figure 2.8. Dependence of $\sigma(f_{\text{NL}}^{\Delta=1})$ on the sky fraction f_{sky} and the maximum redshift z_{max} at fixed observational effort for the double-tracer billion-object survey. The left panel shows the fiducial case, $b_1(z) \propto D(z)^{-1}$ [cf. (2.9)] and the right panel displays the constant bias case, $b_1(z) = \text{const.}$ While we only show the results for $\Delta = 1$, the displayed trends are similar for all scaling exponents Δ .

radial direction.

Higher-redshift objects are generally expected to yield larger values of $b_1(z)$ [see the discussion around (2.9)]. Since $b_\phi(z) \propto (b_1(z) - 1)$ from the universality relation (2.24), the signal can be enhanced by a large amount by extending the survey to redshifts where $b_1(z) > 2$ instead of $b_1(z) \approx 1$. Assuming the bias evolution model of (2.9), we see in Fig. 2.8 that we maximize our sensitivity to f_{NL}^Δ even at constant effort by maximizing z_{max} . While we only show the results for $\Delta = 1$, the qualitative features in the figure are independent of the scaling exponent Δ . This implies that increasing the redshift range at fixed observational effort increases the constraining power for the entire range of Δ .

Increasing the redshift range also increases the number of modes. If we however hold $b_1(z) = b_1$ fixed and increase z_{max} , we see in the right panel of Fig. 2.8 that there is a more complex relationship between z_{max} and the sensitivity to f_{NL}^Δ . For instance, keeping $z_{\text{max}} = 3$ and increasing f_{sky} also increases the number of modes without the added observing time per object needed at high redshifts. Having said that, the benefit of a large sky fraction is much weaker than the impact of larger b_1 at high z which can be seen by the lower overall sensitivity to f_{NL}^Δ in the

right panel.

By comparing the two panels in Fig. 2.8, we can see that the optimal configuration is still at the largest maximum redshift and sky fraction, $z_{\text{max}} = 5$ and $f_{\text{sky}} = 0.7$, for either biasing model. For smaller f_{sky} , the large advantage offered by larger z_{max} however disappears when going from evolving to fixed bias due to the lower overall number density of galaxies at higher redshifts. In this case, more emphasis is placed on high-redshift targets without the corresponding boost in the signal. (The drop-off in $\sigma(f_{\text{NL}})$ between $z_{\text{max}} = 3.0$ and 3.5 is the result of our simplified choice of dividing the observational effort at $z = 3$.)

To summarize, at fixed observational effort, increasing the redshift range via a larger z_{max} has the largest influence on the sensitivity of a survey. The reason is that z_{max} increases both the size of the signal and the number of modes. The former is particularly important since it offsets the increased noise associated with a large-volume survey. We however note that this conclusion does depend on the redshift dependence of the bias and the observing time needed for acquiring spectra of these high-redshift objects.

Photometric versus Spectroscopic Surveys

The benefit of a three-dimensional survey can simply be estimated by counting modes for many cosmological parameters, including f_{NL} from the bispectrum. The number of modes in a three-dimensional survey scales as k_{max}^3 , while it scales as k_{max}^2 for a two-dimensional survey. The benefits of a three-dimensional survey are however less clear for scale-dependent bias since the information for this signal manifests itself at low k , in particular for $\Delta \lesssim 0.5$. In reality, the trade-off is of course not between a two- and a three-dimensional survey, but between surveys which measure redshifts spectroscopically and photometrically. While photometric redshifts are less precise, they may be good enough to provide three-dimensional information on the scales needed for scale-dependent bias for a significant range of Δ which we investigate now.¹²

¹²Spectroscopic surveys will have additional benefits for controlling systematics, such as projecting out large-scale effects, but typically at the cost of larger shot noise. Quantifying the importance of spectra in this context is a delicate issue which is beyond the scope of this work.

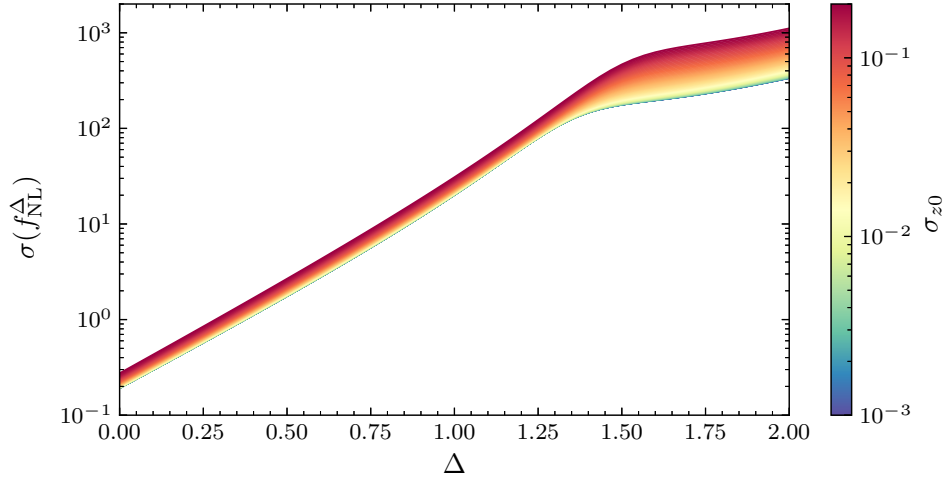


Figure 2.9. Comparison of $\sigma(f_{\text{NL}}^\Delta)$ for a range of photometric redshift errors $\sigma_{z0} \in [0.001, 0.2]$ for the billion-object survey. We observe that photometric redshift errors only mildly degrade the constraining power of a survey for $\Delta \lesssim 1.3$. Photometric surveys therefore seem to be able to access most of the information about the scale-dependent bias with the additional benefit of generally lower noise levels in the galaxy power spectrum.

The impact of redshift errors on the constraining power of the billion-object survey is shown in Fig. 2.9. We see, somewhat surprisingly, that spectroscopic surveys do not provide a significant improvement to the power spectrum measurement of f_{NL}^Δ for $\Delta \lesssim 1 - 1.3$. (The specific value of this transition in Δ depends on the employed biasing model.) In practice, photometric redshifts therefore appear to be good enough to access a lot of the available information about the scale-dependent bias. This conclusion is supported by the similarity of the forecasts for SPHEREx and LSST. While the two surveys pursue very different strategies for obtaining redshifts, the resulting forecasts are quite similar. At first sight, this similarity is still surprising because it does not reflect the impact of the difference in number density shown in Fig. 2.4. The reason for this, however, is just that multi-tracer analyses are not implemented in conventional LSST forecasts. This naturally raises the question of whether a multi-tracer analysis would lead to an advantage of LSST over some spectroscopic surveys given its very high number density (cf. Fig. 2.3). We will explore this in detail in Section 2.6.

Importantly, we have not discussed the impact of systematic effects [121] on photometric and spectroscopic surveys. Given that the signals are the largest on large angular scales (especially

for $\Delta \lesssim 0.5$), a variety of atmospheric, detector and astrophysical effects can contaminate our signal. Having said that, we expect that many of the conclusions drawn from these forecasts will be robust over a large parameter range since the scales on which the signals dominate vary with the scaling exponent Δ .

2.5 Constraints and Data Analysis for General Scaling Exponents

In the previous section, we focused on forecasting current and future constraints on f_{NL}^Δ , and on how to optimize a survey to measure PNG from the galaxy power spectrum for $\Delta \in [0, 2]$. In this section, we will apply this knowledge to constrain f_{NL}^Δ over this same range of scaling exponents Δ using the BOSS DR12 galaxy power spectra. We will first derive and propose an effective and convenient search strategy. We will then perform this measurement and compare it to a direct analysis.

2.5.1 Search Strategy

The scaling exponent Δ is fixed in some inflationary models, while it is a free parameter in others. For the standard local and equilateral shapes, the value of Δ are fixed by their templates to $\Delta = 0$ and $\Delta = 2$, respectively. The physics of these cases is clear since they arise from multi- and single-field dynamics. In quasi-single-field inflation, we have $\Delta = 3/2 - \sqrt{9/4 - m^2/H^2}$ which implies that it is natural to scan over values of $\Delta \in [0, 3/2]$ in an analysis of those models. The natural question we face in a practical analysis of current data therefore is whether it is prudent to vary f_{NL}^Δ and $\Delta \in [0, 2]$ as free parameters, or whether we should hold the scaling exponent fixed to some value as suggested by a given model, $\Delta = \Delta'$, and only vary $f_{\text{NL}}^{\Delta'}$.

The strategy we will follow is to hold Δ fixed to a few discrete values, e.g. $\Delta = 0, 2$ or $\Delta = 0.0, 0.1, \dots, 1.9, 2.0$, and infer the associated constraints on f_{NL}^Δ at each respective Δ . The reason why we prefer this strategy over scanning over both parameters simultaneously is that small changes to Δ have no impact on the signal to noise and, therefore, represent a degenerate

direction in the analysis. This becomes evident in the Fisher matrix formalism, where we have for a fiducial value of $f_{\text{NL}}^\Delta = 0$:

$$\left. \frac{\partial}{\partial \Delta} P_{gg}(k) \right|_{f_{\text{NL}}^\Delta=0} = 6 f_{\text{NL}}^\Delta b_\phi(k) k^{-2} \mathcal{T}(k)^{-1} \log(kR_\star) (kR_\star)^\Delta b(k) P_{\text{lin}}(k) \Big|_{f_{\text{NL}}^\Delta=0} = 0. \quad (2.32)$$

We therefore have no reason to expect marginalizing over Δ would lead to more accurate results until our data are comfortably excluding $f_{\text{NL}}^\Delta = 0$.

We can quantify to what degree the data can distinguish two values of the scaling exponent, Δ_1 and Δ_2 , through the “cosine” between the signals,

$$\cos(f_{\text{NL}}^{\Delta_1}, f_{\text{NL}}^{\Delta_2}) = \frac{F_{f_{\text{NL}}^{\Delta_1} f_{\text{NL}}^{\Delta_2}}}{\sqrt{F_{f_{\text{NL}}^{\Delta_1} f_{\text{NL}}^{\Delta_1}} F_{f_{\text{NL}}^{\Delta_2} f_{\text{NL}}^{\Delta_2}}}} \equiv \frac{F_{12}}{\sqrt{F_{11} F_{22}}}, \quad (2.33)$$

where we defined $F_{ij} \equiv F_{f_{\text{NL}}^{\Delta_i} f_{\text{NL}}^{\Delta_j}}$. This is precisely the same definition of the cosine used to define the PNG shapes in a bispectrum analysis [122]. We can also generalize this to the Fisher matrix after marginalizing over the bias parameters by inverting the Fisher matrix, truncating to $f_{\text{NL}}^{\Delta_1}$ and $f_{\text{NL}}^{\Delta_2}$, and inverting the truncated matrix to get the effective Fisher matrix for these two parameters. We therefore take F_{ij} to be the marginalized, two-dimensional Fisher matrix for the parameters $f_{\text{NL}}^{\Delta_1}$ and $f_{\text{NL}}^{\Delta_2}$ from now on.

The correlation (2.33) between the measurements of $f_{\text{NL}}^{\Delta_1}$ and $f_{\text{NL}}^{\Delta_2}$ is shown in Fig. 2.10 using the experimental specifications for the BOSS survey of Table A.1 which have yielded forecasted results consistent with performed data analyses in the past (see e.g. [18]).¹³ The cosine computed for other surveys is similar (see Appendix A.2). The key take-away from this correlation is that we can effectively constrain almost the entire range $\Delta \in [0, 2]$ using

¹³Note that for the calculation of the correlation matrix for BOSS, we assumed fiducial values of b_{s2} and $b_{\text{III}[2]}$ based on the Lagrangian local-in-matter-density biasing model [25] and b_{s2} based on halo simulation fit in [81]. This choice (rather than setting them to 0 as in most other cases in this paper) leads to minimal differences in the correlation coefficient. In addition, we impose the same Gaussian priors on the loop biases as quoted in [123] which barely affect the forecasted constraints for small Δ and up to a factor of four at large Δ , as expected. Finally, we employ a slightly larger minimum wavenumber $k_{\text{min}} = 0.01 h \text{Mpc}^{-1}$. All these choices are guided by the data analysis performed in §2.5.2.

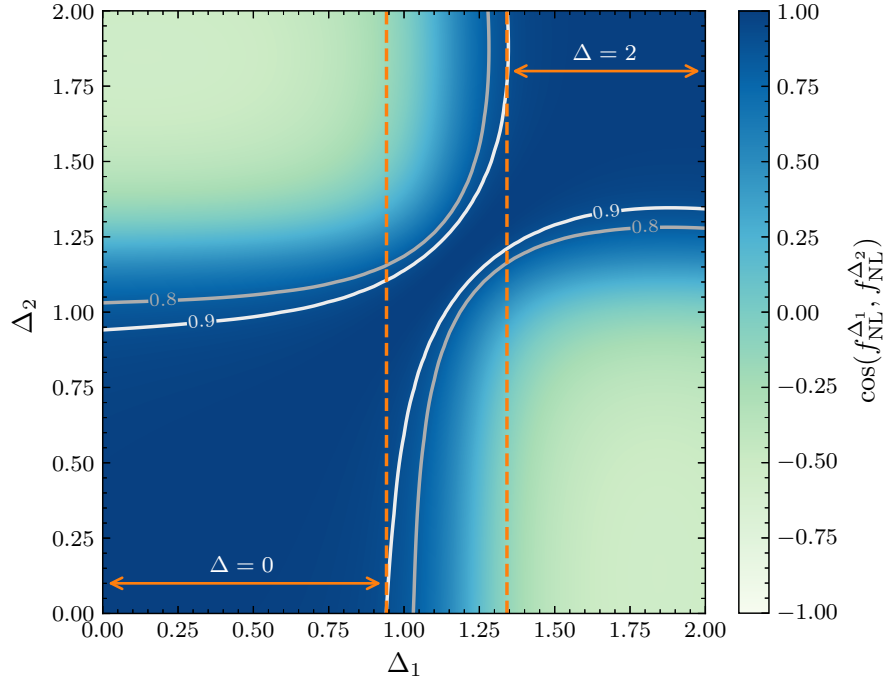


Figure 2.10. Correlation matrix for measurements of galaxy power spectra with different values of the non-Gaussian exponent Δ as defined in (2.33). This was computed for BOSS with fiducial values of $f_{\text{NL}}^{\Delta_i} = 0$, fixed Λ CDM parameters and marginalized over all biases. The orange arrows indicate the coverage of the local ($\Delta = 0$) and equilateral ($\Delta = 2$) templates with a correlation coefficient of larger than 0.9.

measurements only at $\Delta = 0$ and 2. Specifically, using the Cauchy-Schwartz inequality and the Cramér-Rao (CR) bound, we have

$$(F_{22})^{-1} \leq \frac{F_{11}}{F_{12}^2} \leq \frac{\sigma^2(f_{\text{NL}}^{\Delta_1})}{\cos^2(f_{\text{NL}}^{\Delta_1}, f_{\text{NL}}^{\Delta_2})} \frac{F_{11}}{F_{22}}, \quad (2.34)$$

where $\sigma^2(f_{\text{NL}}^{\Delta_1})$ is the variance of the measurement of $f_{\text{NL}}^{\Delta_1}$ while holding $f_{\text{NL}}^{\Delta_2}$ fixed. The best possible constraint we could place on $f_{\text{NL}}^{\Delta_2}$, $\sigma^2(f_{\text{NL}}^{\Delta_2})_{\text{CR}} = F_{22}^{-1}$, is therefore bounded by the measurement of $f_{\text{NL}}^{\Delta_1}$ via

$$\sigma^2(f_{\text{NL}}^{\Delta_2})_{\text{CR}} \leq \frac{\sigma^2(f_{\text{NL}}^{\Delta_1})}{\cos^2(f_{\text{NL}}^{\Delta_1}, f_{\text{NL}}^{\Delta_2})} \frac{F_{11}}{F_{22}}. \quad (2.35)$$

As a result, we can place an approximate constraint on $f_{\text{NL}}^{\Delta_2}$ by rescaling the measurement of $f_{\text{NL}}^{\Delta_1}$ using the elements of the marginalized Fisher matrix F_{ij} .

It is important to highlight that this strategy is only effective if our parameter inference is consistent with $f_{\text{NL}}^{\Delta} = 0$. To be more precise, all values of Δ look similar at low signal to noise for f_{NL}^{Δ} . On the other hand, the data is no longer independent of Δ if one or more of your f_{NL}^{Δ} measurements at specific Δ show significant evidence for $f_{\text{NL}}^{\Delta} \neq 0$. In that case, a measurement at a different value Δ' could produce a much larger or much smaller signal to noise. We refer to Appendix A.2 for more details on this scenario.

2.5.2 BOSS DR12 Analysis

Based on the strategy laid out above, we inferred constraints on f_{NL}^{Δ} from the BOSS DR12 dataset following the general setup of the analysis for local, orthogonal and equilateral PNG performed in [66, 124] (see also [65]) with a few different (generally more conservative) choices. In the following, we provide a brief overview of our data analysis, describe these differences and report the limits on PNG that we inferred directly and via the correlation matrix (2.33).

We analyzed the power spectrum from galaxy clustering data of the twelfth and final release of BOSS, referred to as DR12 [77]. This dataset contains the positions of about

1.2×10^6 galaxies between redshifts 0.2 and 0.75 in a cosmic volume of roughly $5.8 h^{-3} \text{Gpc}^3$ divided into four subsets: the Northern and Southern galactic caps of the two non-overlapping redshift bins $[0.2, 0.5]$ and $[0.5, 0.75]$ with effective mean redshifts $z = 0.38$ and 0.61 . We employ the galaxy power spectrum multipoles $P_{gg}^{(\ell)}(k)$, with $\ell = 0, 2, 4$, measured using a quadratic window-function-free estimator [125]. The corresponding covariance matrices were computed from MultiDark-Patchy mock catalogs [126]. This follows the analyses in [66, 124] restricted to the power spectrum multipoles. We model the nonlinear galaxy power spectrum in redshift space as implemented in CLASS-PT [127] (based on CLASS [128]) which follows the analyses of [129, 130] based on the effective field theory of large-scale structure (EFTofLSS) [131–134] (see [27, 135] for recent reviews) and galaxy bias expansion [25], including the Alcock-Paczynski effect. We additionally include the non-Gaussian contributions from the scale-dependent bias (2.23) due to $f_{\text{NL}}^{\Delta} \neq 0$ similar to the analyses in [66, 124].¹⁴

We performed a Markov chain Monte Carlo (MCMC) analysis of the data using the Metropolis-Hastings sampler of MontePython [136, 137], varying the theoretical galaxy power spectrum multipoles in each step. We fix the ΛCDM parameters to the Planck 2018 best-fit values [1] (with the sum of neutrino masses $\sum m_\nu = 0.06 \text{eV}$) instead of including Planck information on ΛCDM since we saw essentially no change in the forecasted constraints on f_{NL}^{Δ} of Section 2.4 when marginalizing over all of our galaxy bias parameters. Per redshift bin and galactic cap, we however separately vary the bias parameters $b_1^{(i)}$, $b_2^{(i)}$ and $b_{\mathcal{G}_2}^{(i)}$ as defined in [138, 139],¹⁵ and analytically marginalize over the other eight biases and EFTofLSS counterterms which appear linearly in the theoretical galaxy power spectrum [140]. We impose flat priors on the linear biases $b_1^{(i)} \in [1, 4]$ and the non-Gaussian amplitude f_{NL}^{Δ} (infinitely wide), and Gaussian priors on the other nuisance parameters as described in [123]. Following [124], we also marginalize over the non-Gaussian bias $b_\phi \rightarrow N_{b_\phi} b_\phi$ separately for each data subset and

¹⁴In this paper, we have not considered the imprints of light inflationary fields in the bispectrum, but focused on the scale-dependent bias in the power spectrum. To remain model agnostic, we also did not include the contribution of a non-zero primordial bispectrum to the galaxy power spectrum.

¹⁵The bias expansions used in our forecasts and data analysis are equivalent.

impose a wide Gaussian prior on its normalization, $N_{b_\phi} \sim \mathcal{N}(1, 5)$, which is motivated by the peak-background split model [94].¹⁶ On the other hand, the non-Gaussian parameter f_{NL}^Δ is commonly varied for all four data subsets at fixed exponent Δ . We (more) conservatively limit the range of wavenumbers to $k \in 0.01 h \text{Mpc}^{-1} \text{ to } 0.13 h \text{Mpc}^{-1}$ and $0.01 h \text{Mpc}^{-1} \text{ to } 0.16 h \text{Mpc}^{-1}$ for the low- and high-redshift bin, respectively, after initial tests. We therefore employ a smaller maximum wavenumber than in [66, 124], who tested the validity of the analysis for $\Delta = 0$ and 2 on mock catalogs, and a slightly higher (same) minimum (maximum) wavenumber than the forecasts shown in Fig. 2.3. Since $\Delta k = 0.005 h \text{Mpc}^{-1}$, this choice implies that we employ 24 and 30 k -bins for the respective redshift bins, each power spectrum multipole and both galactic caps. All chains converged with a Gelman-Rubin criterion of $R - 1 < 0.01$ (usually much smaller) for each parameter.

The results of our separate MCMC data analyses for 21 fixed values of the non-Gaussian exponent Δ are shown in Fig. 2.11 and we provided the numerical constraints for $\Delta = 0.0, 0.5, 1.0, 1.5, 2.0$ in (2.2). Since the posterior distributions for f_{NL}^Δ are slightly non-Gaussian, we display the mean values in gray pentagons, and the upper and lower 2σ constraints in filled and unfilled black diamonds, respectively. We do not find any evidence for any PNG shape with $\Delta \in [0, 2]$ and observe the same characteristic functional dependence of $2\sigma(f_{\text{NL}}^\Delta)$ on Δ as we forecasted in Fig. 2.3.¹⁷ We note that the difference in forecasted constraining power of Fig. 2.3 and the data analysis of Fig. 2.11 can be mainly attributed to the use of Gaussian priors on the higher-order galaxy bias parameters in Section 2.5. In fact, we find reasonable consistency if we include the priors on the Gaussian biases in our Fisher matrices as discussed in §2.5.1.

¹⁶It might appear surprising that we marginalize over b_ϕ with a Gaussian prior given the degeneracy with f_{NL}^Δ . Note that the mean value of the Gaussian is still set by the universality relation (2.24) with $p = 1$, but its variance is very wide. In addition, we marginalize over b_ϕ in each of the four data subsets separately, as we do with the Gaussian bias parameters. This procedure alleviates degeneracies between f_{NL}^Δ and b_1 which can be severe if we assume that the universality relation is exact, especially for $\Delta \sim 1$. This procedure is an imperfect solution to address the uncertainty in b_ϕ which we plan to address as part of [141].

¹⁷The non-Gaussian amplitude f_{NL}^Δ is defined relative to the scalar amplitude, $A_s^{1/2} \approx 5 \times 10^{-5}$ [1], such that weak PNG corresponds to $f_{\text{NL}} A_s^{1/2} \ll 1$. This means that our bounds $2\sigma(f_{\text{NL}}^\Delta) \sim 10^4$ are approaching the limits of the range of validity of our implicit expansion in linear order in f_{NL}^Δ . This concern is however alleviated even with a very conservative inclusion of bispectrum data [65, 124].

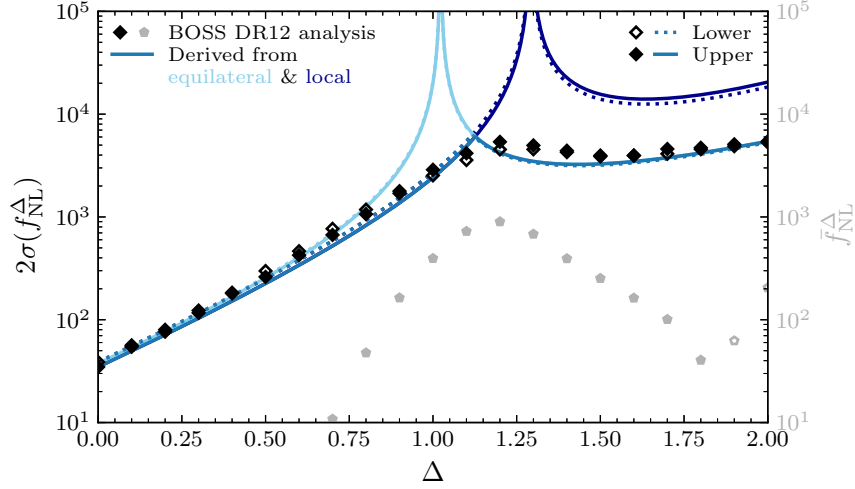


Figure 2.11. Comparison of the measured (diamonds) and derived (lines) values for twice the standard deviation of f_{NL}^Δ , $2\sigma(f_{\text{NL}}^\Delta)$, for the BOSS DR12 dataset. The light (dark) blue lines show the approximate value of $2\sigma(f_{\text{NL}}^\Delta)$ derived from the data with $\Delta = 2$ (0), i.e. the equilateral (local) PNG templates using (2.36), with the minimum of these being shown in blue. The diamonds indicate the constraints directly inferred from BOSS DR12 for fixed values of Δ , while the gray pentagons are the corresponding mean values. Filled (unfilled) symbols and solid (dotted) lines represent positive or upper (negative or lower) values. We do not find evidence for $f_{\text{NL}}^\Delta \neq 0$ for any $\Delta \in [0, 2]$, and generally good agreement between the measured and derived limits on general PNG. This demonstrates that our proposed search strategy works on data.

Compared to [65, 66, 124], our constraints are weaker, as expected due to our more conservative approach and use of less data, but consistent. This is particularly evident for the case of the equilateral shape ($\Delta = 2$) that dramatically benefits from bispectrum information in various ways which is absent in our analysis of the power spectrum alone. On the other hand, our limits on f_{NL}^{Δ} are comparable to the constraints on $f_{\text{NL}}^{\text{eq}}$ when including bispectrum information [124] for $\Delta \lesssim 0.75$. While the bispectrum will always add information, we might however imagine there is a more limited gain in this regime.

We can also compare our bounds on f_{NL}^{Δ} to those obtained in previous analyses of CMB and LSS data. As already anticipated and explained in Section 2.4, the CMB limits on local and equilateral PNG from the Planck 2018 bispectra are significantly better [16]. This is also the case for the constraints for $\Delta \in [0, 3/2]$ derived from Planck 2013 data [142], $|f_{\text{NL}}^{\Delta}| \lesssim 50$, which is dominated by the equilateral bispectrum while our bounds arise from the squeezed limit of the quasi-single-field shapes. The only previous analysis constraining f_{NL}^{Δ} from LSS data was performed in [78] on SDSS-III DR8 data with more limited modeling and range of scales. It is therefore unsurprising that our bounds are considerably more constraining.

In addition to directly inferring constraints on f_{NL}^{Δ} from the data, there exists a second path to inferring the bounds at a given value of Δ based on the correlation of different values of Δ in the marginalized galaxy power spectrum as discussed in §2.5.1. We take $\Delta = \Delta_1 = 0$ and 2 as the fiducial values for this inference since these scaling exponents correspond to the signal from local and equilateral PNG, respectively, which are usually considered in data analyses. Based on (2.35), we obtain these derived bounds as follows:

$$\left[2\sigma(f_{\text{NL}}^{\Delta})\right]_{\text{MCMC}} = \left[\frac{1}{\cos(f_{\text{NL}}^{\Delta}, f_{\text{NL}}^{\Delta=0,2})} \frac{\sigma(f_{\text{NL}}^{\Delta})}{\sigma(f_{\text{NL}}^{\Delta=0,2})}\right]_{\text{Fisher}} \left[2\sigma(f_{\text{NL}}^{\Delta=0,2})\right]_{\text{MCMC}}, \quad (2.36)$$

where the subscript ‘Fisher’ (‘MCMC’) indicates that these quantities are calculated based on the Fisher matrix (from the MCMC analysis). The correlation term $\cos(f_{\text{NL}}^{\Delta}, f_{\text{NL}}^{\Delta=0,2})$ is displayed in Fig. 2.10, while the standard deviations $\sigma(f_{\text{NL}}^{\Delta})$ are computed from the marginalized

one-dimensional Fisher matrix as in Section 2.4, but including the same Gaussian priors as in the correlation term and the data analysis. The constraints derived from the measurement at $\Delta = 0$ and 2 are displayed in Fig. 2.11 as the dark and light blue curves. We see that the constraint derived from $2\sigma(f_{\text{NL}}^{\Delta=0})$ dominates for $\Delta \lesssim 1.1$, while the constraint derived from $2\sigma(f_{\text{NL}}^{\Delta=2})$ dominates for larger Δ , which is consistent with our expectations from Fig. 2.10.

Overall, we find generally reasonable agreement between the directly inferred limits on f_{NL}^{Δ} , and those jointly derived from the local and equilateral measurements. This is especially true for those values of Δ with large correlation and negligible mean values. This implies that we can use (2.36), with the correlation term calculated in a Fisher matrix approach for a given survey, to derive $2\sigma(f_{\text{NL}}^{\Delta \neq 0,2})$ in the absence of a detection of $f_{\text{NL}}^{\text{loc}}, f_{\text{NL}}^{\text{eq}} \neq 0$. In other words, we have demonstrated that the search strategy that we proposed in §2.5.1 works on BOSS data and allows to dramatically reduce the number of analyses needed to constrain general forms of primordial non-Gaussianity and light inflationary fields as parametrized by Δ from measurements of local and equilateral non-Gaussianity. The agreement between our analysis and forecasts, together with the consistency with prior analyses, is an additional sign that searches for light fields with the galaxy power spectrum will be competitive with the CMB for near-term galaxy surveys.

2.6 Multiple Tracers and the Dependence on Astrophysics

Sample-variance cancellation plays an important part in improving the measurement of primordial non-Gaussianity from the galaxy power spectrum. Applying this technique however requires knowledge of the details of the specific galaxy samples used in the multi-tracer forecast or analysis. In this section, we will first investigate how an optimal survey design can maximize the scientific return of multiple tracers in spectroscopic surveys. Then, we will consider the untapped potential of a multi-tracer configuration of LSST and compare it to SPHEREx.

2.6.1 Optimizing Spectroscopic Multi-Tracer Analyses

We now explore the potential of multi-tracer analyses and how astrophysical details of the galaxy samples affect the constraints on f_{NL}^{Δ} . Given our biasing model (2.8) with (2.20), this comes down to the description of the non-Gaussian bias parameter $b_{\phi}(z)$, which depends on three parameters when using the universality relation (2.24): δ_c , $b_1(z)$ and p . Since the parameter δ_c is completely degenerate with f_{NL}^{Δ} , we will focus on exploring what combination of $b_1(z)$ and p leads to the biggest improvement on $\sigma(f_{\text{NL}}^{\Delta})$.

For galaxy power spectrum forecasts to recover all the information in the primordial statistics, we need to apply the multi-tracer technique and its ability for sample-variance cancellation at very high number density. For example, we know the information in the primordial bispectrum for $\Delta > 0$ is dominated by $k \sim k_{\text{max}}$ [122]. On the other hand, the non-Gaussian information in the single-tracer power-spectrum forecasts are limited by sample variance at small wavenumbers for $\Delta \lesssim 0.5$. Sample-variance cancellation eliminates this artificial dependence on the noise at low k , limiting a measurement instead by the shot noise associated with galaxy formation [143], which corresponds to modes at $k_{\text{max}} \sim k_{\text{halo}}$.

Realistic spectroscopic surveys are unfortunately far from being in the fully multi-tracer regime as we have seen in Fig. 2.4. As a result, it is harder to get reliable intuition for how to optimize the measurement of f_{NL}^{Δ} . We have previously explored how to maximize the science of a spectroscopic survey by changing the experimental configuration in terms of f_{sky} , z_{max} , etc. Here, we will assume a fixed survey geometry and investigate what additional information can be extracted by taking advantage of sample-variance cancellation.

In the fully multi-tracer regime, the inverse variance depends on the linear bias $b_1^{(i)}$ and non-Gaussian bias $b_{\phi}^{(i)}$ as [144]

$$\sigma(f_{\text{NL}}^{\Delta})^{-2} \propto \left| b_1^{(1)} b_{\phi}^{(2)} - b_1^{(2)} b_{\phi}^{(1)} \right|. \quad (2.37)$$

On the other hand, we simply maximize $b_1 b_\phi$ in the single-tracer regime. The question therefore is how we can choose the samples to minimize $\sigma(f_{\text{NL}}^\Delta)$ in realistic surveys. In Figure 2.12,¹⁸ we show the relative comparison of $\sigma(f_{\text{NL}}^\Delta)$ between a double- and single-tracer forecast of the billion-object survey while holding the Λ CDM parameters fixed.¹⁹ We show this for $\Delta = 0$ (see also e.g. [145]) since we expect this case to benefit the most from a multi-tracer analysis, but explicitly verified that the forecasts with $\Delta > 0$ show at most comparable levels of improvement.

We vary the linear biases of the two subsamples of the survey in the left column of Fig. 2.12, while fixing $p = 1$, as in all of our forecasts in Section 2.4. Note that we only vary the bias at $z = 0$, as labeled in the figure, with the biases of the actual galaxy samples then being calculated based on the bias evolution model of (2.9). The values along the diagonal $b_1^{(1)} = b_1^{(2)}$ correspond to a single-tracer forecast, in which $\sigma(f_{\text{NL}}^\Delta)$ simply decreases with increasing $b_1^{(i)}$. We observe that the multi-tracer improvement is most pronounced here when the difference between $b_1^{(1)}$ and $b_1^{(2)}$ is maximized. We additionally see that the improvement of the double-over the single-tracer approach is limited to about 30%.

So far, we have assumed a universal halo mass function which corresponds to $p = 1$ in the universality relation (2.24). Even though this has been widely adopted in the literature, this assumption is not a perfect description of halos in simulations and will likely not be generally true for actual galaxy samples, cf. e.g. [146–155].²⁰ This is why we now explore the possibility of improving the constraints on f_{NL}^Δ by selecting tracers within the parametrization

¹⁸Figure 2.12 only shows the forecasted constraints and improvements for the scaling exponent $\Delta = 0$, i.e. local PNG. When we marginalize over all biases, the scaling behavior of the scale-dependent bias however has little effect on the improvement factor and the marginalized biasing model only changes the overall constraining power as indicated by the color bar in the upper panels. In other words, the overall qualitative behavior conveyed in Fig. 2.12 remains intact for different biasing models and scaling exponents.

¹⁹When marginalizing over the biases, the difference between the forecasts for fixed and marginalized Λ CDM parameters (with a Planck prior) is at most a few percent. We therefore fix the Λ CDM parameters to their fiducial values for all forecasts in this section.

²⁰We note that, as long as $b_\phi \neq 0$, a detection of scale-dependent bias for a single redshift bin and tracer implies a detection of primordial non-Gaussianity, $f_{\text{NL}}^\Delta \neq 0$. However, reaching the full sensitivity of a survey requires combining multiple tracers and/or redshift bins which necessitates knowledge of how $b_\phi(z)$ changes across the samples. Most analyses will therefore make assumptions about b_ϕ , such as the universality relation (2.24) and the values of p .

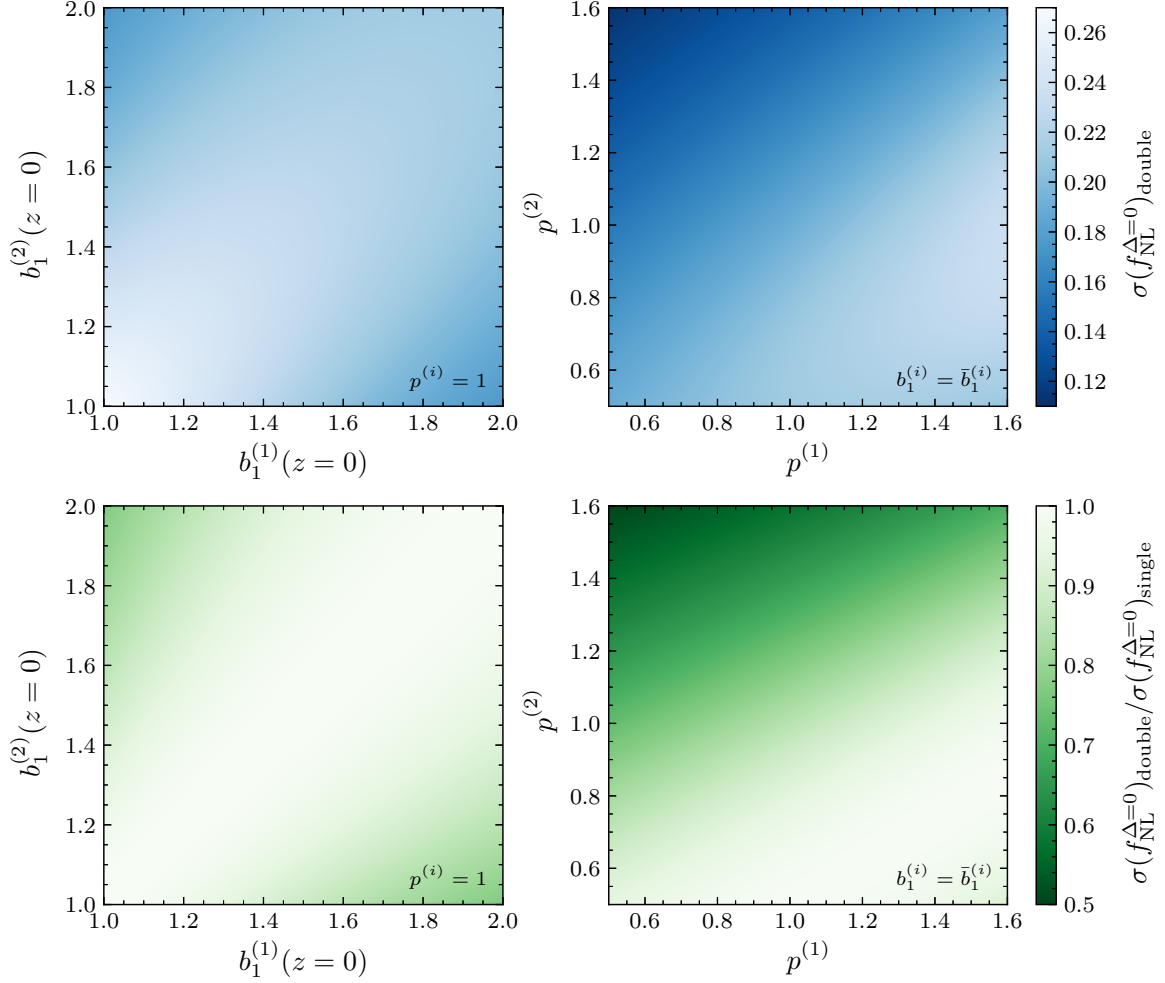


Figure 2.12. Double-tracer forecast for $\sigma(f_{\text{NL}}^{\Delta})$ with $\Delta = 0$ for the billion-object survey (*top*) and relative comparison to a single-tracer forecast, $\sigma(f_{\text{NL}}^{\Delta})_{\text{double}} / \sigma(f_{\text{NL}}^{\Delta})_{\text{single}}$ (*bottom*), with fixed Λ CDM parameters. In the left column, we vary the linear biases $b_1^{(i)}$ of the two samples at fixed $p^{(i)} = 1$, i.e. similar to our forecasts in Section 2.4. In the right column, we use the fiducial biases of the samples [cf. (2.9)], with $\bar{b}_1^{(1)}(z=0) = 2.0 > 1.2 = \bar{b}_1^{(2)}(z=0)$, and vary the parameter $p^{(i)}$ in the universality relation for $b_\phi^{(i)}$ [cf. (2.24)]. While we only show these results for $\Delta = 0$, the qualitative features of these panels remain the same for other values of the scaling exponent.

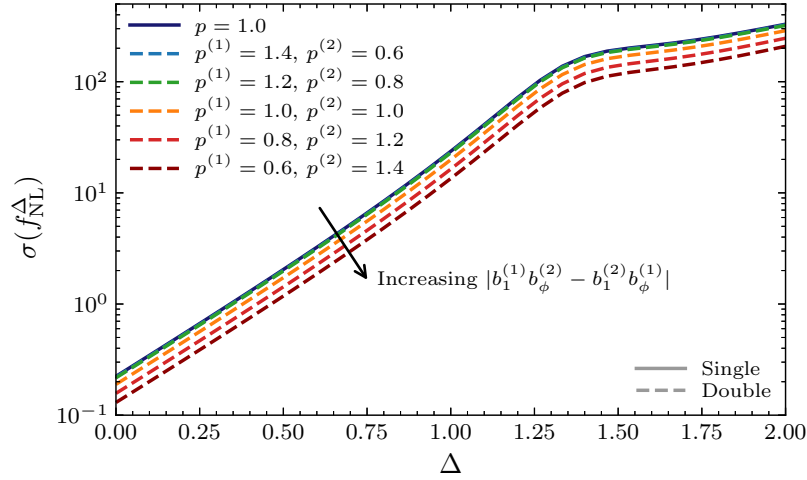


Figure 2.13. Forecasted constraints on f_{NL}^{Δ} for different values of the parameter $p^{(i)}$ in the universality relation (2.24) of b_{ϕ} for the double-tracer configuration of the billion-object survey and fixed Λ CDM parameters. We observe that $\sigma(f_{\text{NL}}^{\Delta})$ improves for all Δ when $|b_1^{(1)}b_{\phi}^{(2)} - b_1^{(2)}b_{\phi}^{(1)}|$ increases which is the direction indicated by the arrow.

of b_{ϕ} through p (assuming such knowledge is already reliably known from simulation results).

We therefore vary the parameters $p^{(i)}$ of the two galaxy samples while using their fiducial biases in the right column of Fig. 2.12. Note that the fact that the linear biases for the two tracers are different, with $b_1^{(1)}(z=0) = 2.0$ and $b_1^{(2)}(z=0) = 1.2$, is the reason for the asymmetry of these panels. We observe the largest improvement for small $p^{(1)}$ and large $p^{(2)}$ which is exactly when the combination shown in (2.37) is maximized. In this case, the double-tracer approach can improve the constraints by as much as a factor of two over a single-tracer analysis which is much larger than what we saw for fixed $p^{(i)}$ and varying $b_1^{(i)}$ in the left panels. Figure 2.13 displays the constraining power as a function of the scaling exponent Δ for a few pairs of tracers characterized by different values of $p^{(i)}$.²¹ This is further confirmation of our findings and we explicitly see that maximizing the combination of biases given in (2.37) leads to the same relative improvement in $\sigma(f_{\text{NL}}^{\Delta})$ for all Δ .

The results shown in the previous figures suggest that it may be possible to improve

²¹We also tested the possibility of using galaxy samples for which the parameter $p^{(i)}$ is redshift dependent. The results shown here are robust to these changes.

$\sigma(f_{\text{NL}}^{\Delta})$ by more than a factor of two if we had knowledge about the parameter $p^{(i)}$ of the tracers. As previously discussed, however, most forecasts for current or near-future galaxy surveys assume the scale-dependent bias takes the form predicted by a universal halo mass function. In particular, these surveys do not select targets based on the value of p (or b_{ϕ}) which implies that there is some level of uncertainty in the actual value of p for the employed galaxy samples. Since the parameter p is degenerate with f_{NL}^{Δ} , this uncertainty in p could complicate the measurement of f_{NL}^{Δ} . While complete ignorance of p (i.e. marginalizing over it as a free parameter in each redshift bin) would make constraining f_{NL}^{Δ} impossible, it would be helpful to know the required level of precision so that the universality assumption does not significantly affect the inference of f_{NL}^{Δ} .

We investigate this question by imposing different Gaussian priors on the parameter p for a fiducial value of $p = 1$. Figure 2.14 shows that an uncertainty in p mostly affects $\Delta \lesssim 0.5$ when marginalization over the biases. In addition, the degradation in the constraining power on f_{NL}^{Δ} is limited to within a factor of less than three (two) for an extremely conservative prior with width $\sigma(p) = 10$ (5). This implies that a lack of knowledge of the precise value of p for the given galaxy samples does not significantly affect the forecasts. On the other hand, a careful consideration of the values of p in target selection could potentially lead to important improvements in $\sigma(f_{\text{NL}}^{\Delta})$.

2.6.2 Multi-Tracer LSST and Comparison to SPHEREx

So far, we have seen that the multi-tracer forecasts offer somewhat limited potential for improvements in the context of both current and future, realistic spectroscopic surveys. Acquiring spectra takes more observational time than imaging alone which makes it harder for spectroscopic surveys to reach the high number densities needed for sample-variance cancellation to be most effective. In principle, photometric surveys, such as LSST, are therefore not limited in the same way, as indicated in Fig. 2.4, and have the potential to benefit significantly from a multi-tracer analysis. Following the LSST science book [107], our LSST fore-

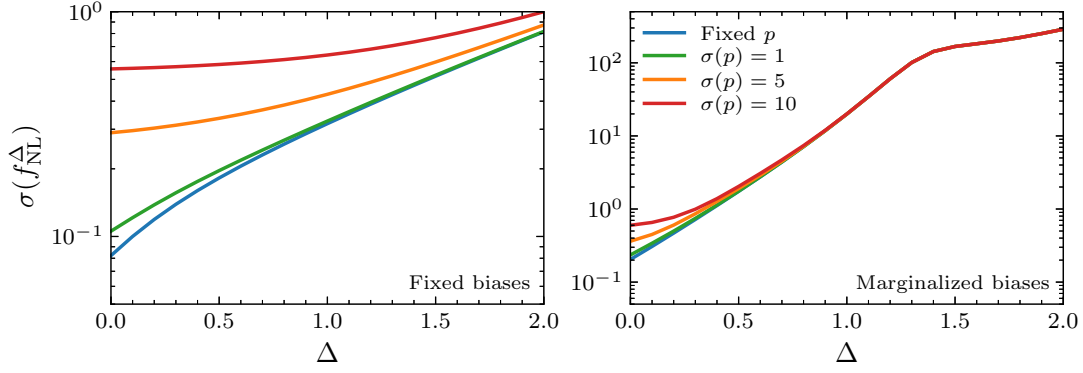


Figure 2.14. Forecast for $\sigma(f_{\text{NL}}^\Delta)$ for different Gaussian priors $\sigma(p)$ on the parameter p in the universality relation (2.24) for the billion-object survey (at fixed Λ CDM parameters and with fiducial value $\bar{f}_{\text{NL}}^\Delta = 1$ so that p can be independently varied). While marginalizing over p with wide priors significantly impacts the constraints for small Δ in the case of fixed bias parameters (*left*), the realistic case of marginalizing over the biases (*right*) shows a relatively small degradation even for very wide priors on p .

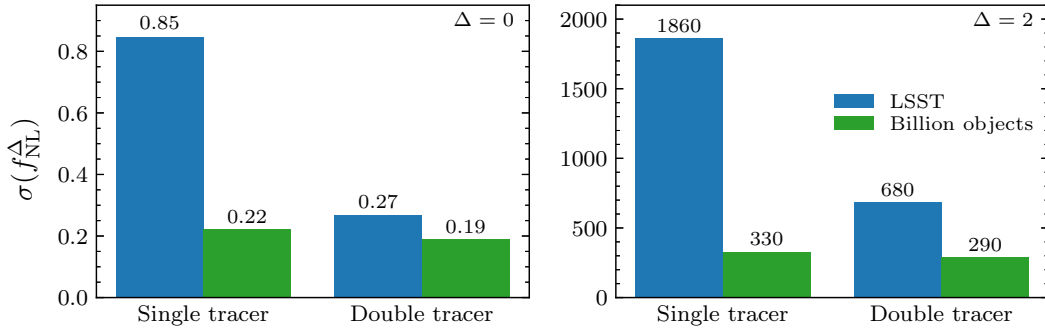


Figure 2.15. Comparison of the constraining power of single- and double-tracer versions of LSST and the billion-object survey (with fixed Λ CDM parameters) for $\Delta = 0$ (*left*) and $\Delta = 2$ (*right*). We can clearly see the significant improvement that a double-tracer analysis of LSST could provide due to its large number density for the entire range of scaling exponents Δ .

casts however use a single tracer as shown in Table A.7 and, consequently, do not capture the full potential of this survey which we will now investigate.

In order to explore the potential for an LSST multi-tracer analysis, we consider a simplified case in which we split the LSST sample of Table A.7 into two subsamples of equal number density but different linear biases $b_1^{(i)}$.²² In Figure 2.15, we compare the constraints $\sigma(f_{\text{NL}}^\Delta)$

²²For simplicity, we split the LSST sample in the same way as the sample for the billion-object survey, i.e. we take $b_1^{(1,2)}(z=0) = b_1^{\text{single}}(z=0) \pm 0.4$ with equal number densities, $\bar{n}_g^{(1)} = \bar{n}_g^{(2)}$, to retain the single-tracer LSST sample when combined. We refer to [156] for similar forecasts for an LSST sample split into “blue” and “red” galaxies.

from single- and double-tracer configurations of LSST and the billion-object survey. As for all of our spectroscopic surveys with their given number densities, there is a relatively small change in the constraining power for the billion-object survey.²³ On the other hand, LSST shows a remarkable improvement which is consistent with the factor of roughly three seen in Fig. 2.4. This increase in sensitivity however depends of course on the biases of the samples and would be diminished to less than a factor of two if we had instead taken $b_1^{(1,2)}(z=0) = b_1^{\text{single}}(z=0) \pm 0.2$, for instance. We additionally note that the much larger number density of the LSST samples appears to outweigh the larger number of (linear) modes in the billion-object survey.

In this regard, LSST has much more potential for testing inflationary physics than is typically seen in forecasts. Of course, spectroscopic information may prove essential in eliminating systematics or reliably splitting the LSST sample. The improvements that we see in Fig. 2.15 however strongly encourage a future investigation that accounts for systematic effects. In addition, we have not included the information from cosmic-shear observations, which are sensitive to primordial non-Gaussianity in their own right (see e.g. [157]) and may be useful in conjunction with the galaxy power spectra considered here. Cross-correlations with other observables, such as CMB lensing and the kinetic Sunyaev-Zel’dovich effect, would additionally provide complementary constraining power [158–160].

Given the potential of LSST as a multi-tracer survey, we should also revisit how it compares to other near-term surveys. We noted in §2.4.2 that LSST and SPHEREx result in similar forecasts despite having different strengths and weaknesses in their designs. SPHEREx is a spectro-photometric survey with a larger sky coverage, while LSST is a photometric survey with a larger number density. The comparison in §2.4.2 is additionally complicated by the fact that we treated LSST as a single-tracer survey following [107] while multi-tracer forecasts would maximize the benefit of the higher number density.

²³In this context, it is important to remember the large maximum redshift for this survey ($z_{\text{max}} = 5$) which results in a number density more or less comparable to DESI/Euclid for many redshift bins. The same number of objects in a smaller volume, as would for example be the case for a spectroscopic follow-up survey to LSST with $z_{\text{max}} = 3$, would not have the same limitation, but its effectiveness for most other science targets would be reduced [48].

We might have naively imagined that the slight advantage of SPHEREx was due to the smaller photometric redshift error of most of its samples. (Specifically, the redshift uncertainty of LSST is assumed to be $\sigma_{z0} = 0.05$, while the SPHEREx samples are binned with maximum errors of $\sigma_{z0} = \{0.003, 0.01, 0.03, 0.1, 0.2\}$.) This is however not what is found in our forecasts. In Figure 2.9, we see that the potential impact of photometric redshift errors on f_{NL}^{Δ} for $\Delta \lesssim 1.2$ is at most 50%. At the same time, we observed in Fig. 2.15 that the potential improvements in LSST from splitting the single sample into two is potentially a factor of three or more. This suggests that the number density is a larger effect than the quality of the photometric redshifts.²⁴ At the same time, there are however many other factors that differ between these surveys and we would like to identify the most important individual design factors that set the two surveys apart (at least for the purpose of constraining f_{NL}^{Δ} from the scale-dependent bias in the power spectrum).

In order to isolate the differences and control the individual factors in survey design, we compare LSST and SPHEREx by their biases and number densities in Fig. 2.16, while ignoring the photometric redshift errors for both surveys, $\sigma_{z0} = 0$. To be precise, we take SPHEREx as the fiducial survey and substitute either the linear biases or the number densities from a survey equivalent to LSST (see Appendix A.1 for details). When we keep the number densities of SPHEREx and use the LSST-like biases, i.e. two samples with equal number densities, which add to the total number density of SPHEREx, and the same biases as in our earlier LSST forecasts just over the larger redshift range of SPHEREx with the same scaling proportional to $D(z)^{-1}$, SPHEREx still outperforms in the single-tracer case, but underperforms in the double-tracer scenario. In this sense, SPHEREx has a built-in advantage from the fiducial biases used in our forecasts, while our relatively large splitting in the linear biases of the LSST-like samples leads to a considerable improvement in the regime of sample-variance cancellation. If we instead keep the biases of SPHEREx and use the number densities of LSST, i.e. we include objects

²⁴While the observed volume of SPHEREx is in principle more than twice as large (see Appendix A.1), its number densities are much smaller, especially at high redshifts, which results in an effective volume that is smaller by a factor of almost four compared to LSST.

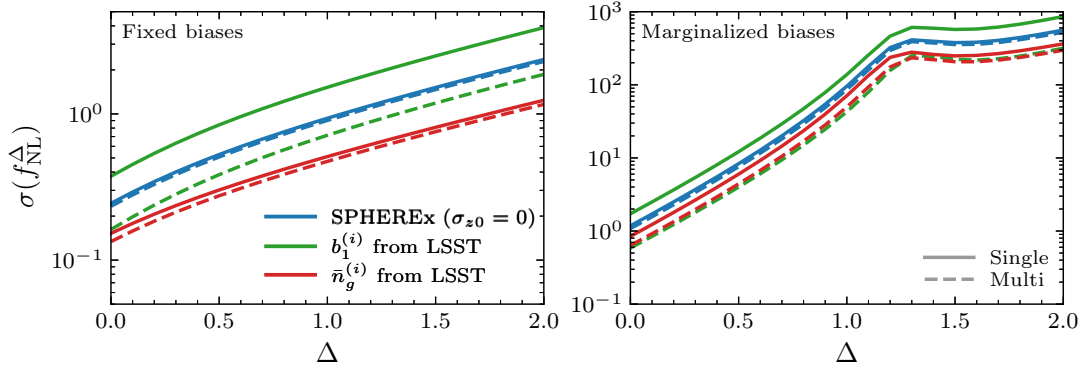


Figure 2.16. Constraining power of SPHEREx with its fiducial biases $b_1^{(i)}$ and number densities $\bar{n}_g^{(i)}$ compared to a survey with either of these observational design factors replaced by LSST-like numbers for fixed (*left*) and marginalized (*right*) Gaussian bias parameters. The solid/dashed lines show the forecasts for a single-/multi-tracer version of these surveys (five samples for SPHEREx and two samples when considering LSST-like design factors) for which we ignore the photometric redshift errors in order to directly infer the sizable impact of these design choices.

with LSST-like number densities according to (A.13) beyond its z_{max} , the latter becomes more sensitive. In fact, this advantage becomes larger when we additionally allow for a double-tracer configuration of LSST, which takes full advantage of the difference in number densities. As a result, we should understand that the slightly better sensitivity of SPHEREx indicated in Fig. 2.3 is primarily due to the single-tracer nature of the LSST forecast and the larger biases of the SPHEREx samples.

Before concluding, we have to however state that the largest challenge for both LSST and SPHEREx will of course be identifying and mitigating the impact of large-angle (and other potential) systematics. In this regard, the survey which will be most sensitive to f_{NL}^Δ cannot be easily anticipated ahead of time. Either way, both surveys have the potential to in principle place constraints on general PNG from the galaxy power spectrum alone (i.e. without the galaxy bispectrum) that exceed those inferred from the Planck bispectrum for $\Delta \lesssim 1$.

2.7 Conclusions and Outlook

Primordial non-Gaussianity from scale-dependent bias offers one of the best avenues to test the physics of inflation in the coming decade and beyond. The signal imprinted by light particles coupled to the inflaton is localized on relatively large scales where analyses are less influenced by gravitational nonlinearities, baryonic physics and modeling errors. Equally importantly, the scaling behavior of this signal encodes the mass of these light particles, offering a unique window into the spectrum and interactions of particles during the inflationary epoch.

Pragmatically, the scale-dependent bias offers a method to test for new physics within the power spectra measured in large-scale structure. One might expect that the best measurements of primordial physics should ultimately come from higher- N -point statistics or even from analyses performed directly at the level of the observed maps.²⁵ Technical obstacles have however slowed down progress and obscured the ultimate sensitivity of LSS surveys. In fact, analyses of the bispectrum only recently became possible for single-field inflation (equilateral PNG) and remain far from the demonstrated sensitivity of the CMB. The established theoretical and observational control of LSS power spectra therefore offer a reliable approach for near-term surveys to impact our understanding of inflation.

Previous work has largely focused on local PNG for which the signal to noise is dominated by only the largest angular scales accessible in a given survey. These measurements are exciting since they have the potential to exceed the sensitivity of the CMB and explore the interesting theoretical threshold of $f_{\text{NL}}^{\text{loc}} \sim O(1)$. The much richer phenomenology of inflation however raises the question if optimizing a survey for $f_{\text{NL}}^{\text{loc}}$ is the same as optimizing for understanding inflation more broadly. For example, degeneracies [167–169] and large-angle (systematic) effects [121, 170–174] that limit the local PNG measurement may not impact other changes to the statistics. Similarly, some variations in the observational strategy or target selection may not

²⁵In some cases, it has already been shown that the signals of extra fields are best understood as map-level features [161–166] that may carry more signal to noise than low-order statistics.

significantly affect the constraining power on $f_{\text{NL}}^{\text{loc}}$, but could yield a vast increase in the insights that we can gain on other inflationary observables.

In order to extend the reach of current and future surveys to these compelling inflationary targets, we studied primordial non-Gaussianity associated with light fields during inflation over a wide range of their masses. We found that surveys optimized for $f_{\text{NL}}^{\text{loc}}$ tend to perform equally well over a broader mass range beyond the massless (local) limit. Furthermore, we placed new constraints on these types of PNG by analyzing the BOSS DR12 dataset in two different and consistent ways. These results are consistent with previous analyses in the case of massless fields. In addition, they extend to heavier fields and other types of non-Gaussianity. They also show that bounds on these general non-Gaussian shapes can be simply inferred from scale-dependent bias measurements of $f_{\text{NL}}^{\text{loc}}$ and $f_{\text{NL}}^{\text{eq}}$ if they are consistent with zero. Our measured constraints from BOSS data are also consistent with our BOSS forecasts which suggests that our understanding of the constraining power of these surveys, which we described in detail, is reflected in real data. These insights are likely relevant to the analysis strategies for current or near-term observations and the design of future surveys. For instance, we demonstrated the significant improvements in sensitivity that may be available from a multi-tracer analysis of LSST data.

More broadly speaking, additional fields present during inflation are interesting in their own right, far beyond their potential impact on the galaxy power spectrum. While we focused on models leading to power-law corrections to the power spectrum, oscillatory contributions to inflationary spectra can arise from (more) massive fields [14, 92, 175] or chemical potentials [176, 177], for instance. This vast range of phenomenology therefore indicates that there are many opportunities in the analysis of galaxy surveys [156, 178] that remain under-explored and, as we also show in this paper, minor additional work could result in much broader and deeper insights into the primordial universe.

Acknowledgements

We are grateful to Alexandre Barreira, Daniel Baumann, Matteo Biagetti, Kyle Dawson, Olivier Doré, Simone Ferraro, Marilena LoVerde, Azadeh Moradinezhad Dizgah, and Charuhas Shiveshwarkar for helpful discussions. The authors were supported by the US Department of Energy under Grant DE-SC0009919. B.W. also acknowledges support from the Swedish Research Council (Contract No. 638-2013-8993). Part of this work is based on observations obtained by the Sloan Digital Sky Survey III (SDSS-III, <http://www.sdss3.org/>). Funding for SDSS-III has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation and the US Department of Energy Office of Science. We acknowledge the use of CAMB [82], CLASS [128], CLASS-PT [127], FAST-PT [179], IPython [180] and MontePython [136, 137], and the Python packages Matplotlib [181], NumPy [182] and SciPy [183].

Chapter 2, in full, is a reprint of the material as it appears in Light Fields during Inflation from BOSS and Future Galaxy Surveys, Green, Daniel; Guo, Yi; Han, Jiashu; Wallisch, Benjamin, JCAP, 2024. The dissertation author was one of the primary investigators and authors of this paper.

Chapter 3

Extending the Cosmological Collider: New Scaling Regimes and Constraints from BOSS

3.1 Synopsis

Primordial non-Gaussianity generated by additional fields during inflation offers a compelling observational target. Heavy fields imprint characteristic oscillatory signals in non-Gaussian correlation functions of the inflaton, a process sometimes referred to as cosmological-collider physics. These distinct signatures are compelling windows into ultra-high-energy physics, but are often suppressed, making standard equilateral non-Gaussianity the most promising discovery channel in many scenarios. In this paper, we show that direct couplings between the inflaton and additional fields can lead to a wide variety of novel, observationally relevant signals which open new parameter regimes that simultaneously exhibit the characteristics of light and heavy fields. We identify these primordial signatures in the late-time observables of the large-scale structure of the Universe, where they most significantly modify the scale-dependent bias of the galaxy power spectrum to include an oscillatory modulation around a non-trivial power law. We explore the full range of parameters that phenomenologically arise in these models and study the sensitivity of current and future galaxy surveys, finding that this new class of primordial non-Gaussianity is particularly accessible in near-term surveys due to its oscillatory

feature. Finally, we perform an analysis of existing data from the final release of the Baryon Oscillation Spectroscopic Survey (BOSS DR12). While we find no evidence for a signal, we demonstrate significant improvements in sensitivity over respective non-oscillatory scenarios and place the first constraints on this extended parameter space of oscillatory non-Gaussianity.

3.2 Introduction

One of the central goals of cosmic surveys is to extract the physics of the early universe from the statistics of primordial fluctuations [48]. The expansion rate, particle content and interactions throughout cosmic history dictate the statistics of the primordial density perturbations, which are in turn encoded in cosmological maps [49]. These maps therefore contain information about the standard cosmological model, the nearly Gaussian fluctuations and their evolution after the hot big bang, and extensions of this model including additional primordial signatures such as isocurvature modes, features and non-Gaussianity [17]. A key challenge in cosmology therefore is to completely decode these maps in order to reveal all the physical laws that shaped our Universe.

Inflation provides a particularly compelling opportunity to translate maps of the Universe into insights about the particle content and interactions relevant at inflationary energy scales [17]. For additional fields weakly coupled to the inflaton, the statistics of non-Gaussian correlation functions encode the masses and spins of the excitations of these fields [11–14, 51, 92, 156, 175, 177, 184–192]. This probe of fundamental physics using maps of the Universe, often referred to as *cosmological-collider physics*, is potentially sensitive to physics up to the scale of grand unified theories and beyond. Moreover, these processes introduce violations of the single-field consistency conditions [19, 34] and are therefore distinguishable from effective self-interactions of the inflaton mediated by the massive fields.

Despite the compelling theoretical motivation and broad space of signals, only a few observational analyses relevant to the cosmological collider have been performed with data from

the distribution of galaxies [193, 194] and the cosmic microwave background (CMB) [195–198]. Unfortunately, for conventional models like quasi-single-field inflation [12], the signals in the squeezed limit for massive fields are generally subdominant to the equilateral contribution to the bispectrum [193]. As a result, there has been limited effort in extracting these exciting signatures from data. However, this conclusion applies only to models with approximate de Sitter invariance. In the presence of stronger interactions between the additional fields and the inflaton, the soft limits are more complex [23, 24] and the actual potential of current and future data is unknown. These signals of the *extended cosmological collider* therefore strongly motivate that we revisit how fundamental physics is encoded in and extracted from data.

Isolating cosmological information about the history of the Universe in maps of the CMB has been a remarkably successful enterprise [1, 16, 199–201]. This is due to both the CMB being an effectively linear system, which makes isolating primordial physics particularly straightforward, and the development of powerful statistical tools, which enable the data analysis. In other words, the fact that linear transfer functions directly relate primordial fluctuations to observed modes in the CMB dramatically helps with the modeling of the signals. These insights into the CMB are now standard practice [122, 202] and have been applied to summary statistics for a wide range of primordial signals [203–207]. These measurements have confirmed that fluctuations are predominantly adiabatic, nearly Gaussian and close to scale invariant, which has already ruled out broad classes of inflationary scenarios. In principle, there remain significant opportunities beyond correlation functions and searches for more complex signals directly in the maps have opened additional pathways to extracting primordial physics [161, 208–211]. Ultimately, the two-dimensional CMB sky provides access to roughly 10^7 independent modes and cosmic variance will therefore soon limit constraints from the primary CMB.

Galaxy surveys can allow us to access vastly more information about the Universe than is available in the CMB [48]. Since these surveys cover a similar range of scales, but in three dimensions, near-term surveys are expected to contain at least one to two orders of

magnitude more modes [54]. Unlike the CMB, however, the large-scale structure (LSS) of the Universe is generally nonlinear, and primordial information can additionally be mixed with late-time gravitational evolution and astrophysics [29]. This not only moves information into higher-point correlation functions, but also complicates the modeling of the observed maps and inferred summary statistics. While considerable progress has been made through effective field theory approaches [27, 131, 135], simulation-based forward modeling [212–218], and their combination [219–223], decoding these maps has proven to be quite challenging. This may explain why a much smaller space of possible primordial signals has been explored in LSS analyses. Even looking forward, restricting to the quasi-linear modes for which the gravitational and astrophysical nonlinearities are under control often leaves the sensitivity of current and future galaxy surveys comparable to that of the CMB today.

One successful strategy for inferring primordial information from LSS maps has been to focus on protected observables [29]. Local primordial non-Gaussianity (PNG) offers one such example since it leads to modifications to the power spectrum at large scales, due to the generation of scale-dependent bias [9, 10, 17]. The presence of an additional massless field beyond the inflaton is generically expected to give rise to local PNG, making it a compelling observable. Moreover, the imprint of this signal in the form of a scale-dependent bias on large scales looks like an apparent violation of the equivalence principle during structure formation and cannot be produced locally with conventional gravity. While current bounds on the amplitude of local PNG, $f_{\text{NL}}^{\text{loc}}$, which parametrizes the size of this signal,¹ from galaxy surveys are weaker than those from the CMB, SPHEREx is expected to change this with a significant improvement [71, 224, 225].

Primordial features provide a second compelling example since they imprint oscillatory patterns in the power [204] and higher-point spectra [205, 226]: for example, sharp events

¹We emphasize that any form of primordial non-Gaussianity will generate scale-dependent bias as we will also explicitly see throughout this paper. In this paragraph, we focus on its local type since the scale-dependent bias has the potential to provide the best measurement in this case [9, 17].

during inflation or resonant interactions with heavy fields lead to a strong departure from scale invariance during inflation (or its alternatives) producing distinct signals in the correlation functions [17, 227]. Late-time astrophysical and gravitational processes however cannot give rise to these high-frequency oscillations. This allows us to measure these features without careful modeling of structure formation, making them especially robust probes of the primordial universe [18] (cf. also [228–232]). The Baryon Oscillation Spectroscopic Survey (BOSS) has already provided the strongest constraints on such oscillatory features over a wide range of frequencies [18, 233–235] and upcoming surveys will further extend these bounds to levels that CMB observations will never be able to achieve (cf. e.g. [18, 236, 237]).

One major reason to be optimistic about the potential for discovering extra fields during inflation is that they can introduce a hybrid of these two robust signals [11, 12]. Specifically, if we introduce a nonzero interaction of the inflaton with a heavy scalar field with mass $m > 3H/2$, where H is the Hubble scale during inflation, one finds an oscillatory non-Gaussian signal. In the squeezed limit of the primordial bispectrum, it takes the form²

$$\lim_{k_1 \rightarrow 0} \langle \zeta(\vec{k}_1) \zeta(\vec{k}_2) \zeta(\vec{k}_3) \rangle' = \lambda f_{\text{NL}} \left[e^{i\phi} \left(\frac{k_1}{k_2} \right)^{i\nu} + e^{-i\phi} \left(\frac{k_1}{k_2} \right)^{-i\nu} \right] \left(\frac{k_1}{k_2} \right)^{3/2} P(k_1) P(k_2), \quad (3.1)$$

where we employed the correlation function defined as $\langle \dots \rangle = (2\pi)^3 \delta_D(\sum_i \vec{k}_i) \langle \dots \rangle'$, with the Dirac delta function δ_D , ϕ is a calculable phase and $\nu = \sqrt{m^2/H^2 - 9/4}$. Since the non-Gaussian amplitude f_{NL} is set in the equilateral limit, we introduced λ to capture the relative amplitude in the squeezed configuration. Unfortunately, λ is exponentially suppressed in simple models with heavy fields, $\lambda \propto e^{-\pi\nu}$ [92]. Nevertheless, it is important that the signal in (3.1) exhibits logarithmic oscillations in k_1/k_2 with a frequency ν , which is set by the mass m of the field. These oscillations arise from the interference term between the wavefunction of the heavy field, which oscillates at a frequency set by its mass, and the massless mode of the inflaton [14].

²In a standard abuse of notation, $\lim_{k_1 \rightarrow 0}$ indicates that we take the squeezed limit $k_1 \ll k_2, k_3$ while keeping the leading term. The (implicit) momentum conservation on the left-hand side of the equation then also forces $\vec{k}_2 = -\vec{k}_3$ so that k_1 and k_2 on the right-hand side are the wavenumbers of the short and long mode, respectively.

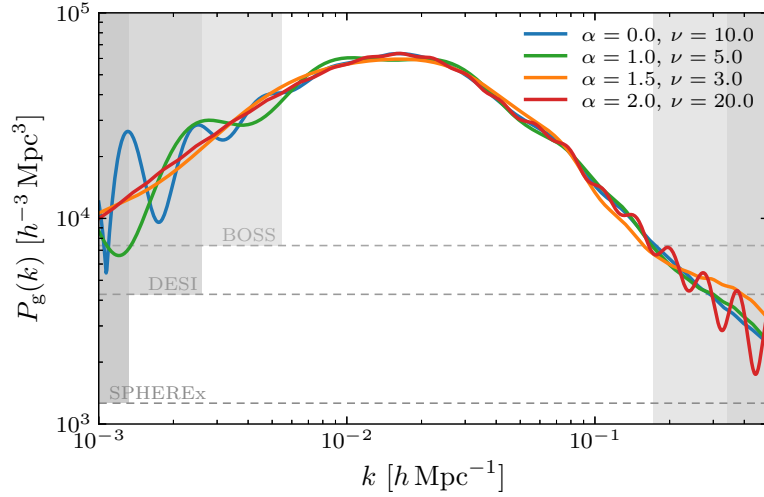


Figure 3.1. Illustration of the effect of oscillatory primordial non-Gaussianity on the linear galaxy power spectrum through the scale-dependent bias for different scaling dimensions $\Delta = \alpha + i\nu$. In addition to the standard cosmological collider with $\alpha = 3/2$ as in (3.1), we also show examples for the wider range of values considered in this work. The employed non-Gaussian amplitudes are $f_{\text{NL}}^{\alpha,\nu} = 3, 500, 2000$ and 2000 for the shown cases (with absorbed suppression factors; see especially Section 3.4 for its definition and details). We observe that the oscillatory imprint with logarithmic frequency ν moves from large to small scales as we increase the scaling exponent α , which also sets the envelope of the oscillations. The horizontal dashed lines indicate the effective shot noise level for BOSS, DESI and SPHEREx after scaling them to the displayed redshift $z = 0$ and linear bias $b_1 = 1.6$. The gray-shaded regions on large scales indicate the wavenumbers below the minimum wavenumber of these surveys, $k < k_{\text{min}}$, computed based on their entire spherical volume. The gray-shaded regions for large wavenumbers indicate the regimes where the scales exceed the nonlinear scale, $k > k_{\text{NL}}$, for the maximum redshift of BOSS and DESI, respectively.

The squeezed bispectrum in (3.1) is responsible for the robustness of the cosmological-collider signal: it induces a scale-dependent bias and imprints characteristic oscillations, as illustrated in Fig. 3.1. It therefore benefits from both forms of protection. These oscillations in addition to the smooth power-law bias are a particularly compelling theoretical and observational target for two reasons. First, the oscillation frequency ν directly encodes the mass of the additional fields which could provide access to the particle spectrum during inflation. Second, this type of oscillations can generally neither arise without new degrees of freedom in the very early universe nor be produced by known astrophysical processes at late times. As a result, it is typically easier to robustly isolate this oscillatory signal in cosmological data.

Unfortunately, the amplitude of this compelling signal is doubly suppressed in most models. First, the squeezed limit is Boltzmann suppressed $\sim e^{-\pi m/H}$ due to the high energy of the additional particles compared to the de Sitter temperature. Second, the signal scales as $(k_1/k_2)^{3/2}$. This means that it effectively vanishes as $k_1 \rightarrow 0$ which is a particularly significant observation since resolving an oscillation in the ratio of wavenumbers k_1/k_2 requires taking this squeezed limit. Moreover, the suppression by at least $(k_1/k_2)^{3/2}$ in the squeezed limit is unavoidable for massive fields that effectively see a de Sitter background, even in strongly interacting theories: the effective scaling dimension $\Delta = \alpha + i\nu$, with $\alpha \equiv \text{Re}\Delta = 3/2$ and $\nu \equiv \text{Im}\Delta$, in (3.1) can receive corrections in interacting theories, but the optical theorem (or, in other words, the positivity of probabilities) always requires $\alpha \geq 3/2$ [23, 187, 238–244]. The resulting suppression renders the oscillations encoded in (3.1) difficult to observe in practice, unless an equilateral non-Gaussian signal has already been detected with high significance.

On the other hand, light fields with mass $m \leq 3H/2$ have a real-valued effective scaling dimension $0 \leq \Delta = \alpha \leq 3/2$, which alleviates both suppression factors. While this comes at the expense of generating non-oscillatory signals, these lighter fields have therefore been considered the more promising target for current and future LSS observations [40, 193]. At the same time, however, inflation does not occur in a pure de Sitter space and allows for models that escape the discussed squeezed-limit suppression. In fact, due to the time evolution of the inflaton, the de Sitter isometries are broken and the optical theorem no longer applies. In particular, explicit examples with Lorentz-violating or time-dependent mixing between fields have been shown to allow for complex scaling dimensions with $\text{Re}\Delta = \alpha < 3/2$ and $\text{Im}\Delta = \nu > 0$ [23, 24]. In such cases, high-frequency oscillations can appear in the squeezed limit without the mentioned exponential or power-law suppressions of the standard cosmological-collider scenario. In consequence, such models would be easier to search for in cosmological data, opening a new and observationally accessible window into the particle content and interactions of the inflationary epoch.

In this paper, we explore the theoretical and observational implications of this broader class of models characterized by general complex-valued scaling dimension Δ . We first analyze the two-field model of [24] in Section 3.3, demonstrating that large frequencies ν are achievable for $0 \leq \text{Re}\Delta = \alpha \leq 3/2$. We in particular show that while the amplitude of the oscillatory signal in the galaxy power spectrum remains suppressed for vacuum fluctuations, this can be partially compensated by the enhanced observational sensitivity to these oscillations. We then deduce the observational signatures imprinted in galaxy power spectra, as exemplified in Fig. 3.1 for a range of scaling exponents α and frequencies ν , and assess the detectability of such signals through the scale-dependent bias in current and upcoming galaxy surveys in Section 3.4. Using Fisher-matrix forecasts, we demonstrate and quantify the increased sensitivity of the galaxy power spectrum to these high-frequency oscillations in the squeezed limit. In Section 3.5, we subsequently place the first constraints on these bispectra for general complex values of Δ from the BOSS DR12 dataset. Finally, Section 3.6 contains our conclusions, and a set of appendices provide additional details on our forecasts (Appendix B.1) and our data analysis (Appendix B.2).

Throughout the paper, we will use the following notation: A comoving vector $\vec{k}$ is defined to have length $k \equiv |\vec{k}|$. We denote the n -point correlation function of operators $\mathcal{O}_i$ without momentum-conserving Dirac delta function δ_D as

$$\langle \mathcal{O}_1(\vec{k}_1) \dots \mathcal{O}_n(\vec{k}_n) \rangle = (2\pi)^3 \delta_D \left(\sum_{i=1}^n \vec{k}_i \right) \langle \mathcal{O}_1(\vec{k}_1) \dots \mathcal{O}_n(\vec{k}_n) \rangle'. \quad (3.2)$$

The inflationary adiabatic mode is ζ , the additional scalar fields are σ_i and the correlation function in a classical σ background is written as $\langle \zeta(\vec{k}_1) \dots \zeta(\vec{k}_n) \rangle_\sigma$. This should not be confused with the squeezed limit of a correlator $\langle \mathcal{O}_1(\vec{k}_1) \mathcal{O}_2(\vec{k}_2) \rangle_{k_1 \ll k_2} \equiv \lim_{k_1 \ll k_2} \langle \mathcal{O}_1(\vec{k}_1) \mathcal{O}_2(\vec{k}_2) \rangle = \lim_{k_1 \rightarrow 0} \langle \mathcal{O}_1(\vec{k}_1) \mathcal{O}_2(\vec{k}_2) \rangle$ (see also footnote 2). The scaling behavior of the additional fields σ_i coupled to the inflaton is parameterized by the complex dimension $\Delta = \alpha + i\nu$, with real-valued exponent α and frequency ν , so that the power spectrum scales as $k^{2\Delta-3}$ in three spatial dimensions. For a single principal-series scalar ($m^2 > 9H^2/4$) in de Sitter space, we therefore

have $\alpha = 3/2$ and $\nu = \sqrt{m^2/H^2 - 9/4}$. We refer to the non-Gaussian amplitude for the oscillatory ($\nu \neq 0$) and non-oscillatory ($\nu = 0$) scenarios as $f_{\text{NL}}^{\alpha,\nu}$ and $f_{\text{NL}}^{\alpha,\psi}$, respectively. Finally, we use the density perturbation δ , the smoothed density field δ_M for (halo) mass M and the galaxy overdensity δ_g .

3.3 Cosmological-Collider Models

The framework of the cosmological collider provides a natural origin for the distinct oscillatory and power-law signatures [14, 92, 156, 175, 177, 185, 186]. The simplest examples arise in quasi-single-field inflation [11, 12], in which the inflaton interacts with an additional scalar of mass m . For heavy fields, the signal is characterized by a frequency which is controlled only by this mass providing a direct connection between the inflationary particle spectrum and observable non-Gaussianity. However, the space of possible signals is large and we would consequently like to understand the viable parameter space in which to look for these primordial signals. This is why we will in this theoretical section characterize the range of frequencies (§3.3.1) and amplitudes (§3.3.2) accessible in the two-field model of [24], and evaluate the impact of the squeezed limit of the adiabatic modes (§3.3.3). While many aspects of our discussion were presented in [24], our focus on the range of viable values of the scaling exponent α , oscillatory frequency ν and PNG amplitude $f_{\text{NL}}^{\alpha,\nu}$ is both new and relevant for observations.

3.3.1 Frequency Range of Two-Field Models

The frequency of oscillatory signals in cosmological-collider scenarios is determined by the scaling behavior of the fields in the super-horizon limit [cf. the frequency ν in the exponents of (3.1)]. We therefore start by analyzing this behavior in a simple model with two mixing fields which already exhibits a broad spectrum of scaling behaviors. To be specific, we consider two

scalar fields σ_i , $i = 1, 2$, which mix according to the following quadratic Lagrangian:

$$\mathcal{L}_\sigma = -\frac{1}{2} \sum_{i=1}^2 (\partial_\mu \sigma_i \partial^\mu \sigma_i - m_i^2 \sigma_i^2) + \rho (\dot{\sigma}_1 \sigma_2 - \dot{\sigma}_2 \sigma_1), \quad (3.3)$$

with the masses m_i of the two fields, the mixing parameter ρ and dots denoting derivatives with respect to physical time t . The last term is responsible for the mixing and explicitly breaks the de Sitter isometries. This is particularly significant because the allowed space of scaling behaviors (values of Δ) is highly constrained by symmetries and unitarity in a de Sitter background. Once de Sitter invariance is broken, much wider ranges of Δ are (at least in principle) allowed. As we will see, the relative sizes of the masses m_i , the mixing ρ and the Hubble scale H will determine whether the scaling dimensions are real or complex and what the size of the imaginary part is, which sets the characteristic frequencies that we are interested in.

We start by computing the scaling dimension of the model (3.3) as a function of m_i , ρ and H . The equations of motion are given by

$$\begin{aligned} \ddot{\sigma}_1 + 3H\dot{\sigma}_1 + \left(\frac{k^2}{a^2} + m_1^2\right)\sigma_1 + \rho(2\dot{\sigma}_2 + 3H\sigma_2) &= 0, \\ \ddot{\sigma}_2 + 3H\dot{\sigma}_2 + \left(\frac{k^2}{a^2} + m_2^2\right)\sigma_2 - \rho(2\dot{\sigma}_1 + 3H\sigma_1) &= 0. \end{aligned} \quad (3.4)$$

Since the Fourier modes are described by the time evolution in the super-horizon limit, $k \ll aH$, we make the ansatz $\vec{\sigma} = (\sigma_1, \sigma_2) = a(t)^{-\Delta} \vec{v}$. We consequently have to solve

$$\begin{pmatrix} \Delta(\Delta-3)H^2 + m_1^2 & H\rho(-2\Delta+3) \\ H\rho(2\Delta-3) & \Delta(\Delta-3)H^2 + m_2^2 \end{pmatrix} \vec{v} = 0. \quad (3.5)$$

The four solutions to this equation are given by

$$\Delta_\pm = \frac{3}{2} - \sqrt{\frac{9}{4} + a_\pm} \quad \text{and} \quad \bar{\Delta}_\pm = 3 - \Delta_\pm, \quad (3.6)$$

where we introduced

$$a_{\pm} \equiv -\frac{1}{2} \frac{m_1^2 + m_2^2 + 4\rho^2}{H^2} \pm 2\sqrt{\frac{1}{2} \frac{\rho^2}{H^2} \left(\frac{m_1^2 + m_2^2 + 2\rho^2}{H^2} - \frac{9}{2} \right) + \frac{1}{16} \frac{(m_1^2 - m_2^2)^2}{H^4}}. \quad (3.7)$$

We therefore see that $\Delta = \alpha + i\nu$ is in general complex. Now, we are interested in the ranges of α and ν that are possible within this model, in particular whether small α and large ν may occur concurrently.

To gain further insight by simplifying the expressions of these four solutions, it is helpful to define

$$x \equiv \frac{1}{2} \frac{m_1^2 + m_2^2}{H^2}, \quad y \equiv \frac{1}{16} \frac{(m_1^2 - m_2^2)^2}{H^2 \rho^2}, \quad z \equiv \frac{\rho^2}{H^2}, \quad (3.8)$$

so that we have

$$a_{\pm} = -x - 2z \pm 2\sqrt{z} \sqrt{x + y + z - \frac{9}{4}}. \quad (3.9)$$

Given the assumption of real fields, the parameters x , y and z are also real, and $y, z > 0$. If we demand that both squared masses are positive, $m_i^2 > 0$, we also have $x > 0$. This limits $a_{\pm}$ to be real-valued, except at small x , y and z such that $x + y + z < 9/4$. While this does not eliminate complex dimensions Δ since $a_{\pm}$ can be negative, it does not lead to novel phenomenology.

In contrast, for tachyonic masses, $m_i^2 < 0$,³ we find a much wider range of viable values of Δ . Considering the special case for which $y = 0$ and $x + z = 0$, i.e. $\rho^2 = -m_1^2 = -m_2^2 > 0$, we get $a_{\pm} = -z \pm 3i\sqrt{z}$, which implies the following four complex-valued solutions:

$$\Delta \in \left\{ \frac{3}{2} - \sqrt{\left(\frac{3}{2} \pm i\sqrt{z}\right)^2}, \frac{3}{2} + \sqrt{\left(\frac{3}{2} \pm i\sqrt{z}\right)^2} \right\} = \left\{ \pm i \frac{\rho}{H}, 3 \pm i \frac{\rho}{H} \right\}. \quad (3.10)$$

As a result, allowing $m_i^2 < 0$ leads to solutions where $\Delta = 0 \pm i\nu$. In other words, we found

³The use of negative squared masses for the fields in the Lagrangian (3.3) does not imply any non-attractor behavior in the solutions we will discuss. Tachyonic matter without mixing ($\rho = 0$) has long been known to have unique phenomenology (see e.g. [245]), due in part to the fact that the background is unstable. We however emphasize the novel signals are stable and distinct from the usual differences between attractor and non-attractor phenomenology.

at least one region of the parameter space $\{m_i, \rho\}$ of the two-field model (3.3) in which the frequency ν may be large while the real-valued scaling exponent α vanishes. Small deviations from our special case will generally preserve this behavior with a combination of potentially large frequency ν and scaling exponent $0 \leq \alpha \ll 3/2$. We reiterate that this does not violate the optical theorem since the mixing ρ between the two fields breaks the de Sitter isometries. This is therefore a phenomenologically compelling region of viable parameter space that has so far been under-explored.

3.3.2 The Amplitude of Primordial Fluctuations

To establish whether this phenomenologically compelling region of parameter space is also observationally interesting, we need to compute the expected amplitude of the oscillatory signal, which depends on how efficiently the massive modes are excited during inflation. We therefore need to calculate the normalization of the fluctuations produced during inflation, focusing on the regime where $\alpha \approx 0$ and $\nu \gg 1$. While the scaling exponent α and the frequency ν are determined by the long-wavelength behavior, as we saw above, calculating the amplitude requires matching to the sub-horizon limit. This means that the full dynamics of the coupled system must be solved. In practice, solving for the mode functions analytically is difficult in this case, but a numerical solution is fortunately sufficient to determine the amplitude. In the following, we will set up the differential equations to be solved and discuss how we set their initial conditions using the WKB approximation, numerically solve these equations and validate our procedure.

In order to calculate the amplitude of the oscillatory signal, we return to the quadratic equations of motion (3.4). To make solving these coupled differential equations numerically more straightforward, we define the dimensionless fields $\tilde{\sigma}_i \equiv k^{3/2} \sigma_i / H$, and employ the dimensionless

quantities $u \equiv k\tau$, $\tilde{m}_i \equiv m_i/H$ and $\tilde{\rho} \equiv \rho/H$ so that

$$\begin{aligned}\tilde{\sigma}_1'' - \frac{2}{u}\tilde{\sigma}_1' + \left(1 + \frac{\tilde{m}_1^2}{u^2}\right)\tilde{\sigma}_1 + \tilde{\rho}\left(-\frac{2}{u}\tilde{\sigma}_2' + \frac{3}{u^2}\tilde{\sigma}_2\right) &= 0, \\ \tilde{\sigma}_2'' - \frac{2}{u}\tilde{\sigma}_2' + \left(1 + \frac{\tilde{m}_2^2}{u^2}\right)\tilde{\sigma}_2 - \tilde{\rho}\left(-\frac{2}{u}\tilde{\sigma}_1' + \frac{3}{u^2}\tilde{\sigma}_1\right) &= 0,\end{aligned}\tag{3.11}$$

where primes denote derivatives with respect to u . Following [53], these equations can be solved numerically by starting at a time $u \ll -1$ when the mode functions are well inside the horizon. Since we took the Fourier transform, it is in principle straightforward to numerically compute the solutions in terms of u using any integrator for ordinary differential equations.

The non-trivial step in determining the amplitude of the vacuum fluctuations is imposing numerically that the modes are in the Bunch-Davies vacuum. We implement this as an initial condition for our numerical analysis as follows: in the early-time limit, $u = k\tau \rightarrow -\infty$ ($a(t) \rightarrow 0$), the WKB approximation in physical time t should correspond to a massless field with dispersion relation $\omega(t) \rightarrow ka^{-1}(t) + O(a^0)$. The WKB calculation incorporates the first sub-leading correction to (3.11), i.e. terms of $O(u^{-1})$ that are suppressed by $(k\tau)^{-1}$, but neglects the second-order terms, $O(u^{-2} = (k\tau)^{-2})$. Instead of the $\sigma_{1,2}$ basis used in (3.5), the WKB solutions for the initial mode functions are more conveniently expressed in a different basis, $\sigma_{A,B}$, which are defined with respect to the original basis by

$$\sigma_1 = \frac{\sigma_A^{(+)} - \sigma_A^{(-)}}{\sqrt{2}i}, \quad \sigma_2 = \frac{\sigma_B^{(+)} + \sigma_B^{(-)}}{\sqrt{2}}.\tag{3.12}$$

The dimensionless WKB mode functions $\tilde{\sigma}_{A,B} \equiv k^{3/2}\sigma_{A,B}/H$ are then given by

$$\tilde{\sigma}_A^{(\pm)}(u) = \mp \frac{iu}{2} e^{-i[u \pm \rho \log(-u)]}, \quad \tilde{\sigma}_B^{(\pm)}(u) = -\frac{u}{2} e^{-i[u \pm \rho \log(-u)]}.\tag{3.13}$$

A priori, one might have expected the modes to be defined in the $\sigma_{1,2}$ basis since the mixing term is subdominant when the WKB approximation holds. However, if we include only the canonical

kinetic term, we can rotate to any linear combination of $\sigma_{1,2}$ and the action will be unchanged. The benefit of using the $\tilde{\sigma}_{A,B}$ basis is that it diagonalizes the (sub-leading) mixing term $\rho \log(-u)$ in the limit where the WKB approximation is exact, $u \rightarrow -\infty$. Keeping the mixing term in the WKB approximation allows us to match the numerical result to the analytic WKB solution at smaller values of $|u|$. This, in turn, improves the accuracy of our results as we do not have to integrate the equations for as long to reach the super-horizon regime, $|u| \ll 1$.

Having determined the analytic initial conditions that hold when $|u|^2 \gg m_i^2/H^2$,⁴ we can now solve the differential equations numerically starting from the WKB regime. The analytic form of the WKB solution is sufficient to calculate the initial values and first derivatives of $\sigma_{A,B}^{(\pm)}$. Since we have to additionally determine the correct normalizations, we match to the Bunch-Davies solution by isolating the positive frequency modes in the WKB limit and canonically normalize the fields according to $q_i = a\sigma_i = ak^{-3/2}H\tilde{\sigma}_i$, $i = 1, 2$. We can then solve for $\tilde{\sigma}_i$ numerically given these initial conditions. If needed, their (equal-time) power spectra may subsequently be obtained according to

$$H^{-2}k^3P_{\sigma_1} = |\tilde{\sigma}_A^{(+)}|^2 + |\tilde{\sigma}_A^{(-)}|^2, \quad H^{-2}k^3P_{\sigma_2} = |\tilde{\sigma}_B^{(+)}|^2 + |\tilde{\sigma}_B^{(-)}|^2, \quad (3.14)$$

and we naturally have $P_{\sigma_1} = P_{\sigma_2}$ for the case of no mixing and equal masses, $\rho = 0$ and $m_1 = m_2$.

As a first step, we consider the special case of a principal-series field σ with no mixing, $m^2 > 9H^2/4$ and $\rho = 0$, since it can serve as an important point of comparison between the analytic and numerical results. In this particular case, the mode functions have an analytic solution,

$$\tilde{\sigma}(u) = \frac{\sqrt{\pi}}{2} e^{\frac{\pi\nu}{2}} (-u)^{3/2} H_{i\nu}^{(2)}(-u), \quad (3.15)$$

where $H_n^{(2)}(x)$ are Hankel functions of the second kind. In the limit $u \rightarrow 0$, we can expand the

⁴Note that in this representation, the mass term is time dependent (rather than the physical frequency) and is therefore relevant to the WKB approximation.

solution to find (see e.g. [242])

$$\tilde{\sigma}(u \rightarrow 0) \rightarrow i \frac{(-u)^{3/2}}{2\sqrt{\pi}} \left[\Gamma(-iv) \left(-\frac{i}{2}u \right)^{iv} + \Gamma(iv) \left(-\frac{i}{2}u \right)^{-iv} \right], \quad (3.16)$$

with the gamma function Γ . Since $i^{\pm iv} = e^{\mp v\pi/2}$ and the asymptotic behavior of the gamma function is $|\Gamma(iv)| \rightarrow \sqrt{2\pi}v^{-1/2}e^{-|v|\pi/2}$ as $v \rightarrow \infty$, the two terms of the mode function have an exponential hierarchy for large v ,

$$\tilde{\sigma}(u \rightarrow 0) \rightarrow i \frac{(-u)^{3/2}}{\sqrt{2v}} \left[e^{-\pi v} \left(-\frac{1}{2}u \right)^{iv} + \left(-\frac{1}{2}u \right)^{-iv} \right]. \quad (3.17)$$

As a result, the power spectrum of the original field σ in the super-horizon and large-frequency limits, $u \rightarrow 0$ and $v \gg 1$, is

$$P_{\sigma}(k, \tau) \approx \frac{H^2 \tau^3}{2v} \left[1 + 2e^{-\pi v} \cos(2v \log(k\tau)) + \dots \right]. \quad (3.18)$$

We see that the leading term is analytic in k and, consequently, is a pure contact term. The leading oscillatory (non-analytic) behavior in k is therefore suppressed by $e^{-\pi v}$. This behavior highlights an important feature of the principal-series field which is that the wavefunctions are highly oscillatory as $u \rightarrow 0$ while the power spectrum is mostly analytic. This is because the oscillations in the power spectrum arise from the interference between the two oscillatory terms of the wavefunctions. We additionally note that the exponential suppression of the massive power spectrum makes it also useful as a test of our numerical implementation in a regime with known analytic behavior. In Fig. 3.2, we show this comparison between the analytic and numerical results for both the mode function $\tilde{\sigma}(u)$ and its power spectrum $P_{\sigma}(u)$, both of which generally agree with one another within about two percent or (much) less. This therefore serves as a validation of our numerical approach that employs the Dormand-Prince method to solve the ordinary differential equations (3.11) for which we set the initial conditions at $u_0 = -5000$ and

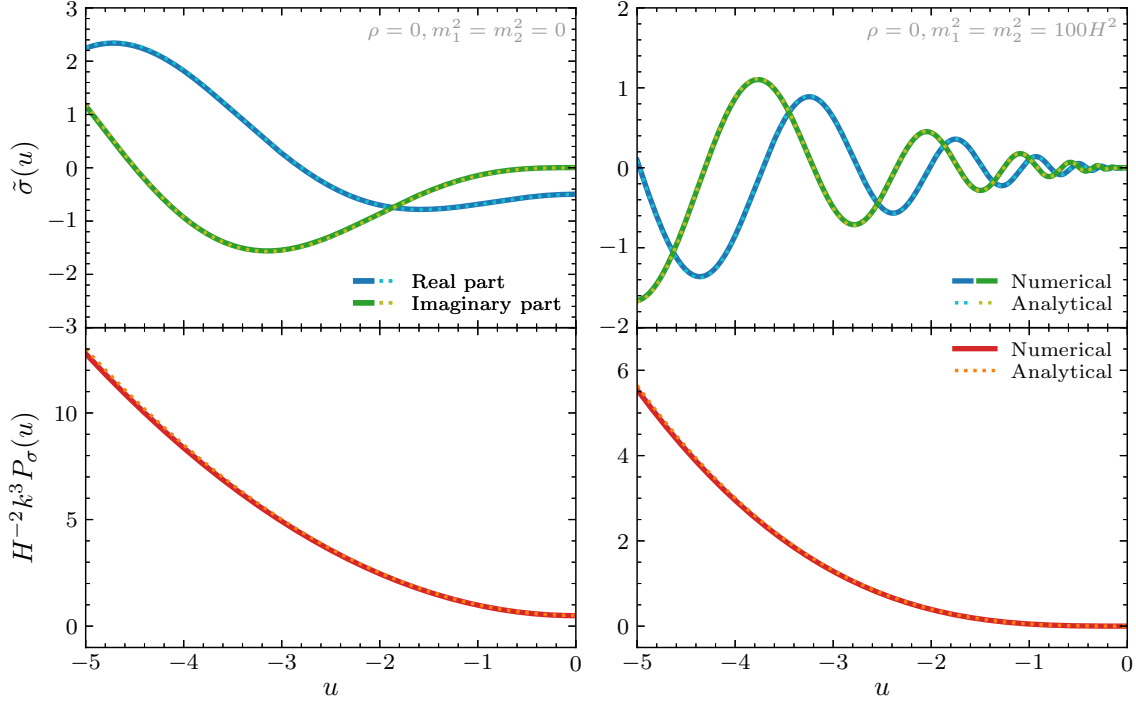


Figure 3.2. Comparison of the numerical and analytical solutions for the dimensionless mode function $\tilde{\sigma}(u)$ (*top*) and its respective power spectrum $H^{-2}k^3 P_{\sigma}(u)$ (*bottom*) as a function of $u = k\tau$ for the case of no mixing, $\rho = 0$, and equal masses of the two fields, $m_1 = m_2$. We display both the real and imaginary parts of $\tilde{\sigma}$, and note that $\tilde{\sigma} \equiv \tilde{\sigma}_A^{(+)} = -\tilde{\sigma}_A^{(-)} = i\tilde{\sigma}_B^{(\pm)}$ [cf. (3.13) since $\sigma_1 = \sigma_2$]. The left and right panels illustrate the massless scenario and an example for the massive case with $m = 10H$, respectively. The numerical solution is computed using the Dormand-Prince method with initial step size $\Delta u = 0.01$ and initial point $u_0 = -5000$ to minimize the phase error, and exhibits very good agreement with the analytic result.

use an initial step size of $\Delta u = 0.01$.

Having established that the accuracy of our numerical methods for massive fields is sufficient, we will now re-introduce $\rho \neq 0$ to find the solutions in this novel parameter regime. As discussed in §3.3.1, the mode function in the $u \rightarrow 0$ limit is a linear combination of four power-law solutions, with their dimensions being two complex-conjugate pairs. In general, the dimensionless mode functions $\tilde{\sigma}_A^{(\pm)}(u)$ and $\tilde{\sigma}_B^{(\pm)}(u)$ can consequently be parameterized as

$$\tilde{\sigma}_{A,B}^{(\pm)}(u) = a_{A,B}^{(\pm)}(-u)^\Delta + b_{A,B}^{(\pm)}(-u)^{\Delta^*} + c_{A,B}^{(\pm)}(-u)^{\bar{\Delta}} + d_{A,B}^{(\pm)}(-u)^{\bar{\Delta}^*}, \quad (3.19)$$

where Δ^* and $\bar{\Delta}^*$ are the complex conjugates of Δ and $\bar{\Delta}$, respectively, which are defined in (3.6), and the coefficients $a_{A,B}^{(\pm)}, b_{A,B}^{(\pm)}, c_{A,B}^{(\pm)}$ and $d_{A,B}^{(\pm)}$ are all complex. In the limit $u \rightarrow 0$, the two solutions with smaller α become dominant, which allows us to drop two out of the four terms in (3.19).⁵ Our expectation therefore is that our numerical solutions for $|u| \ll 1$ can be used to determine these coefficients.

We fit the power laws in (3.19) to the numerical mode functions and show several cases in Fig. 3.3, including their dominant power laws. We find that the parametrization (3.19) provides excellent fits to the numerical solutions, with the relative differences being at most at the sub-percent level. The central take-away from these solutions consequently is that the mode functions corresponding to the Bunch-Davies vacuum populate the long-wavelength scaling solutions associated with Δ and Δ^* for $\alpha \rightarrow 0$ and $\nu > 1$. Concretely, our analytic calculations confirm the existence of these solutions in the super-horizon limit, but we need the numerical computations to confirm these solutions are realized with a sufficient amplitude to dominate the super-horizon evolution. We additionally see that we can generate positive and negative scaling exponents in this two-field model and in particular observe that the novel regime where $\alpha \approx 0$ and $\nu \gg 1$ is indeed generated from vacuum fluctuations with an amplitude

⁵For $\alpha = 3/2$, some of the scaling dimensions become degenerate, i.e. $\Delta^* = \bar{\Delta}$, which requires a more careful treatment of the solutions and fitting all four coefficients in (3.19).

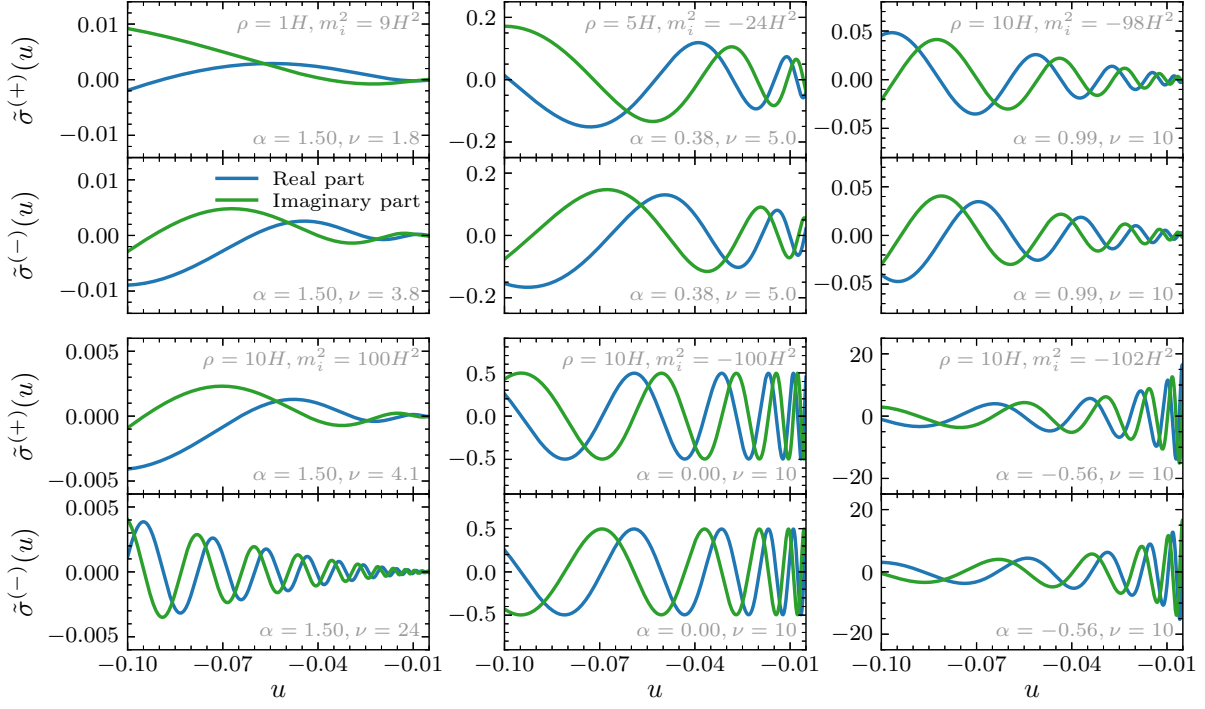


Figure 3.3. Numerical solutions for the mode functions $\tilde{\sigma}^{(\pm)}(u) \equiv \pm \tilde{\sigma}_A^{(\pm)}(u) = i \tilde{\sigma}_B^{(\pm)}(u)$ for six different cases of non-zero mixing, $\rho \neq 0$. We show the real (blue) and imaginary (green) parts for a few different scenarios, taking the masses to be equal for simplicity. Fitting the parametrization (3.19) to these curves over the shown range of u leads to excellent agreement, with differences being at most at the sub-percent level. We explicitly highlight the dominant scaling solution as parameterized by $\Delta = \alpha + i\nu$ in each panel. We observe that the oscillatory mode functions are not only non-zero in all cases, but may in fact have an appreciably large amplitude, especially for tachyonic masses in the new regime of interest with small scaling exponent α and large frequency ν .

of $O(1)$. Figure 3.3 also illustrates the sensitivity of α to the precise values of ρ and m^2 . In this regard, there is nothing special about the case $\alpha = 0$, but we see instead that α can take on a wide range of positive and negative values for a narrow range of ρ and m_i^2 .

While the power spectrum provides a useful benchmark for numerical results, the observational signatures will be driven by the unequal-time two-point functions. As a result, it is primarily the mode functions that directly control observational predictions. To avoid distractions, we therefore do not show the (equal-time) power spectra of the numerical mode functions in Fig. 3.3. Importantly, we have now explicitly seen that the mode functions can be reliably determined numerically and indeed exhibit the large oscillatory signals that we anticipated analytically. From here, we will see that the amplitude of the observational imprints are driven more by the coupling to the inflaton, rather than the intrinsic size of the statistics of the field σ .

3.3.3 Non-Gaussianity and the Squeezed Limit

Having established the existence of a phenomenologically viable parameter space for $\Delta = \alpha + i\nu$ with general α and ν , particularly small α and large ν , we want to know the amplitude of any observational signatures. In general, the signal of primordial non-Gaussianity in the scale-dependent bias of galaxies arises from the squeezed and collapsed limits of the primordial correlation functions of the adiabatic mode ζ . Since the correlations are induced entirely when the massive field is outside the horizon, these limits are fortunately particularly straightforward to theoretically understand and calculate which we do in the following.

Suppose we have a cubic interaction between our massive fields σ_i and the adiabatic inflaton mode ζ , such as given by the interaction Hamiltonian $\mathcal{H}_{\text{int}} = \kappa M_{\text{pl}}^2 \dot{H} \zeta^2 (a\sigma_1 + b\sigma_2) \equiv \kappa M_{\text{pl}}^2 \dot{H} \zeta^2 \sigma$, with coupling constant κ , Planck mass M_{pl} , and parameters a and b governing the relative coupling strengths of the massive fields to the inflaton. At a very basic level, this gives rise to the non-Gaussian signal because it means that the power spectrum of ζ is modulated by the long-wavelength σ . In this regime, the long-wavelength modes are effectively classical and

we can therefore think about this as a purely statistical correlation of the form [184, 246]

$$\left\langle \sigma(\vec{k}_1, \tau_2) \langle \zeta(\vec{k}_2, \tau_1) \zeta(\vec{k}_3, \tau_2) \rangle'_\sigma \right\rangle'_{k_1 \ll k_2 \approx k_3} \approx \langle \sigma(k_1, \tau_1) \sigma(-\vec{k}_1, \tau_2) \rangle' \frac{\partial}{\partial \sigma} P_\zeta(k_2), \quad (3.20)$$

where $\langle \zeta(\vec{k}_2) \zeta(\vec{k}_3) \rangle_\sigma$ is the two-point correlator in a classical σ background. The derivative on the right-hand side of this expression is easily calculated by treating σ as a background field and the two-point correlator of σ is computed by the methods in §3.3.2. We however have not specified τ_1 and τ_2 which play crucial roles in determining the signal when $\Delta \neq 0$.

In practice, it is the squeezed bispectrum of the adiabatic modes that is the most observationally relevant.⁶ The role of the field σ in this case is to enhance the correlations between adiabatic modes of different wavelengths. We can write the squeezed limit in the suggestive form

$$\langle \zeta(\vec{k}_1) \zeta(\vec{k}_2) \zeta(\vec{k}_3) \rangle_{k_1 \ll k_2 \approx k_3} \approx \langle \langle \zeta(\vec{k}_1) \rangle_\sigma \langle \zeta(\vec{k}_2) \zeta(\vec{k}_3) \rangle_\sigma \rangle. \quad (3.21)$$

In order to generate a large three-point function, we must introduce an additional interaction, a Hamiltonian mixing term $\mathcal{H}_{\text{mix}} = \kappa' M_{\text{pl}}^2 \dot{H} \dot{\zeta} \sigma$.⁷ Note that we have not yet specified in this expression at which time these background fields are evaluated. The correct answer of course is that we integrate over time, but understanding the scaling requires identifying the times that dominate this expression.

Interactions of ζ always involve derivatives and, therefore, exponentially decay in time t outside of the horizon. Similarly, due to the $i\epsilon$ prescription (the interacting Bunch-Davies vacuum), fluctuations inside the horizon are also exponentially suppressed. As a result, the dominant contributions to the correlations of ζ are associated with the fields at the time of horizon crossing of the corresponding mode. This implies that we should evaluate the squeezed

⁶The collapsed limit of the trispectrum is the leading signal in some models [11, 13, 86]. The generalization to that case is straightforward, but we focus on the bispectrum to simplify the presentation.

⁷In principle, there is no reason that the linear combination of σ_i in the mixing term matches the combination in the cubic interaction. To simplify the notation, we assume here that they are both the same $\sigma = a\sigma_1 + b\sigma_2$ without impacting our results.

limits at the times when $\zeta(\vec{k}_i)$ is crossing the horizon, i.e.

$$\langle \zeta(\vec{k}_1) \zeta(\vec{k}_2) \zeta(\vec{k}_3) \rangle'_{k_1 \ll k_2, k_3} \approx \text{Re} \left[\langle \sigma(\vec{k}_1, (aH)_1) \sigma(-\vec{k}_1, (aH)_2) \rangle' \frac{\partial}{\partial \sigma} P_\zeta(\vec{k}_2) \frac{\partial}{\partial \sigma} \zeta(\vec{k}_1) \right], \quad (3.22)$$

where we defined the time of horizon crossing via $(aH)_i = k_i$ and $\frac{\partial \zeta}{\partial \sigma}$ is the linear response of ζ to σ . The unequal-time correlators of σ can be evaluated using the leading power laws in the mode expansion,

$$\sigma(\vec{k}, aH) k^{3/2} = c_+ \left(\frac{k}{aH} \right)^{\alpha + i\nu} + c_- \left(\frac{k}{aH} \right)^{\alpha - i\nu}, \quad (3.23)$$

with the coefficients $c_\pm$ having the potential of being $O(1)$ as explicitly shown in the numerical calculations of §3.3.2, so that

$$\langle \sigma(\vec{k}_1, (aH)_1) \sigma(-\vec{k}_1, (aH)_2) \rangle' \propto \frac{k_1^{\alpha-3}}{k_2^\alpha} (c_+ + c_-) \left[c_+^* \left(\frac{k_1}{k_2} \right)^{-i\nu} + c_-^* \left(\frac{k_1}{k_2} \right)^{i\nu} \right]. \quad (3.24)$$

Since $(aH)_1 = k_1$, we see that the correlator scales as $(k_1/k_2)^\Delta$. Moreover, there is no possibility of canceling the oscillatory behavior unlike in the calculation at equal times. In this sense, the squeezed limit of the primordial bispectrum necessarily takes the form

$$\langle \zeta(\vec{k}_1) \zeta(\vec{k}_2) \zeta(\vec{k}_3) \rangle'_{k_1 \ll k_2, k_3} \propto \frac{1}{k_1^3 k_2^3} \frac{\partial}{\partial \sigma} \zeta(\vec{k}_1) \left(\frac{k_1}{k_2} \right)^\alpha \cos[\nu \log(k_1/k_2) + \varphi], \quad (3.25)$$

where φ is a (model-dependent) phase.

While the amplitude of the oscillations in the unequal-time correlation function of σ in (3.24) is unsuppressed, the same conclusion cannot be reached for the bispectrum. Specifically, the correlation between ζ and σ is an integrated rather than instantaneous quantity and must be evaluated with care. This means that we should actually consider

$$\frac{\partial \zeta}{\partial \sigma} \propto \text{Im} \left\{ k_1^3 \zeta(\vec{k}_1, \tau_0) \int \frac{d\tau}{(-\tau)^4} \dot{\zeta}(k_1, \tau) \left[c_+ (-k_1 \tau)^{\alpha + i\nu} + c_- (-k_1 \tau)^{\alpha - i\nu} \right] \right\}. \quad (3.26)$$

This integral can be evaluated by a Wick rotation [247] noting that

$$\int_0^\infty d\bar{\tau} e^{-\bar{\tau}} \bar{\tau}^{\alpha-1 \pm i\nu} = \Gamma(\alpha - 1 \pm i\nu) \propto \nu^{\alpha-3/2} e^{-\nu\pi/2}, \quad (3.27)$$

where $\bar{\tau} = -k_1 \tau$. Consequently, the mixing term alone is expected to exponentially suppress the amplitude at large frequencies, $\nu \gg 1$. We note that the precise form of the bispectrum is model dependent, including on both the definition of σ and the specific interactions. For an explicit calculation of the bispectrum and the trispectrum in one such model, we refer to [24], which confirms the above scaling. As a result, we expect the scaling of our signal to take the model-independent form

$$\langle \zeta(\vec{k}_1) \zeta(\vec{k}_2) \zeta(\vec{k}_3) \rangle'_{k_1 \ll k_2, k_3} \propto \frac{\nu^{\alpha-3/2} e^{-\nu\pi/2}}{k_1^3 k_2^3} \left(\frac{k_1}{k_2} \right)^\alpha \cos[\nu \log(k_1/k_2) + \varphi], \quad (3.28)$$

with the mentioned exponential suppression with frequency, $e^{-\nu\pi/2}$.

The origin of this exponential suppression has a relatively simple physical origin. In order to measure the two-point function (3.24) of the field σ , we would have to have a detector with perfect time resolution. This requires an injection of energy sufficient to produce physical particles of any mass [248], as we would expect from the uncertainty principle. When the signal is not localized in time, there is correspondingly no energy injection and the signal is Boltzmann suppressed [14]. In this particular model, the same behavior holds, although the “energy” associated with these models is determined by ν , but does not have the familiar interpretation in terms of the mass of a particle.

We have focused on the behavior in the soft limit because it plays an outsized role in the description of the large-scale structure of the Universe. As we can see from (3.21), the long-distance behavior of these soft correlators is controlled by $\sigma(k)$ outside the horizon and is largely independent of the details of how the short-wavelength ζ fluctuations behave. Moreover, these signals violate the single-field consistency relations [19, 34, 38] and, therefore,

cannot be mimicked by loop corrections to correlators, either in the early [184, 249–252] or late universe [20, 21, 253, 254]. This implies that the exponential suppression with ν , the power-law behavior with α and the logarithmic oscillations with frequency ν can generally be expected to be present in late-time cosmological observations.

While this suppression is typical, it is specific to models with an approximate time-translation symmetry that is broken at the scale H by the expansion rate. The fields in our model effectively carry energy $E \approx H\nu \gg H$ in the high-frequency regime and are consequently Boltzmann suppressed in an approximate de Sitter background. However, the inflaton background carries an effective energy of $\dot{\phi}^{1/2} \approx 60H$ [255] which is why direct couplings to the inflaton can overcome this suppression. Explicit examples have realized this idea (see e.g. [93, 176, 177, 205]) and, therefore, suggest our scaling behavior could arise without the exponential suppression.⁸

For the rest of this paper and in line with this phenomenological approach, we will consider a range of scaling exponents of $\alpha \in [-1, 3]$. While the focus of this section was on models with $\alpha \in [0, 3/2]$, this broader range for α allows for connections with a wider range of models [257]. Concretely, for $\nu = 0$, some relevant values of α include: the tachyonic collider ($\alpha = -1$) [245], local PNG ($\alpha = 0$), the conventional cosmological collider ($\alpha = 3/2$) [92], equilateral and orthogonal PNG ($\alpha = 2$), and inflationary loops or composite operators ($\alpha = 3$) [90].

3.4 Oscillations in the Galaxy Power Spectrum

We have established a novel type of oscillatory behavior in the squeezed limit of inflationary correlators in concrete multi-field models. We therefore expect that these primordial signatures are imprinted on the late-time distribution of matter as characteristic modulations in the large-scale clustering of halos and galaxies. Structure formation should translate these

⁸Some classes of models with broken time translations produce additional oscillatory signals in the power spectrum [256]. There are physical reasons these corrections can be subdominant in other examples, including chemical-potential-like couplings [176].

violations of the single-field consistency relations into a distinct scale-dependent bias. In this section, we will explicitly determine the relationship between the primordial statistics of the adiabatic mode ζ , which we determined in the previous section, and the observable statistics of galaxies, which are accessible to cosmological surveys of the large-scale structure of the Universe. We first review how the primordial bispectrum in the squeezed limit affects the halo power spectrum by modifying its bias (§3.4.1). We then translate this modification into the scale-dependent and oscillatory behavior of the galaxy power spectrum (§3.4.2). Finally, we describe the Fisher forecasts and their results for the sensitivity of current and future surveys to this class of oscillatory primordial non-Gaussianity (§3.4.3).

In the following, we primarily focus on the origin of the scale-dependent bias from primordial non-Gaussianity [10]. Due to the novel parameter regime found in the models in Section 3.3, the relationship between the primordial signal and the galaxy power spectrum does not follow directly from many conventional discussions which is why we discuss it in detail. In contrast, the nonlinear biases due to gravity and galaxy formation will be the same as those that appear in any number of analyses of LSS data and have been thoroughly reviewed in [25].

3.4.1 Oscillations in the Scale-Dependent Bias

The essential feature of the models studied in this paper is the complex scaling dimension Δ which encodes the time evolution of the additional fields. Our goal is to infer this parameter from the observed distribution of galaxies. The map from initial conditions to the statistics of galaxies, i.e. galaxy biasing, is a well-established subject with many comprehensive reviews [15, 25, 80]. Having said that, the complex scaling behavior of the primordial bispectrum (3.28) introduces a novel challenge in describing the signals accurately. This section will review some of the well-known tools in modeling galaxies with the goal of making the derivation and assumptions of our signal transparent to all readers. To some degree, the results mirror the scenario in which Δ is real [40, 86, 193]. As we have however already seen when calculating the squeezed limit of the primordial signal in §3.3.3, high-frequency oscillations demand more care

when determining the normalization of the signal.

In order to quantify the impact of primordial non-Gaussianity on the statistics of galaxies, we need a model that relates initial conditions to galaxy abundances. The essential input is the idea that galaxy formation is effectively local on cosmological scales. Specifically, we often assume that the density contrast of galaxies, $\delta_g(\vec{x})$, is determined by some unknown local function,

$$\delta_g(\vec{x}) = F_g(\delta(\vec{x}), \nabla_i \nabla_j \Phi(\vec{x}), \nabla^2 \delta(\vec{x}), \dots). \quad (3.29)$$

Here, we have allowed not only for the density perturbation $\delta(\vec{x})$ and its derivatives to appear in this expression, but also for the tidal tensor, $\nabla_i \nabla_j \Phi(\vec{x})$, with the Newtonian potential Φ , since it is allowed by the equivalence principle. In principle, Taylor expanding this expression tells us that any local operator allowed by the equivalence principle can appear in the relationship between δ_g and δ .

While this is a compelling physical argument, it presents a challenge when it comes to the initial conditions. The above expression is only local on scales larger than the scale of the halos, k_{halo} , or the nonlinear scale, k_{NL} . Since this is better expressed in terms of Fourier modes $\vec{k}$, we are considering local operators

$$\mathcal{O}(\vec{k}) \equiv \int d^3x e^{i\vec{k}\cdot\vec{x}} \mathcal{O}(\vec{x}), \quad (3.30)$$

with $k < k_{\text{NL}}$. However, for non-Gaussian correlators, this constraint does not mean we are neglecting short-wavelength modes. For example, using (3.22), we see that

$$\langle \delta^2(\vec{k}) \delta(\vec{k}') \rangle' = \int \frac{d^3k_s}{(2\pi)^3} \langle \delta(\vec{k}_s) \delta(\vec{k} - \vec{k}_s) \delta(\vec{k}') \rangle' \propto \langle \sigma(\vec{k}, (aH)_s) \sigma(\vec{k}', aH) \rangle'. \quad (3.31)$$

In this sense, in a theory with primordial non-Gaussianity, $\delta^2(\vec{k})$ behaves at a practical level like a proxy for $\sigma(\vec{k})$ as much as it encodes the local density squared.

The unusual dependence on initial conditions is related to the fact that the long-distance

description of the large-scale structure is local in space, but nonlocal in time. One consequence is that $\sigma(\vec{x}, t_*)$ effectively plays a role in the clustering of matter at a time $t \gg t_*$. In addition, the biasing model in (3.29) defines the halos in terms of these locations in the initial conditions (Lagrangian coordinates), which are then mapped to biases in Eulerian coordinates through the motion of the underlying structures through gravitational evolution.

In order to provide a concrete model for the formation of halos (and, therefore, the statistics of halos and galaxies), we will use the barrier-crossing formalism [258]. The key input is the smoothed density field,

$$\delta_M(x) = \int \frac{d^3k}{(2\pi)^3} e^{-i\vec{k}\cdot\vec{x}} W_M(k) \delta(\vec{k}), \quad W_M(k) = 3 \frac{\sin(kR_M) - kR_M \cos(kR_M)}{(kR_M)^3}, \quad (3.32)$$

where R_M is the Lagrangian radius corresponding to the halo mass M and W_M is the Fourier transform of the associated top-hat window function. An important simplifying feature of using the smoothed density field is that it eliminates short modes from the description of the bias and moves the squeezed PNG into the scale dependence of the bias.

The barrier-crossing model defines the number density of halos in terms of the linear density above a critical threshold δ_c . Following [86], it is first useful to define the Edgeworth expansion of the probability distribution of the matter overdensity. To determine the power spectrum of galaxies, we will need the two-point probability distribution

$$p(v, v') = \exp\left(\kappa_{1,1} \frac{\partial^2}{\partial v \partial v'} + \sum_{m+n \geq 3} \frac{(-1)^{m+n}}{m!n!} \kappa_{m,n} \frac{\partial^{m+n}}{\partial v^m \partial (v')^n}\right) p_g(v) p_g(v'), \quad (3.33)$$

where $v^{(r)} = \delta(\vec{x}^{(r)})/\sigma_M$, σ_M is the root-mean-square amplitude of δ_M and $p_g(v)$ is the normal distribution. The coefficients of this expansion are the cumulants

$$\kappa_{n,m}(r = |\vec{x} - \vec{x}'|, M, \bar{M}) = \frac{\langle [\delta_M^n](\vec{x}) [\delta_M^m](\vec{x}') \rangle'}{\sigma_M^n \sigma_{\bar{M}}^m}, \quad (3.34)$$

where the notation $[\mathcal{O}]$ means the operator $\mathcal{O}$ minus all self-contractions (or, equivalently, that we only evaluate the connected correlator). In cosmological applications, we frequently also need to determine the cross-correlation between the galaxies and the matter overdensity,

$$p(\mathbf{v}|\delta(\vec{x}')) = \exp \left[\sum_{n \geq 1} \frac{(-1)^n}{n!} \kappa_{\hat{1},n} \frac{\partial^n}{\partial \mathbf{v}^n} \right] p_g(\mathbf{v}), \quad (3.35)$$

with the cumulants relevant here being defined as

$$\kappa_{\hat{1},m}(r = |\vec{x} - \vec{x}'|, M) = \frac{\langle [\delta^n](\vec{x}) [\delta_M^m](\vec{x}') \rangle'}{\sigma^n \sigma_M^m}. \quad (3.36)$$

This is, in effect, a definition of the linear bias. In order to therefore determine the scale-dependent bias, we need to calculate

$$\kappa_{\hat{1},n}(k, M) \equiv \int d^3x e^{i\vec{k} \cdot \vec{x}} \kappa_{\hat{1},n}(r, M) = \int d^3x e^{i\vec{k} \cdot \vec{x}} \frac{\langle \delta^{(1)}(\vec{x}) [\delta_M^n](\vec{x}') \rangle'}{\hat{\sigma} \sigma_M^n}, \quad (3.37)$$

where $\delta^{(1)}$ is the linear matter overdensity and $\hat{\sigma}$ is its root-mean-square amplitude.

In addition, we now define $\delta^{(1)}(\vec{k}, z) = k^2 \mathcal{T}(k, z) \Phi(\vec{k})$, with $\mathcal{T}(k, z) = \frac{2T(k)D(z)}{3H_0^2 \Omega_m}$, where $T(k)$ is the transfer function defined such that $T(k \rightarrow 0) \rightarrow 1$, $D(z)$ is the linear growth factor normalized as $D(z) = 1/(1+z)$ during matter domination, H_0 is the Hubble constant and Ω_m is the matter density. Although measuring the matter-galaxy cross-correlation is not possible with BOSS (or other galaxy surveys) directly, the calculation of the cross-correlation is important for distinguishing the stochastic and deterministic scale-dependent biases.

We can determine the relevant deterministic bias parameters by taking the $k \rightarrow 0$ limit

of $\kappa_{\hat{1},n}$. We will only need the first two cumulants,

$$\kappa_{\hat{1},1}(k, M) \xrightarrow{k \rightarrow 0} \frac{P_{\text{mm}}(k)}{\hat{\sigma} \sigma_M}, \quad (3.38)$$

$$\begin{aligned} \kappa_{\hat{1},2}(k, M) &\xrightarrow{k \rightarrow 0} \hat{\sigma}^{-1} \sigma_M^{-2} \int_{\vec{q}_1} \int_{\vec{q}_2} k^6 W_M(q_1) W_M(q_2) \mathcal{T}(q_1) \mathcal{T}(q_2) \mathcal{T}(k) \langle [\Phi](\vec{k}) [\Phi](\vec{q}_1) [\Phi](\vec{q}_2) \rangle \\ &= \frac{4\lambda f_{\text{NL}}^{\alpha, \nu}}{\hat{\sigma} \sigma_M^2} \frac{P_{\text{mm}}(k)}{k^2 \mathcal{T}(k)} (kR_M)^\Delta \int_{\vec{q}} W_M(q)^2 P_{\text{mm}}(q) (qR_M)^{-\Delta}, \end{aligned} \quad (3.39)$$

where $\int_{\vec{q}} \equiv \int \frac{d^3 q}{(2\pi)^3}$ and P_{mm} is the standard (Gaussian) matter autocorrelation power spectrum. As we will explain in §3.4.2, we will need the real part of this expression for complex-valued exponents Δ . We note that we reintroduced here the relative normalization of the squeezed limit in terms of the parameter λ which is defined in (3.1). To simplify the equations, we will hereafter denote the integral on the right-hand side as

$$\sigma_M^2(\Delta) \equiv \int_{\vec{q}} W_M(q)^2 P_{\text{mm}}(q) (qR_M)^{-\Delta}. \quad (3.40)$$

This integral isolates the bispectrum shapes that contribute to the long-distance scale-dependent bias and are more easily measured.

The leading contribution to the halo-matter power spectrum is then given by

$$\begin{aligned} P_{\text{mh}}(k) &\xrightarrow{k \rightarrow 0} \left[b_G + \frac{1}{2} \delta_c b_G^L \frac{\kappa_{\hat{1},2}(k, M)}{\kappa_{\hat{1},1}(k, M) \sigma_M} + \dots \right] P_{\text{mm}}(k) \\ &= \left[b_G + 2\lambda f_{\text{NL}}^{\alpha, \nu} \delta_c (b_G - 1) \frac{(kR_M)^\Delta}{k^2 \mathcal{T}(k)} \frac{\sigma_M^2(\Delta)}{\sigma_M^2} + \dots \right] P_{\text{mm}}(k), \end{aligned} \quad (3.41)$$

where b_G is the linear bias in the limit of purely Gaussian initial conditions and the additional terms denoted by $\dots$ are of $O(f_{\text{NL}}^2)$, with the contributions of higher-order biases vanishing in the $k \rightarrow 0$ limit. For later convenience, we will define the second term in the bracket as the

non-Gaussian bias

$$b_{\text{NG}}(k; M, \Delta) \equiv 2\lambda f_{\text{NL}}^{\alpha, \nu} \delta_c(b_{\text{G}} - 1) \frac{(kR_M)^\Delta}{k^2 \mathcal{T}(k)} \frac{\sigma_M^2(\Delta)}{\sigma_M^2} \propto \frac{k^\Delta}{k^2 \mathcal{T}(k)}. \quad (3.42)$$

The essential feature of this expression is that the scale-dependent term scales as $k^{\Delta-2}$ at small wavenumbers k since $\mathcal{T}(k \rightarrow 0) \rightarrow \text{const.}$ In addition, the amplitude of the signal is controlled by $\sigma_M^2(\Delta)$.

What we actually observe from galaxy surveys is the halo-halo correlation, which involves an additional contribution from the primordial trispectrum whose amplitude is $O(f_{\text{NL}}^2)$ and should be sub-dominant compared to the term linear in f_{NL} . For simplicity, we will approximate this contribution using b_{NG}^2 , so that the halo-halo correlation in Fourier space is

$$P_{\text{hh}}(k) \xrightarrow{k \rightarrow 0} [b_{\text{G}} + b_{\text{NG}}(k; M, \Delta) + \dots]^2 P_{\text{mm}}(k). \quad (3.43)$$

Here, we are working in the small f_{NL} approximation and are ignoring the stochastic halo biases that are of $O(f_{\text{NL}}^2)$. Some of these terms contribute to the stochastic bias, which is given a more rigorous treatment in [86].

The key take-away from this result is that the shape and behavior of the scale-dependent bias are determined by the scaling dimension Δ . For instance, the scale-dependent bias has a power-law behavior in k when the field coupled to the inflaton is light so that $\Delta = \alpha$ is real. On the other hand, when the scaling dimension is complex, $\Delta = \alpha + i\nu$, we can rewrite $(kR_M)^\Delta = (kR_M)^\alpha e^{i\nu \log(kR_M)}$ which shows that the imaginary component manifests itself in the power spectrum as an oscillatory feature in $\log k$. It is also useful to note that this oscillation arises from a scale-invariant initial condition due to the breaking of scale invariance associated with the size of the halo.

3.4.2 Modeling the Galaxy Power Spectrum

Having derived the contribution to the halo-halo power spectrum from a generic primordial non-Gaussian signal, we can now discuss how to parameterize the galaxy-galaxy power spectrum in order to extract this signal. Our goal is to measure the amplitude of the non-Gaussian signal in a model-independent way while accounting for uncertainties in galaxy biasing. We will therefore discuss the new non-Gaussian contributions in detail and mention the standard nonlinear biases at the end.

We parameterize the theoretical galaxy power spectrum by separating the scale-dependent bias contribution into a smooth power-law piece, $S(k)$, and an oscillatory modulation, $O(k)$, such that (3.43) is written as⁹

$$P_g(k) = [b_1 + 3f_{\text{NL}}^{\alpha, \nu} b_\phi |\Sigma_M^2(\Delta)| S(k; M, \Delta) O(k; M, \Delta) + \dots]^2 P_{\text{mm}}(k), \quad (3.44)$$

where $b_\phi = 2\delta_c(b_1 - 1)$ following the universality relation,¹⁰ $\Sigma_M^2(\Delta) \equiv \sigma_M^2(\Delta)/\sigma_M^2$ and we have taken $\lambda = 3$ to match the normalization for the equilateral template used especially in [40, 94, 193] when $\Delta = 2$.¹¹ To make a distinction between the bias for halos and galaxies, we switched the notation from b_G to the standard notation for the linear galaxy bias as b_1 . We note that the form and parameter values of the universality relation may change for different galaxy populations or types of primordial non-Gaussianity, for instance, and should therefore be used with care (cf. [98, 100, 144, 146–148, 152–155, 259–263]). Having said that, we will generally

⁹Since we are only considering the galaxy auto-power spectrum and no cross-spectra, we are denoting it by $P_g \equiv P_{\text{gg}}$ for simplicity.

¹⁰A simple relation between the non-Gaussian bias b_ϕ and the linear Gaussian bias b_G can be derived in halo-bias theory [25]. For local PNG, $b_\phi = 2\delta_c(b_G - p)$ can be obtained, where $\delta_c \approx 1.686$ is the critical overdensity for spherical collapse and $p = 1$ when assuming a universal halo mass function. While this relationship was originally derived for dark matter halos, it is usually also applied to galaxies by using their linear bias b_1 instead of b_G .

¹¹Even for equilateral non-Gaussianity, the normalization in the squeezed limit is model dependent and only has $\lambda = 3$ for the factorizable template. The normalization obeys a non-trivial soft theorem that depends on the specific interaction [88] and is exponentially sensitive to the frequency ν which we implicitly absorb into the PNG amplitude $f_{\text{NL}}^{\alpha, \nu}$.

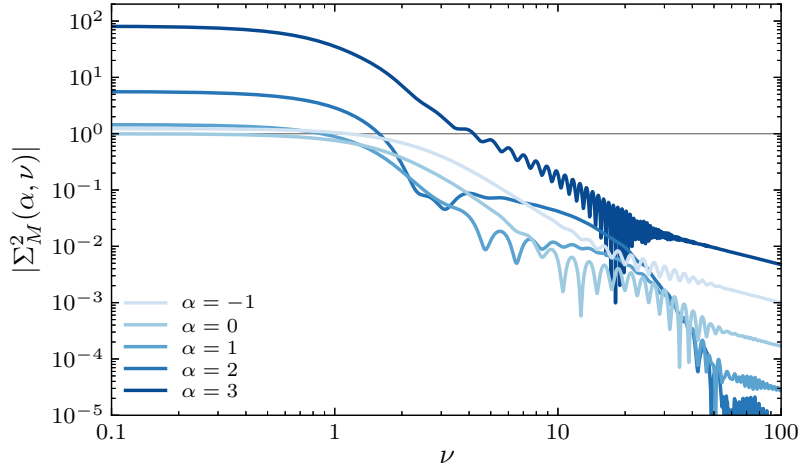


Figure 3.4. Suppression of the overall amplitude of the oscillatory PNG signal in the galaxy power spectrum as captured by $|\Sigma_M^2(\alpha, \nu)|$ in (3.44) for a halo mass of $M = 10^{13} M_\odot$. For all values of the scaling exponent α , $|\Sigma_M^2|$ sharply decreases at large frequencies ν since we integrate over the highly oscillatory signal. It approaches a constant for small frequencies which is (close to) unity for (around) $\alpha = 0$, as expected, but is $O(100)$ for $\alpha = 3$ since the integrand in (3.40) peaks around the scale of matter-radiation equality and is additionally enhanced by $(qR_M)^{-\alpha} \gg 1$ at that scale. We note that we absorb this suppression factor into the PNG amplitude $f_{\text{NL}}^{\alpha, \nu}$ for the rest of our discussion, forecasts and data analysis. This is because we are mostly interested in investigating the sensitivity to a potential signal in this new PNG class and this suppression factor can easily be reinstated to get specific model predictions and results.

apply and use this particular form of the relation throughout this work.¹²

The numerical results for $|\Sigma_M^2(\Delta)| = |\Sigma_M^2(\alpha, \nu)|$ are shown in Fig. 3.4 as a function of frequency ν for the range of scaling exponents α considered in this work. We note that $\Sigma_M(\Delta)$ may be complex valued since Δ may be complex. We therefore factored out the absolute value $|\Sigma_M^2(\Delta)|$ in (3.44) so that the term $S(k; M, \Delta)O(k; M, \Delta)$ carries only the phase information of $\sigma_M^2(\Delta)$. The overall magnitude of $|\Sigma_M^2(\Delta)|$, especially for small ν , depends on the range of wavenumbers in which the integrand in (3.40) peaks. For $\alpha = 0$, this is the case around $0.5 h \text{Mpc}^{-1}$ which shifts to larger (smaller) scales for positive (negative) α , reaching approximately the scale of matter-radiation equality for $\alpha = 3$. In the latter case, the peak is additionally enhanced since $(qR_M)^{-\alpha} \gg 1$ for those scales. For $\nu > O(1)$, we see a large suppression of the signal due to the

¹²It is interesting to note in this context that when we marginalized over the universality relation with a broad Gaussian prior for non-oscillatory PNG in [193], we found the same qualitative behavior with some degradation in constraining power.

integration over the rapidly oscillating correlator.

Comparing (3.44) with (3.42) and (3.43), it remains to define the two terms $S(k; M, \Delta)$ and $O(k; M, \Delta)$. We take the former to be the power-law part of the scale-dependent bias, which only depends on the real part of Δ and the halo mass M ,

$$S(k; M, \Delta) = S(k; M, \alpha) = \frac{(kR_M)^\alpha}{k^2 \mathcal{T}(k)}. \quad (3.45)$$

For complex Δ , i.e. if $\nu \neq 0$, the latter describes the oscillatory part of the scale-dependent bias, which only depends on the frequency ν (and the halo mass M),

$$f_{\text{NL}}^{\alpha, \nu} O(k; M, \Delta) = f_{\text{NL}}^{\alpha, \nu} O(k; M, \nu) = A_{\cos} \cos[\nu \log(kR_M)] + A_{\sin} \sin[\nu \log(kR_M)], \quad (3.46)$$

i.e. we parameterize the oscillations in the amplitude parametrization with two amplitudes $A_{\cos}$ and $A_{\sin}$. For notational convenience, we defined these two parameters to effectively absorb the PNG parameter $f_{\text{NL}}^{\alpha, \nu}$, but we will only use this as an intermediate step and state our results again in terms of $f_{\text{NL}}^{\alpha, \nu}$. In addition, it is implied that $f_{\text{NL}}^{\alpha, \nu} O(k) = f_{\text{NL}}^{\alpha, \nu}$ for $\nu = 0$ and a change of the pivot scale in the logarithmic oscillations, which we took as R_M , would only lead to a physically irrelevant shift in the phase.

We note that it would have been equally possible to describe the oscillations in terms of one amplitude and a phase, i.e. in the phase parametrization. While this can be a conceptually simpler way to understand the signal, it can complicate a parameter inference, whether for a forecast or a data analysis. This same challenge is present in searches for primordial features in the power spectrum and we refer to [18] for a detailed discussion of the merits and shortcomings of these parametrizations. In short, the posterior distributions for $A_{\cos}$ and $A_{\sin}$ are both Gaussians that are mostly uncorrelated at large ν and peak at zero for a vanishing signal. This contrasts with the phase parametrization, with a phase $\varphi \in [0, 2\pi)$, for which the phase will be uniformly distributed in the absence of a signal, while the amplitude will have a positive prior. In this regard,

Fisher forecasts in the phase parametrization may not accurately capture these non-Gaussian priors and can therefore be a weaker approximation to the results of Markov chain Monte Carlo analyses like those used in Section 3.5.

Throughout the rest of this paper, we adopt two approximations for the theoretical non-Gaussian signal or, equivalently, the normalization of $f_{\text{NL}}^{\alpha, \nu}$. First, we know that large-scale gravitational nonlinearities exponentially damp any oscillatory signal in the late universe [18] (see also [45–47, 228–232]). While it is important to take this damping into account in general, we are able to neglect it in the following because we make conservative scale cuts given the rest of our modeling (and, more generally, the original signal can be reconstructed on smaller scales to some level). This means that the damping effect remains small over the range of wavenumbers that we are taking into account (see e.g. Fig. 19 of [18] for an illustration). More generally, our work is of a more exploratory nature and with results shown on a logarithmic scale for the PNG amplitude which means that corrections to the amplitude at the $O(10\%)$ level will not have a visible impact. Second, we saw in Fig. 3.4 that $|\Sigma_M^2(\alpha, \nu)|$ is suppressing the non-Gaussian oscillations by many orders of magnitude for large frequencies ν . In our forecasts and data analysis, we are however mostly interested in investigating the sensitivity of the surveys to a potential signal. This is why we set this term to unity, $|\Sigma_M^2(\alpha, \nu)| \rightarrow 1$, or alternatively absorb it into a redefinition of the PNG parameter, $f_{\text{NL}}^{\alpha, \nu} \rightarrow f_{\text{NL}}^{\alpha, \nu} / |\Sigma_M^2(\alpha, \nu)|$.¹³ As discussed in Section 3.3, the primordial signal itself may be exponentially suppressed or enhanced in a given inflationary model so that these factors can be reinstated when projecting the inferred constraints on the model space or specific scenarios.

The shape of the signal is shown in Fig. 3.5 for a range of (real-valued) scaling exponents α and frequencies ν (see also Fig. 3.1). We note that the oscillatory signal looks the same

¹³We therefore effectively remove an $O(1)$ number for small and vanishing ν (and scaling exponents around zero), cf. Fig. 3.4, which we had implicitly done in [193] following [40]. For $\alpha \gtrsim 2$ and small ν , we however remove a larger factor. For the two-field model of Section 3.3, this is consequently the second exponential factor in ν that we absorb into the PNG amplitude. More generally speaking, our convention isolates the survey sensitivity to the phenomenological signature of the combined power law and logarithmic oscillations, and model-dependent suppression factors can be reinstated when mapping inferred constraints to specific inflationary scenarios.

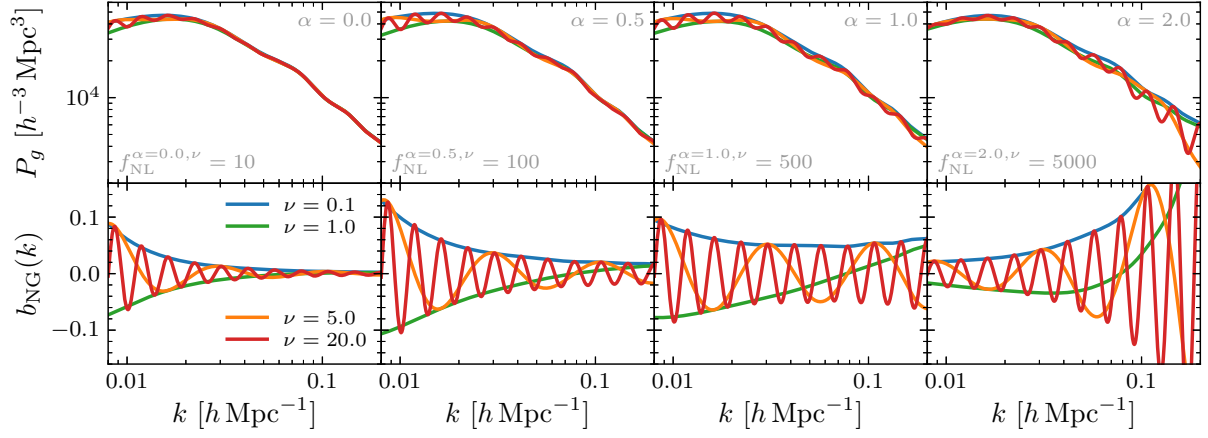


Figure 3.5. Illustration of the effect of the oscillatory scale-dependent bias on the linear BOSS galaxy power spectrum at $z = 0.35$, with the phase chosen so that $A_{\text{sin}} = 0$. The top panels display the power spectra $P_g(k)$ for different values of the scaling exponent α and frequency ν , while the bottom panels show the non-Gaussian bias $b_{\text{NG}}(k)$. We observe how the non-Gaussian signature moves from the largest to the smallest scales as we increase the scaling exponent α from the leftmost to the rightmost panels. (Note that the PNG amplitude $f_{\text{NL}}^{\alpha, \nu}$ also increases between the panels and we have already absorbed the suppression factor $|\Sigma_M^2|$ into its normalization as explained in the main text.) We can also see how the logarithmic oscillations with frequency ν are modulated by the power-law contribution, which is governed by α , and that we would only measure part of an oscillation for small frequencies given the limited range of wavenumbers (see main text).

on all scales since these are logarithmic oscillations on a logarithmic wavenumber axis. We can already anticipate degeneracies with the smooth broadband galaxy power spectrum, particularly for small frequencies, because only a fraction of a period is covered in the finite range of k . On the other hand, the oscillatory behavior of the signal, especially for high frequencies, will clearly break these degeneracies. At the same time, the signal-to-noise ratio is not uniform in k both for sample variance due to the finite cosmic volume and the finite number density of objects in a realistic cosmological survey.¹⁴ As a result, the leading constraints may arise from limited ranges in k for which the oscillations have the largest signal-to-noise ratio. As can be anticipated from Figures 3.1 and 3.5, this directly depends on the value of α which will also be evident in our forecasts below.

Our treatment of the non-Gaussian signal as a complex power law is a simplification that is valid for small $f_{\text{NL}}^{\alpha, \nu}$. Nonlinear corrections will generate additional frequencies, $n\nu$ for integer n , that are suppressed by f_{NL}^n . This phenomenon arises for any nonlinear amplifier and is suppressed by the amplitude of the oscillatory signal (and not just the nonlinear scale). Additionally, there are stochastic bias terms proportional to the collapsed trispectrum that can contain oscillating and power-law behavior that we have neglected. This should not qualitatively affect our results given the existing bounds on primordial non-Gaussianity and features which allow us to anticipate that the potential non-Gaussian signal at any given frequency is small and uncorrelated across frequencies. This is also why we can analyze one pair of (α, ν) at a time. Having said that, the constraints between frequencies will themselves be correlated since we cannot isolate each logarithmic frequency in our analysis ahead of time (unlike for a time-series analysis with white noise).

In addition to the linear and scale-dependent biases, we also have to model nonlinear corrections to the galaxy spectrum. We will follow the same bias expansion as [193] which

¹⁴The wavenumber dependence of the signal-to-noise ratio depends primarily on the shape of the scale-dependent bias and power spectrum relative to the sample variance and shot noise. This means that shot noise being independent of k does not simplify our understanding of the constraints.

includes operators up to third order in δ_m . At the field level, the bias parametrization is given by

$$\begin{aligned} \delta_g(\vec{x}) = & b_1 \delta_m(\vec{x}) + b_{\nabla^2} R_M^2 \nabla^2 \delta_m(\vec{x}) + b_{\delta^2} (\delta_m^2(\vec{x}) - \langle \delta_m^2 \rangle) + b_{s^2} (s^2(\vec{x}) - \langle s^2 \rangle) \\ & + b_{\Pi\Pi^{[2]}} \text{Tr} \left[\Pi^{[1]} \Pi^{[2]} \right] \Big|^{(3)}(\vec{x}) + \dots, \end{aligned} \quad (3.47)$$

where $s_{ij} = (\nabla_i \nabla_j / \nabla^2 - \delta_{ij} / 3) \delta_m(\vec{x})$ is the tidal tensor and the tensors $\Pi_{ij}^{[n]}$ are defined in [25], with the superscript (3) indicating that the operator includes terms up to third order in perturbation theory. We additionally include redshift-space distortions and higher-derivative terms through a change to the linear bias,

$$b_1(z) \rightarrow b(k, \mu, z) = b_1(z) + f(z) \mu^2 + \sum_{n>0} b_{k^{2n}}(z) (kR_M)^{2n}, \quad (3.48)$$

with the cosine μ between the wavevector $\vec{k}$ and the line of sight, and the growth rate $f(a) = \text{dlog } D(a) / \text{dlog } a$, where $D(a)$ is the linear growth factor, while keeping $b_\phi = b_\phi(z)$ as is.

While the details of our modeling, forecasting and eventual analysis are largely the same as in [193], it is noteworthy that in the regime $v > 1$, the detailed description of the nonlinear galaxy power spectrum has a limited impact on the sensitivity to our non-Gaussian signal. This is an expected property of oscillatory features [18, 45]. Since these details do not significantly affect our main results, we refer to §2.3.1 and Appendix A.1 for a complete description of the nonlinear galaxy power spectrum with Gaussian initial conditions, including the observational Alcock-Paczynski effect and the theoretical one-loop corrections.

The key take-away of this section is that we describe our non-Gaussian oscillatory signal in the galaxy power spectrum at small wavenumbers k by

$$P_g(k) = \left[b_1 + 3b_\phi \frac{(kR_M)^\alpha}{k^2 \mathcal{T}(k)} (A_{\cos} \cos[v \log(kR_M)] + A_{\sin} \sin[v \log(kR_M)]) \right]^2 P_{\text{mm}}(k), \quad (3.49)$$

where the PNG amplitude $f_{\text{NL}}^{\alpha, v}$ is determined from the oscillatory amplitudes $A_{\cos}$ and $A_{\sin}$

by (3.46). This expression receives corrections from higher-order biasing given by (3.47) and (3.48) that are important and taken into account at larger wavenumbers and in redshift space.

3.4.3 Forecasts for Current and Future Surveys

Having translated the phenomenological two-field model of Section 3.3 into a specific, but generalized observational signal as a combination of power-law and oscillatory PNG in the galaxy power spectrum, we are now ready to examine the signal in realistic surveys. Our forecasting methodology generally follows the detailed description in [193], with some additional aspects explained in Appendix B.1. In the following, we will summarize our approach based on the Fisher information matrix, provide our forecasted results for current and future surveys, and discuss their main implications and differences to the case of power-law PNG ($\nu = 0$), which we comprehensively studied in [193].

We compute the standard Fisher matrix for four galaxy surveys: BOSS [77], which is completed, the Dark Energy Spectroscopic Instrument (DESI) [104]¹⁵ and the Spectro-Photometer for the History of the Universe, Epoch of Reionization and Ices Explorer (SPHEREx) [71], which are currently taking data, and the billion-object survey, which we introduced in [193] to investigate the potential capabilities of a more futuristic survey. The former are modeled to roughly replicate the specifications of the existing surveys once completed, while the latter assumes one billion spectroscopic objects split into two target samples with different linear galaxy biases and distributed with a constant number density over 10 bins in the redshift range $0 \leq z \leq 5$ covering half the sky. This could therefore be thought of as a spectroscopic follow-up survey to the Vera Rubin Observatory’s Legacy Survey of Space and Time (LSST [107]; cf. the MUltiplexed Survey Telescope [264] and Spec-S5 [54] as steps in this direction) and/or an estimation for a 21 cm or other line intensity mapping survey over these redshifts (see e.g. [115–120, 178, 265]). Our detailed specifications for these surveys, including the values of the linear bias and number

¹⁵Since DESI and Euclid [106] are expected to have very similar constraining power for our non-Gaussian signal (cf. [193]), we only include one of these current surveys explicitly as part of our reported results.

density in each redshift bin, can be found in Appendix A of [193].

We model the galaxy power spectrum including our non-Gaussian signatures as described in §3.4.2, cf. (3.49). Unless stated otherwise, we vary the cosmological parameters $\{\omega_b, \omega_c, \theta_s, \log(10^{10}A_s), n_s\}$ and oscillatory PNG parameters $\{A_{\cos}, A_{\sin}\}$ defined in (3.46) for a given survey, while the bias parameters $\{b_1, b_{k^2}, b_{k^4}, b_{\delta^2}, b_{s^2}, b_{\text{III}[2]}\}$ are independently varied in each redshift bin.¹⁶ The fiducial cosmology is the same Λ CDM cosmology with Gaussian initial conditions as in [193],¹⁷ i.e. $A_{\cos} = A_{\sin} = 0$, and we also impose a marginalized Planck-like CMB Fisher matrix as a Λ CDM prior.¹⁸ We set the minimum wavenumber $k_{\min}$ of a given redshift bin by its volume assuming a spherical geometry. The maximum wavenumber $k_{\max}$ is taken to be the minimum of the nonlinear scale $k_{\text{NL}} = \frac{\pi}{2R_{\text{NL}}}$, with the radius R_{NL} at which the variance of linear fluctuations is halved at a given z , and the wavenumber $k_{\text{halo}} \approx 0.19 h \text{Mpc}^{-1}$ associated with the Lagrangian scale $R_M \approx 2.66 h^{-1} \text{Mpc}$ corresponding to the minimum halo mass of the surveyed population which we generally assume to be $10^{13} M_{\odot}$. Our choices for both $k_{\min}$ and $k_{\max}$ are conservative. In this work, we implement two extensions of our forecasting methodology compared to [193] which are both necessitated by the oscillations on top of the previously studied smooth PNG signal, which we discuss in the following and refer to Appendix B.1 for further details: using bandpowers and reporting 95% upper limits on the PNG amplitude.

Surveys can only observe a finite cosmic volume which implies that we can only access

¹⁶The Λ CDM parameters are the physical baryon density ω_b , the physical cold dark matter density ω_c , the angular size θ_s of the sound horizon at recombination, the (logarithm of the) primordial scalar amplitude A_s and the scalar spectral index n_s . The bias parameters that we include are the linear bias b_1 , the gradient biases up to quartic order b_{k^2} and b_{k^4} , the tidal bias b_{s^2} and the evolution bias $b_{\text{III}[2]}$ defined in (3.47) or explicitly in terms of the theoretical galaxy power spectrum in (2.11).

¹⁷The fiducial values of the linear biases are provided in Appendix A for all surveys together with their other specifications. The fiducial values of the nonlinear bias parameters are set to zero except for BOSS, for which we assumed fiducial values for b_{s^2} and $b_{\text{III}[2]}$ based on the Lagrangian local-in-matter-density biasing model [25] and for b_{δ^2} based on the halo-simulation fit of [81] to align with previous and current analyses. As expected from [193], this different choice in the fiducial values for BOSS has only minimal effects on the forecasted constraints.

¹⁸For BOSS, we impose the same Gaussian priors on the biases b_{δ^2} , b_{s^2} and $b_{\text{III}[2]}$ as [123]. For all other surveys, we take the priors to be uniform for all biases. Similar to the findings in [193], imposing priors on the loop biases results in $O(1)$ improvements in the forecasted constraints for small v , but leads to only minor changes of around a few percent for $v \gtrsim 10$.

a limited number of discrete wavenumbers k_i . For sufficiently small spacing Δk between these wavenumbers, we can approximate a smooth enough (theoretical) power spectrum $P(k)$, such as the Λ CDM power spectrum with Gaussian or power-law non-Gaussian initial conditions, in a given band $[k_i - \Delta/2, k_i + \Delta/2]$ by taking the value at its midpoint, $P(k_i)$. On the other hand, the logarithmic oscillations (3.46) lead to significant variations within bands depending on the frequency ν and wavenumber k_i . This is why we compute the finite-size bandpowers P_i as linear averages of the galaxy power spectrum within these bands taking $\Delta k = k_{\min}$. We refer to Appendix B.1 for additional details, and to Fig. 4 of [18] for an illustration of the effect on the observed signal and the reason why we do not additionally need to take the survey window function into account.

As a result of our forecasts and analyses, we would like to report constraints on the overall non-Gaussian signal. Given its oscillatory nature and our use of the amplitude parametrization with the two amplitudes $A_{\cos}$ and $A_{\sin}$ instead of the phase parametrization with an amplitude and a phase for good reasons [see our discussion below (3.46)], this requires more care than in the case of simple power-law PNG. Motivated by the phase parametrization, we define the non-Gaussian amplitude of our signal (3.44) as

$$Af_{\text{NL}}^{\alpha, \nu} \equiv \sqrt{A_{\cos}^2 + A_{\sin}^2}, \quad (3.50)$$

which is positive semi-definite. Phenomenologically, there is no preferred phase and the values of the amplitudes are unknown. We therefore do not assume any prior knowledge of $A_{\cos}$ or $A_{\sin}$ and the forecasted constraints on $Af_{\text{NL}}^{\alpha, \nu}$ should be defined so that they are independent of the fiducial values for $(A_{\cos}, A_{\sin})$ or, equivalently, the fiducial phase. First, as noted, we take the fiducial values for both amplitudes to be zero, i.e. the fiducial cosmology has Gaussian initial conditions. Second, following [18], we describe the sensitivity of the surveys in terms of the 95% exclusion limit of $Af_{\text{NL}}^{\alpha, \nu}$ which is a convenient estimate of the constraining power that however does not quite articulate the statistical power at small frequencies due to the strong degeneracy between

the two amplitudes $A_{\cos}$ and $A_{\sin}$. Given the definition (3.50), this means in practice that we compute the radius of the circle in the marginalized $(A_{\cos}, A_{\sin})$ parameter plane that contains 95% of the forecasted probability. In practice, we perform this calculation by Monte Carlo integration of the two-dimensional Gaussian posterior distribution for $(A_{\cos}, A_{\sin})$ around their origin given their marginalized Fisher matrix.¹⁹ This way of computing the survey sensitivity to our oscillatory PNG signal is both robust to changes of the fiducial values and implicitly marginalizes over the phase information (see Appendix B.1 for more details). We additionally note that the sensitivity metric to compare to for the non-oscillatory power-law non-Gaussianity of [193] with $\nu = 0$ is the 95% limit on the absolute value of its amplitude, $|f_{\text{NL}}^{\alpha, \nu}|$ given the positive semi-definite nature of (3.50).

We present the forecasted 95% exclusion limits on the amplitude $Af_{\text{NL}}^{\alpha, \nu}$ of our oscillatory PNG signal of (3.49) in Fig. 3.6. For all displayed experiments (BOSS, DESI, SPHEREx and the billion-object survey), we can observe a few general trends and regimes as a function of the scaling exponent α and frequency ν . First, we can generally see that the constraints at small and intermediate frequencies change by several orders of magnitude with increasing α , but remain within about one order of magnitude at large frequencies indicating the impact of the oscillations. More generally speaking, we expect that oscillatory features in a power spectrum make the primordial signal easier to identify and distinguish from nonlinear physics. As a result, we might anticipate that introducing $\nu > 0$ will strengthen our constraints for all frequencies, but the actual situation is more nuanced as can be seen in Fig. 3.6 when comparing the solid to the dotted line with the 95% limit on $|f_{\text{NL}}^{\alpha, \nu}|$, which is the absolute value of the power-law PNG amplitude considered in [193]. We additionally present their ratio in Fig. 3.7. Taken together, we can see distinct behavior at small, intermediate and large frequencies, and for positive and negative scaling exponents.

¹⁹While the inherent scatter of the inferred mean values around the truth in an observational measurement was also taken into account in the forecasts of [18] by additionally sampling the fiducial cosmology, we keep the fiducial values fixed to the origin in this work.

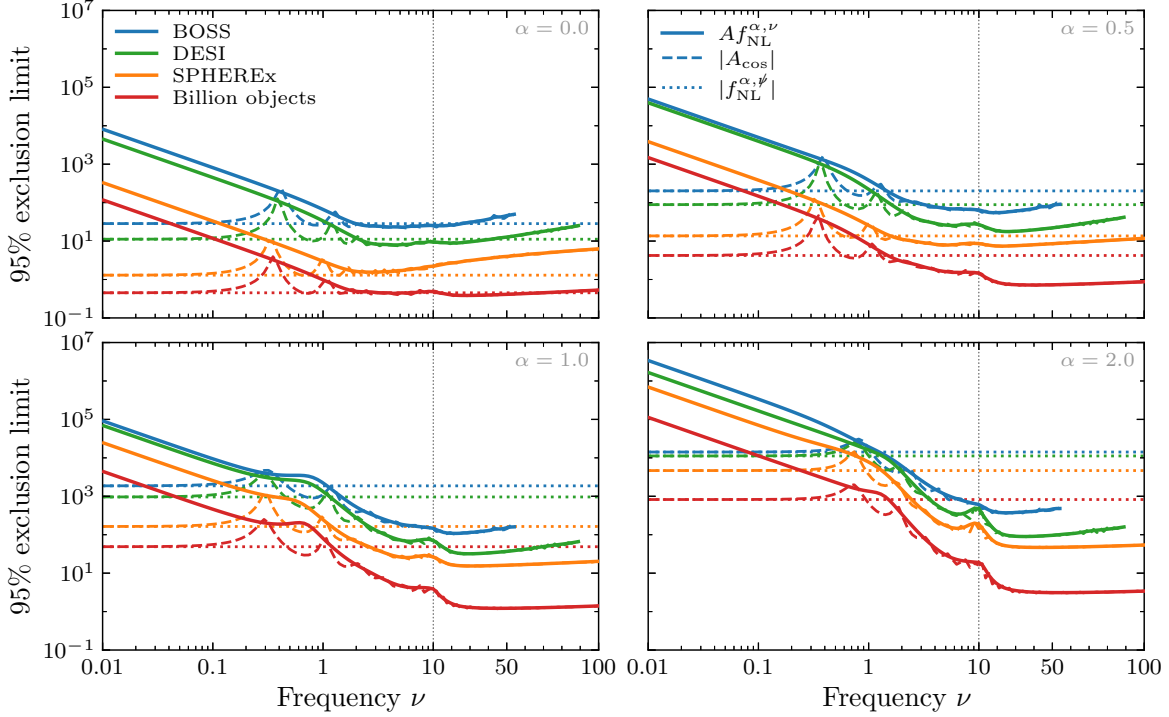


Figure 3.6. Forecasted constraints on the amplitude of primordial non-Gaussianity in terms of its 95% exclusion limit as a function of the oscillation frequency ν for four values of the scaling exponent α . We note the logarithmic-linear scaling of the frequency axis as separated by the vertical dotted line. We show the Fisher-based limits on the overall oscillatory amplitude $Af_{\text{NL}}^{\alpha,\nu}$ (solid), the oscillatory cosine amplitude $|A_{\text{cos}}|$ with fixed $A_{\text{sin}} = 0$ (dashed) and non-oscillatory power-law amplitude $|f_{\text{NL}}^{\alpha,\nu}|$ (dotted) for the current and future surveys BOSS (blue), DESI (green; Euclid has a very similar forecasted sensitivity), SPHEREx (orange) and the billion-object survey (red) described in the main text. The fiducial cosmology has Gaussian initial conditions, and we marginalized over the galaxy biases and Λ CDM parameters, with priors on the latter from a Planck-like CMB experiment. For small frequencies ν , the constraints on the overall oscillatory amplitude $Af_{\text{NL}}^{\alpha,\nu}$ are worse than those for the power-law case with $\nu = 0$ denoted by $|f_{\text{NL}}^{\alpha,\nu}|$ due to the implicit marginalization over the phase exhibiting the phase degeneracy, while we recover the $\nu = 0$ result when the phase is fixed to zero by setting $A_{\text{sin}} = 0$. On the other hand, the forecasted limits are much more similar for different values of α at large ν since the main constraining power comes from the imprinted oscillations. The frequency range available to the experiments is limited by the probed volume since it sets the minimum wavenumber and bandwidths which we conservatively estimate for each redshift bin (see Fig. B.1 in Appendix B.1 for a visualization). We also note that we compare the PNG amplitude $Af_{\text{NL}}^{\alpha,\nu}$ to the absolute values of A_{cos} and $f_{\text{NL}}^{\alpha,\nu}$ since $Af_{\text{NL}}^{\alpha,\nu}$ is a positive semi-definite quantity.

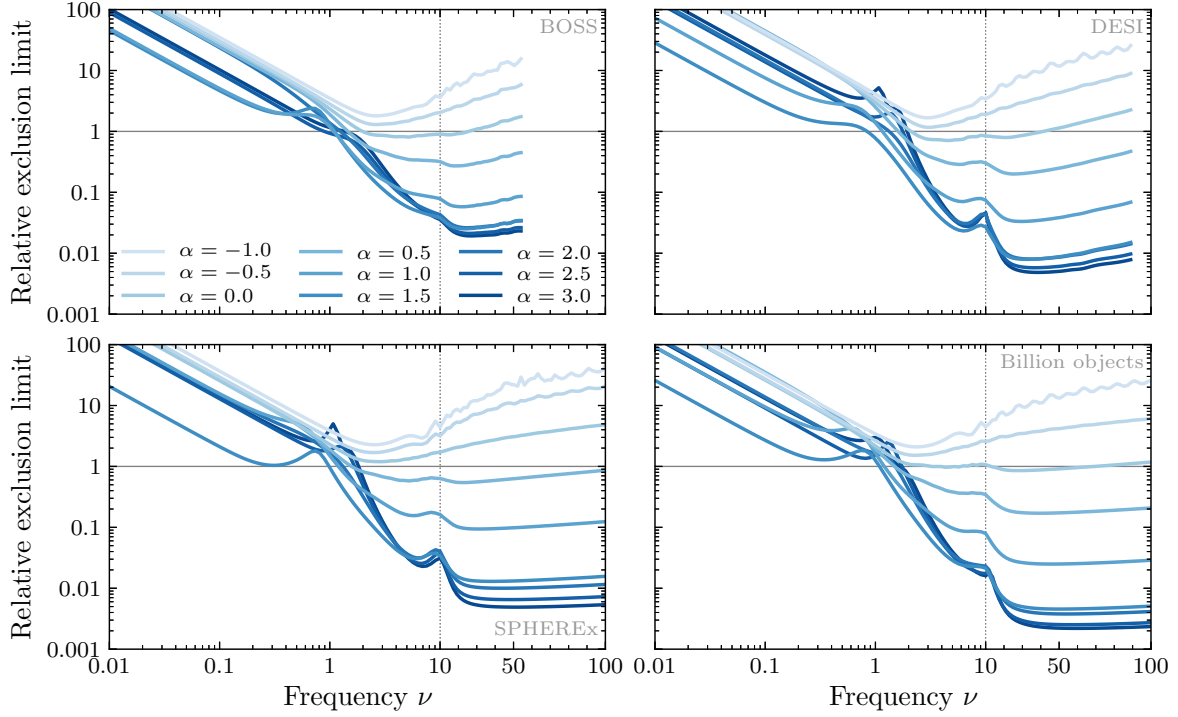


Figure 3.7. Ratio of the 95% exclusion limits for the oscillatory PNG amplitude $Af_{\text{NL}}^{\alpha,\nu}$ with $\nu \neq 0$ and the (absolute value of the) power-law PNG amplitude $|f_{\text{NL}}^{\alpha,\nu}|$ with $\nu = 0$ as a function of frequency ν for a range of scaling exponents α . We note the logarithmic-linear scaling of the frequency axis as separated by the vertical dotted line. The four panels show the relative improvement (or degradation) of the oscillatory signature for the four current and future surveys BOSS, DESI, SPHEREx and the billion-object survey. While the forecasted limits are degraded by the inclusion of the oscillations for small ν and $\alpha \lesssim 0$ due to the phase degeneracy and the signal peaking at the lowest wavenumbers, respectively, we observe a significant improvement in particular for large frequencies and scaling exponents since the oscillatory signal breaks the degeneracies with the broadband galaxy power spectrum.

For scaling exponents $\alpha \lesssim 0$, the overall constraining power on the PNG amplitude $f_{\text{NL}}^{\alpha, \nu}$ is most stringent in absolute terms, reminiscent of the light-field case of [193], which we have now extended to negative values of α . At the same time, there is effectively no improvement or even a severe degradation of the sensitivity by more than an order of magnitude even for large frequencies compared to the $\nu = 0$ case. The reason is that the PNG signal peaks at small wavenumbers $k \rightarrow 0$, leaving little room for our oscillations to be visible in the observed power spectra (see Fig. 3.5) and providing little benefit. On the other hand, due to the power-law term $S(k; M, \alpha)$ in (3.44), the signal moves to larger wavenumbers as the scaling exponent increases and peaks at k_{max} for the largest α under consideration. While this degrades the overall constraints on $f_{\text{NL}}^{\alpha, \nu}$, the imprinted oscillations can as a result dramatically improve the relative PNG sensitivity by up to two orders of magnitude for BOSS and even almost three orders of magnitude for the billion-object survey for $\nu \gtrsim 30$. So, as for primordial features [18], introducing oscillations allows the PNG signal to be measured at $k \sim k_{\text{NL}}$ while being robust to nonlinearities. In fact, as mentioned above, our treatment is conservative since we only consider linear wavenumbers, i.e. we limit the potential reach of the observations that would be possible following [18] in this regime.²⁰ Even with our limited range of wavenumbers, we however clearly see the improvement over $\nu = 0$ for $\alpha \gtrsim 0$ and $\nu \gtrsim 1$ precisely due to the fact that the oscillations break the degeneracies of the power-law term with biasing and short-distance physics more generally.

At small frequencies, $\nu \lesssim 1$, the forecasts weaken significantly, effectively independently of the experimental specifications. We can see this both when comparing the limits for $\nu \neq 0$ and $\nu = 0$ in Fig. 3.6 or in their ratio in Fig. 3.7. This is due to marginalizing over the phase. Concretely, without seeing multiple periods of the oscillation at a large signal-to-noise ratio, changing the phase allows to escape the more stringent bounds. Put differently, the amplitude

²⁰The range of considered wavenumbers could be significantly extended to all observed modes of a survey for the oscillatory imprint with the additional modeling as for primordial features in [18]. While we left this combined analysis to future work, we can already anticipate that it would further improve the sensitivity considerably in this regime of (α, ν) .

parameter $A_{\sin}$ becomes nearly unconstrained at $\nu \rightarrow 0$, which makes constraining the overall amplitude more difficult. The role of this degeneracy is clearly demonstrated in Fig. 3.8, in which the forecasted contours in the $(A_{\cos}, A_{\sin})$ plane are shown as a function of ν for $\alpha = 0$.²¹ We see that the degeneracy line rotates as we decrease the frequency. As $\nu \rightarrow 0$, the $A_{\cos}$ direction remains as constraining as the non-oscillatory scenario (see also the dashed line in Fig. 3.6 for the 95% exclusion limit on $|A_{\cos}|$ which is essentially the same as the limit on $f_{\text{NL}}^{\alpha, \nu}$ for $\nu \lesssim 0.1$), while constraints along the $A_{\sin}$ direction degrade rapidly. Since we also display the circle indicating the 95% limit on $A f_{\text{NL}}^{\alpha, \nu}$, we can also directly see how this degeneracy impacts the sensitivity to the overall PNG amplitude.

When we increase the frequency, we can observe in Fig. 3.8 that the contours are nearly circular for $\nu \gtrsim 2$ which indicates that this phase degeneracy is effectively broken at these larger frequencies. We note that the same rotation and vanishing of the degeneracy line with increasing ν is not only present for $\alpha = 0$ as shown, but all scaling exponents under consideration. This therefore also explains why we then start being able to see the improvements in the forecasted limits for $\alpha \gtrsim 0$ since the power-law term moves the signal to the largest wavenumbers in the analysis where the oscillations really help to distinguish between the primordial signature and the nonlinear galaxy power spectrum with Gaussian initial conditions. Comparing the limits on $A f_{\text{NL}}^{\alpha, \nu}$ and $|f_{\text{NL}}^{\alpha, \nu}|$ in Fig. 3.6 or their ratio in Fig. 3.7 at fixed α confirms our expectation that oscillations make it easier to measure this signal at large k .

Additional features of Figures 3.6 and 3.7 are the bumps around $\nu \approx 1$ and 10, with their significance depending on the scaling exponent. These bumps can be attributed to degeneracies with cosmological and bias parameters since the oscillatory signal may become degenerate with the broadband and/or the standard baryon acoustic oscillations. We further discuss this in Appendix B.1, in particular around Fig. B.4, which shows that these features depend on the marginalization over both the bias and Λ CDM parameters. We additionally note that we cut off

²¹Note that the light-blue two-dimensional 2σ contours enclose about 86.4% of the probability while the orange circles visualize the respective (one-dimensional) 95% limits.

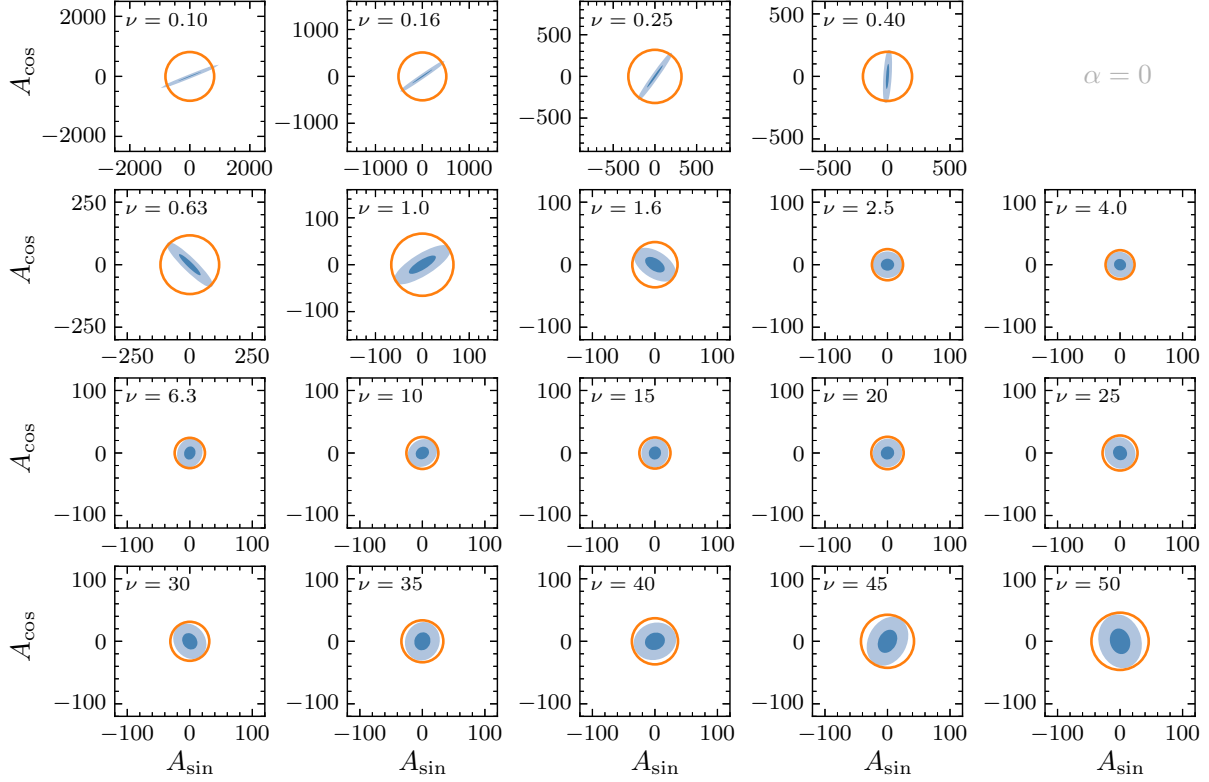


Figure 3.8. Forecasted posterior distributions of $(A_{\cos}, A_{\sin})$ for BOSS, a scaling exponent of $\alpha = 0$ and different frequencies ν . The standard two-dimensional 1σ and 2σ limits around the fiducial value of zero are included as the blue contours. The orange circle encloses 95% of the probability, with its radius being the corresponding limit on the overall oscillatory PNG amplitude $A f_{\text{NL}}^{\alpha, \nu}$. While the sine and cosine amplitudes are strongly degenerate for small ν , and the degeneracy line rotates as we increase the frequency, the degeneracy effectively vanishes for $\nu \gtrsim 2$ as evidenced by the (nearly) circular posterior contours. We note that the light-blue 2σ contours and the orange circle do not coincide even for the frequencies for which $A_{\cos}$ and $A_{\sin}$ are essentially independent since the former (two-dimensional 2σ) contours enclose approximately 86.4% of the probability and not 95% as in the case of the latter to display the (one-dimensional) limit on the PNG amplitude.

the frequency range for BOSS and DESI since the oscillations for frequencies exceeding the effective Nyquist frequency, which is set by the employed conservative bandwidths for each redshift bin,²² are aliased for all wavenumbers (see Fig. B.1 in Appendix B.1 for a visualization). This implies that the forecasts for BOSS and DESI should not be trusted for $\nu \gtrsim 55$ and 90, respectively, while the other surveys cover a larger cosmic volume with correspondingly larger Nyquist frequencies far beyond the frequencies considered here.

The potential reach of SPHEREx, which will have results within the next few years, is of particular interest in Figures 3.6 and 3.7. When $\alpha \approx 0$, the improvement from BOSS and DESI to SPHEREx is more than an order of magnitude and persists across the full range for frequencies. This improvement is largely compatible with the reach in $f_{\text{NL}}^{\text{loc}}$ since the higher-frequency results do not meaningfully exceed the $\nu = 0$ sensitivity. In contrast, for $\alpha = O(1)$, the relative improvement is the same within a factor of a few for small ν , but we see substantial improvements at higher frequencies relative to the $\nu = 0$ forecasts. Both results are compatible with the idea that SPHEREx has a larger advantage when the signal peaks at low k , while the benefit of large frequencies is pronounced for larger α when the signal is pushed to higher wavenumbers.

The relative importance of the smallest modes can also be seen through the dependence on the maximum wavenumber k_{max} as shown in Fig. 3.9. For all values of α , we see that for $\nu \gtrsim 10$, there is a significant improvement in the forecasted limits with $k_{\text{max}} = k_{\text{NL}} > k_{\text{halo}}$. For $\nu < 1$, there is a more pronounced improvement with larger k_{max} only for $\alpha = O(1)$. This reinforces the conclusion that (i) $\nu > 1$ gains sensitivity from small-scale modes since the oscillations are not degenerate with biasing and (ii) $\alpha \lesssim 0.5$ gets more constraining power from $k \ll k_{\text{NL}}$ and is therefore less sensitive to the choice of k_{max} . As confirmed in additional forecasts based on the billion-object survey (as in [193]), this same pattern is important for understanding the

²²The bandwidths employed in our forecasts are inferred from a spherical volume separately for each redshift bin which results in overly cautious estimates. This also means that the bandwidths for BOSS are larger in our forecasts than in the data analysis presented in Section 3.5. Compared to [18], the smaller maximum wavenumber further limits the effective reach in the logarithmic frequency ν .

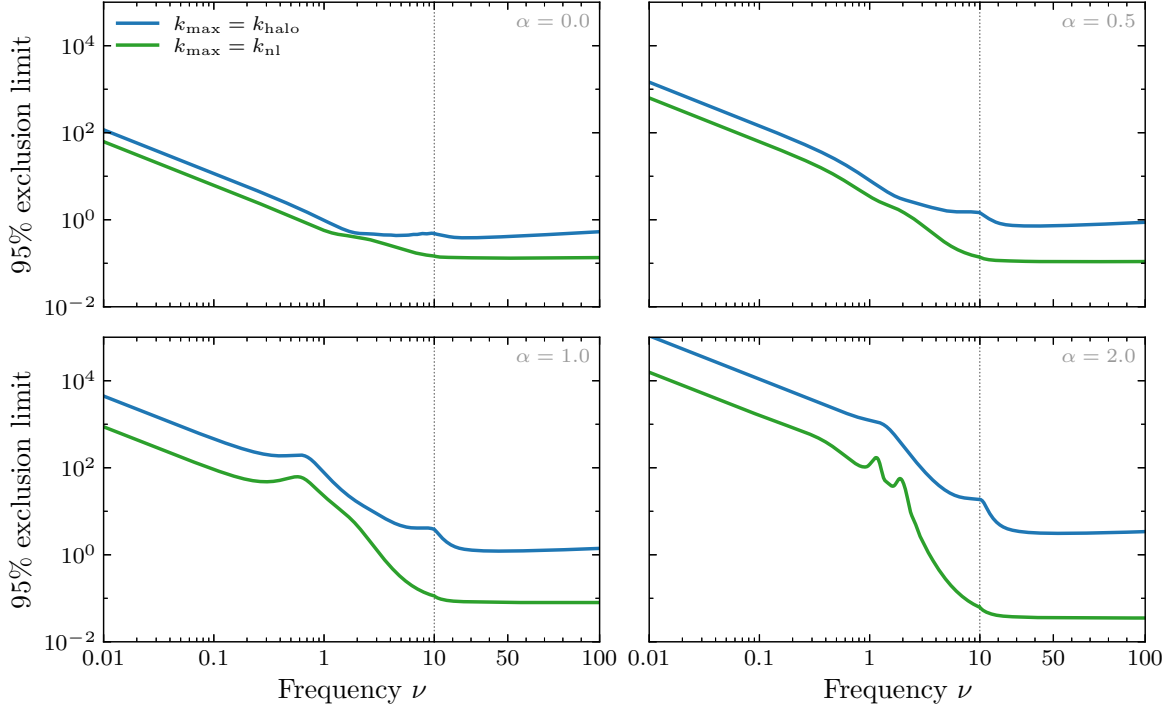


Figure 3.9. Dependence of the 95% exclusion limit of the PNG amplitude $f_{\text{NL}}^{\alpha, \nu}$ on the maximum wavenumber k_{max} of the billion-object survey for different values of the scaling exponent α . We compare the two cases for which we set the maximum wavenumber to correspond to the Lagrangian scale of the minimum halo mass, $k_{\text{max}} = k_{\text{halo}}$, and to the nonlinear scale, $k_{\text{max}} = k_{\text{NL}}(z)$. The latter is in particular much more optimistic at higher redshifts given our modeling of the (Gaussian) galaxy power spectrum. While the impact on the forecasted constraints for small frequencies is limited, as expected, we see decisive improvements when increasing k_{max} to k_{NL} at high frequencies, especially for large α , since many more oscillatory periods are being captured by the survey.

relative importance of the survey volume (or the minimum wavenumber $k_{\min}$) and the number density (or $k_{\max}$) when designing future surveys for these kinds of power-law and/or oscillatory signals.

Of course, larger and deeper surveys have the potential to find ever weaker PNG signals of this kind. We however see in Fig. 3.7 that the shapes of the curves illustrating the relative limits as a function of ν are very similar across the four displayed surveys. As a result, the impact of survey design on the $\nu = 0$ forecasts, which were discussed at length in [193], will be effectively the same for all values of ν . In this regard, the large improvement in going from DESI to SPHEREx can again be attributed to the higher biases and the slightly larger number densities of the latter. Furthermore, we again expect a multi-tracer analysis from LSST to offer improved sensitivity over SPHEREx. Having said that, it would be worthwhile to understand how these non-Gaussian oscillations are affected by systematics. While the raw sensitivity of SPHEREx and LSST offer the potential for large improvements over spectroscopic surveys like DESI and Euclid, the scale-dependent bias from local non-Gaussianity can be particularly sensitive to large-angle effects and (catastrophic) photometric redshift errors among other potential systematic biases (see e.g. [121, 169, 171, 173, 174]). An oscillatory signal can potentially evade this potential sensitivity to systematic errors, making these kinds of improvements more robust than for the scale-dependent bias from light fields parametrized by $f_{\text{NL}}^{\alpha, \nu}$.

Ultimately, inferring constraints on the PNG signal in terms of $f_{\text{NL}}^{\alpha, \nu}$ for a given α and ν is a determination of the strength of the coupling of σ and ζ for a fixed quadratic action for σ (i.e. its masses and mixing). For $\alpha \approx 1$ and $\nu \approx 1$, the improvement in the limits on $f_{\text{NL}}^{\alpha, \nu}$ relative to $\nu = 0$ from Fig. 3.7 is substantially larger than the effect due to the LSS suppression factor $|\Sigma_M^2|$ in Fig. 3.4 (as well as the primordial factor $e^{-\pi\nu}$). This is the regime where our new analysis is expected to be more sensitive to the coupling to additional fields than the models with no oscillations. This might appear to be a small region of parameters space, but we note that there is good reason for dimensionful parameters to be set by the Hubble scale H [13]. Moreover,

as we discussed in §3.3.3, there are phenomenological ways to boost the primordial amplitude to further overcome the suppression factors over a wider frequency range. Either way, it is evident that the improved PNG sensitivity is a consequence of model building and analysis, which highlights the discovery potential that is driven by theoretical work in this area.

To summarize, galaxy surveys are more sensitive to oscillatory signals with $\nu \gtrsim 1$ when $\alpha = O(1)$. In these situations, the signal-to-noise ratio is dominated by large- k modes and is sufficiently distinct from nonlinear corrections to be robust to marginalization over bias parameters. For scaling exponents α near or smaller than zero, the signal is peaked at smaller wavenumbers which results in less improvement in sensitivity relative to the non-oscillatory scenario, $\nu = 0$. These models are not degenerate with more conventional PNG shapes, which means that an analysis of BOSS or DESI data could find a non-zero signal for non-zero α and ν while being consistent with previous analyses.²³

3.5 Constraints from BOSS Galaxy Clustering

Having established the expected sensitivities of current and future LSS surveys to our broad class of oscillatory non-Gaussianity in the previous section, we now apply its phenomenological template to observations and derive constraints from observed galaxy-clustering data. We analyze BOSS DR12 data based on the galaxy power spectrum. The methodology, which we summarize in §3.5.1, closely follows that of recent analyses of standard PNG templates [65, 66, 124, 266] and our previous analysis of non-oscillatory models [193], with extensions to address the presence of the frequency-dependent oscillations. Our inferred limits on the PNG amplitude $A f_{\text{NL}}^{\alpha, \nu}$ are presented in §3.5.2 and exhibit the same behavior as seen in the forecasts. This includes confirming the significance of an oscillatory component in reducing the sensitivity to nonlinear astrophysical effects even for frequencies $\nu = O(1)$. Since we find no evidence for a non-zero oscillatory PNG signal, we place upper limits on the explored

²³The results must also be consistent with the existing constraints on primordial features from BOSS [18, 233–235].

(α, ν) parameter space. The results of this section suggest that this type of analysis should be straightforward to apply to current and future data, for example from DESI (e.g. as an extension of [266]) and SPHEREx.

3.5.1 Analysis Pipeline for BOSS DR12 Data

To extract constraints on oscillatory PNG, we analyze the BOSS DR12 [77] galaxy power spectrum following the same methodology used in [193], which was based on [66, 124] (see also [65]) with a few different (generally more conservative) choices, after extending it to account for the scale-dependent oscillatory modulation predicted by the class of models studied in this work. In particular, we modify the theoretical prediction for the galaxy power spectrum to include the scale- and frequency-dependent bias induced by primordial oscillations in the bispectrum. As discussed in more depth after (3.46), we use the amplitude instead of the phase parametrization since the former results in nearly Gaussian posterior distributions for the oscillatory parameters while the phase is relatively poorly constrained in the latter in the absence of a signal. Compared to [193], we therefore extend our theoretical model of the galaxy power spectrum $P_g(k)$ to include the oscillatory amplitudes $A_{\cos}$ and $A_{\sin}$ as follows:²⁴

$$\begin{aligned} b_{\text{NG}}^{\Delta}(k, z) &= 3f_{\text{NL}}^{\Delta} \frac{b_{\phi}(z)}{k^2 \mathcal{T}(k, z)} (kR_M)^{\Delta} \\ \rightarrow b_{\text{NG}}^{\alpha, \nu}(k, z) &= 3 \frac{b_{\phi}(z)}{k^2 \mathcal{T}(k, z)} (kR_M)^{\alpha} \{A_{\cos} \cos[\nu \log(kR_M)] + A_{\sin} \sin[\nu \log(kR_M)]\}, \end{aligned} \quad (3.51)$$

where we took $R_M \approx 2.66 h^{-1} \text{ Mpc}$ corresponding to the Lagrangian radius of a halo with a mass of $10^{13} M_{\odot}$, as in our Fisher forecasts presented in §3.4.3. Let us in particular reiterate that we effectively absorbed the suppression factor $|\Sigma_M^2(\alpha, \nu)|$ of (3.44) into the PNG amplitudes.

The dataset, the observables and covariances, and the nonlinear and bias modeling are identical to the analysis presented in [193]. In short, we use the BOSS DR12 dataset [77] with

²⁴We also refer to (3.49) and remember that we absorbed the overall PNG amplitude $f_{\text{NL}}^{\alpha, \nu}$ into the amplitudes $A_{\cos}$ and $A_{\sin}$ for notational convenience while ultimately still reporting the results of our analysis in terms of $f_{\text{NL}}^{\alpha, \nu}$.

about 1.2 million galaxies which are split into two redshift bins, $z \in [0.2, 0.5]$ and $[0.5, 0.75]$, and two galactic caps. The monopole, the quadrupole and hexadecapole of the galaxy power spectrum $P_g^{(\ell)}(k)$ are estimated using a quadratic window-function-free estimator [125] while the employed covariance matrices are computed from the MultiDark-Patchy mocks [126]. We limit the range of wavenumbers to $k \in 0.01 h \text{Mpc}^{-1} \text{ to } 0.13 h \text{Mpc}^{-1}$ and $0.01 h \text{Mpc}^{-1} \text{ to } 0.16 h \text{Mpc}^{-1}$ for the low- and high-redshift bin, respectively, with bandpowers $\Delta k = 0.005 h \text{Mpc}^{-1}$. The theoretical power spectra account for redshift-space distortions, nonlinear evolution and the Alcock–Paczynski effect, include parameters from the bias expansion and EFTofLSS description following [129, 130] (cf. [25, 131–134] and [27, 135] for recent reviews) and are calculated using a modified version of CLASS-PT [127] which in turn is based on CLASS [128].

The parameter inference is performed via Markov chain Monte Carlo using the Metropolis-Hastings algorithm implemented in MontePython [136, 137], with all chains converging with a Gelman-Rubin criterion of $R - 1 < 0.01$ (usually much smaller) for each varied parameter. While the cosmological parameters are fixed to the Planck 2018 best-fit ΛCDM values (with the sum of neutrino masses $\sum m_\nu = 0.06 \text{eV}$) [1],²⁵ the bias and EFT parameters are sampled independently for each redshift bin and galactic cap as in [123, 124, 140, 193]. On the other hand, the oscillatory non-Gaussian parameters A_{cos} and A_{sin} are shared across all subsets, with the (real-valued) scaling exponent $\alpha = \text{Re}\Delta$ and frequency $\nu = \text{Im}\Delta$ being fixed for individual runs and scanned over a linear and logarithmic-linear grid with nine and 19 values, respectively: $\alpha \in [-1, 3]$ and $\nu \in [0.1, 10] \cup [10, 50]$. As discussed above, these ranges are motivated by theoretical and observational considerations, respectively: we interpolate between the benchmark scaling exponents α for the tachyonic collider ($\alpha = -1$), local PNG ($\alpha = 0$), the conventional cosmological collider ($\alpha = 3/2$), equilateral and orthogonal PNG ($\alpha = 2$) and originating from inflationary loops or composite operators ($\alpha = 3$); we take $\nu \geq 0.1$ since we saw in our forecasts

²⁵We saw in the forecasts (see Fig. B.3 in Appendix B.1) that fixing the five ΛCDM parameters has essentially the same effect on the constraints on $f_{\text{NL}}^{\alpha, \nu}$ as imposing a ΛCDM prior consistent with Planck 2018. For BOSS, the main difference of fixing these parameters is a slight overestimation of the constraining power around $\nu \approx 10$ which is broadly negligible on our logarithmic scale for $f_{\text{NL}}^{\alpha, \nu}$ and given the main focus of this work. We therefore effectively infer the limits on the PNG amplitude from the BOSS data while imposing a Planck prior on the ΛCDM parameters.

that the constraints further weaken for smaller frequencies given the phase degeneracy and $\nu \leq 50$ to approximately remain within the effective logarithmic Nyquist frequency of the BOSS DR12 volume (cf. [18] and Fig. B.1 in Appendix B.1). We impose flat priors on the non-Gaussian and the linear bias parameters, and Gaussian priors on higher-order bias, the non-Gaussian bias and nuisance parameters. We refer to [193] for further information and details on the data and choices underlying our analysis.

To infer the value of the PNG amplitude $f_{\text{NL}}^{\alpha,\nu}$, rather than the product $b_\phi f_{\text{NL}}^{\alpha,\nu}$, we marginalize over b_ϕ with a Gaussian prior separately in each redshift bin. This prior has a mean value that is set by the universality relation, $b_\phi = 2\delta_c(b_1 - 1)$, and a large standard deviation of five times this mean (see [193] for additional details). This procedure alleviates degeneracies between $f_{\text{NL}}^{\alpha,\nu}$ and b_1 which can be severe if we assumed that the universality relation is exact, especially for $\alpha \sim 1$. Although the constraint on $f_{\text{NL}}^{\alpha,\nu}$ is sensitive to the prior on b_ϕ , the degeneracy is further reduced by the fact that b_ϕ varies in each redshift bin while there is only a single primordial amplitude $f_{\text{NL}}^{\alpha,\nu}$ (or rather A_{cos} and A_{sin}) for the entire survey.

In order to constrain the oscillatory PNG amplitude while marginalizing over the phase, we follow the same method introduced in [18] (and also applied to our forecasts in §3.4.3) to infer the limit on the overall (phase-independent) oscillation amplitude from an analysis based on the amplitude parametrization. This means that we analyze the marginalized posteriors of the oscillatory amplitudes, A_{cos} and A_{sin} , in the two-dimensional plane defined by these parameters. The 95% confidence limit on the total amplitude $A f_{\text{NL}}^{\alpha,\nu} = \sqrt{A_{\text{cos}}^2 + A_{\text{sin}}^2}$, cf. (3.50), is defined as the radius of a circle around the origin that encloses 95% of the posterior probability or, equivalently, 95% of the Monte Carlo samples which implicitly marginalizes over the phase of the non-Gaussian oscillations. This procedure therefore accounts for the fact that the phase is not independently constrained and that the maximum posterior point may be offset from the origin due to noise. While the posteriors are pixelated in the $(A_{\text{cos}}, A_{\text{sin}})$ plane, we explicitly verified that pixelation artifacts are negligible. We note that A_{cos} and A_{sin} are

effectively uncorrelated for frequencies $\nu \gtrsim 1$, but exhibit correlations especially for the smallest frequencies under consideration (see Fig. 3.8 and the top panels of Fig. 3.10 for some example illustrations, and Appendix B.2 for additional details). While our procedure is generally robust and conservative in estimating the PNG amplitude $Af_{\text{NL}}^{\alpha,\nu}$, this correlation implies that the inferred limits may not always reflect the actual constraining power of the data. For full details and validation of this methodology, we refer to Appendix C of [18].

The bounds on A_{cos} and A_{sin} for $\alpha = 0$, along with the derived 95% exclusion limits of $Af_{\text{NL}}^{\alpha,\nu}$, are shown in the bottom panel of Fig. 3.10. Comparing the top panels with the associated frequency illustrates the relationship between the samples in the Monte Carlo Markov chains and the inferred bounds. The mean values of A_{cos} and A_{sin} are consistent with zero across all of the frequencies (and scaling exponents) that we sampled. It is important to keep in mind that the analysis at each individual frequency is correlated with other frequencies, especially for $\alpha = 0$, where the information is at low wavenumbers k . As a result, we do not have an expectation to see several 1σ fluctuations in these 19 frequencies. Similarly, we had already seen and utilized the broad correlations in the scaling exponent in [193].

3.5.2 Bounding Non-Gaussian Oscillations

We can straightforwardly infer the posterior distributions for the parameter $Af_{\text{NL}}^{\alpha,\nu}$ from the Markov chains of the BOSS DR12 data following the procedures laid out above. For each pair of (α, ν) in our grid of values, we calculate the posterior while marginalizing over the additional bias and EFT parameters. These results are all consistent with $Af_{\text{NL}}^{\alpha,\nu} = 0$ and, therefore, do not show any evidence for this kind of underlying non-Gaussian signal. We display the 95% exclusion limits on the non-Gaussian oscillatory amplitude $Af_{\text{NL}}^{\alpha,\nu}$ as a function of the frequency ν for nine values of the scaling exponent $\alpha \in [-1, 3]$ in Fig. 3.11.

We can clearly observe the general trends of our inferred exclusion limits as a function of both α and ν . We infer the tightest bounds for $\alpha = -1$ which is due to the sharp cubic scaling behavior on large scales. As we increase the scaling exponent α , the limits continuously

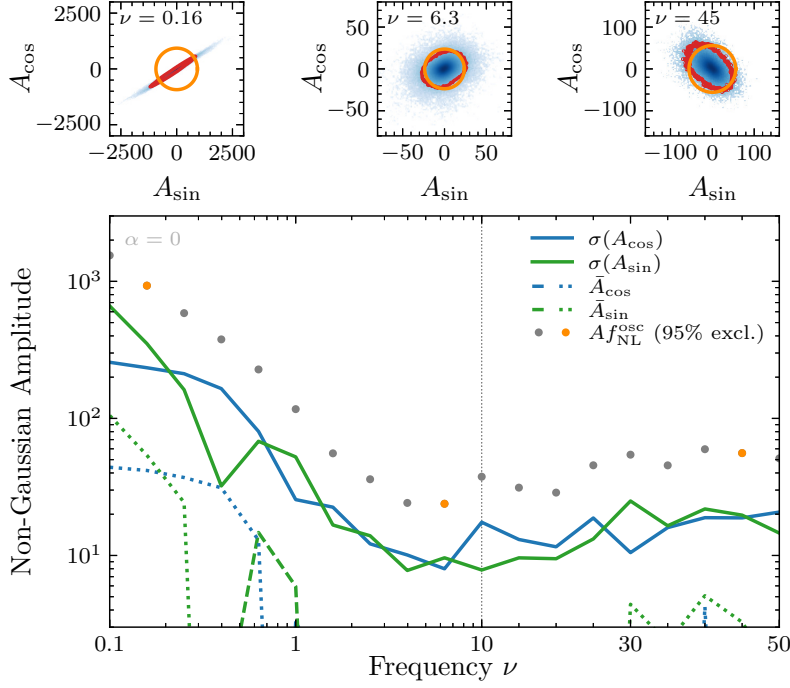


Figure 3.10. Constraints on the non-Gaussian amplitude(s) inferred from BOSS DR12 for a scaling exponent of $\alpha = 0$ as a function of frequency ν . We note that the frequency axis is displayed on a logarithmic-linear scale. This figure illustrates the relationship between the MCMC samples and the inferred mean values, standard deviations and 95% upper limits at each sampled frequency. *Top:* The three panels contain the pixelated posterior distributions in the marginalized $(A_{\cos}, A_{\sin})$ plane based on the MCMC samples in blue (light/dark blue: low/high density) for the three frequencies $\nu = 0.16, 6.3$ and 45 . We see the strong and weak correlation of these two amplitudes at small and intermediate/large frequencies, respectively, which is also illustrated by the red contours which enclose 95% of the samples. This is also the case for the orange circles which display our main statistic, the phase-marginalized 95% (exclusion) limits on the overall amplitude $Af_{\text{NL}}^{\alpha, \nu} = \sqrt{A_{\cos}^2 + A_{\sin}^2}$, which are derived following [18] as summarized in the main text. We additionally refer to Appendix B.2 for further details and these $(A_{\cos}, A_{\sin})$ panels for all sampled frequencies ν . *Bottom:* The standard deviations, positive and negative mean values of $A_{\cos}$ (blue) and $A_{\sin}$ (green) are shown as solid, dashed and dotted lines. The gray dots are the mentioned 95% exclusion limits on the PNG amplitude $Af_{\text{NL}}^{\alpha, \nu}$ for the 19 sampled values of the frequency ν , with the orange points corresponding to the three panels at the top. We can clearly observe that no non-zero PNG amplitude is detected and how the degeneracy line between $A_{\cos}$ and $A_{\sin}$ rotates for $\nu \lesssim 1$ while these amplitudes become nearly uncorrelated for larger frequencies.

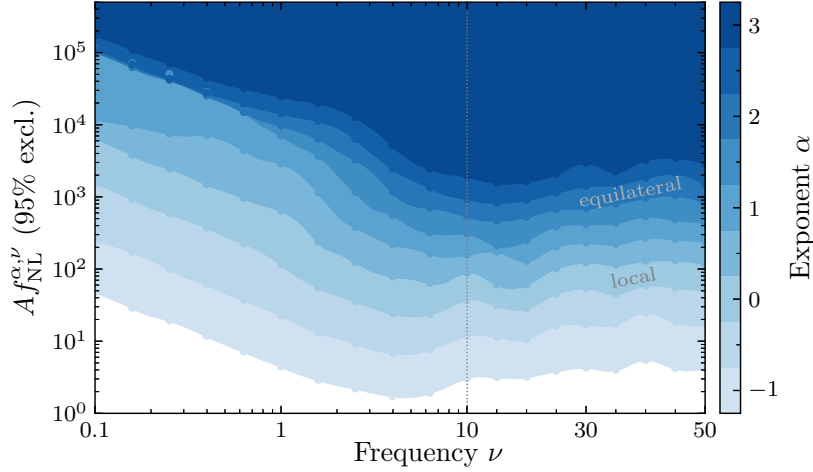


Figure 3.11. Exclusion limits at 95% c.l. on the non-Gaussian oscillatory amplitude $A f_{\text{NL}}^{\alpha, \nu}$ inferred from the BOSS DR12 dataset as a function of the oscillatory frequency ν of the non-Gaussian signature. The excluded parameter space is indicated by the shaded regions for different scaling exponent α . Note that the horizontal frequency axis is on a logarithmic-linear scale, while the vertical axis for the amplitude is logarithmic. The labels ‘local’ and ‘equilateral’ denote the parameter space excluded for $\alpha = 0$ and 2, respectively, which are the scaling exponents of these standard PNG types in the $\nu \rightarrow 0$ limit. These exclusion limits are inferred from 171 separate Markov chain Monte Carlo analyses at fixed (α, ν) using the method introduced in [18], which implicitly marginalizes over the oscillatory phase. On this grid, the tightest 95% bounds with $A f_{\text{NL}}^{\alpha, \nu} < 1.7$ are inferred for $(\alpha, \nu) = (-1, 4)$, while the weakest 95% limits with $A f_{\text{NL}}^{\alpha, \nu} < 1.7 \times 10^5$ are inferred for $(\alpha, \nu) = (3.0, 0.1)$. We observe that the bounds get weaker for both lower frequencies ν and larger exponents α .

degrade, being more than three orders of magnitude weaker for $\alpha = 3$. On the other hand, these limits improve by about one to almost two orders of magnitude from $\nu = 0.1$ to $\nu \sim 10$ before remaining roughly at the same level for $\nu \gtrsim 10$. As discussed in §3.4.3 and §3.5.1, the apparent improvement in constraints from $\nu \ll 1$ to $\nu = O(1)$ is primarily a consequence of marginalizing over the phase of our two-parameter model of the non-Gaussian oscillations. Due to this marginalization, the constraints at small ν are much weaker than the constraint for $\nu = 0$ derived in [193]. In this sense, the data remains very capable of detecting a signal at low frequency, but the overall sensitivity of the survey is not well captured by our statistic of the exclusion limit of $Af_{\text{NL}}^{\alpha,\nu}$ when $\nu \ll 1$ since the sine component becomes poorly constrained.

To better understand the functional dependence of these limits on the frequency ν , we also included the marginalized constraints on the amplitudes $A_{\cos}$ and $A_{\sin}$ for $\alpha = 0$ in Fig. 3.10. We clearly see that the two amplitudes are highly degenerate for small frequencies $\nu \lesssim 1$, with the degeneracy line rotating in the amplitude plane as a function of ν . This additionally implies that our procedure to infer a conservative exclusion limit on the overall amplitude A leads to weaker bounds in this range of parameters. For larger frequencies, the standard deviations of the two amplitudes are very similar and the procedure results in limits that more directly capture the underlying constraining power of the survey. We note that this behavior shown for $\alpha = 0$ is generally similar for other exponents α as can be assumed from Fig. 3.11 (and Fig. B.7 included in Appendix B.2). We also display the two-dimensional histograms for all sampled frequencies at a fixed scaling exponent of $\alpha = 0$ in Fig. B.6 in Appendix B.2.

We can additionally compare these new results with the respective limits for the non-oscillatory forms of primordial non-Gaussianity studied in [193]. We show the relative exclusion limits in Fig. 3.12. To be precise, we present the ratio of the 95% exclusion limits on the oscillatory PNG amplitude $Af_{\text{NL}}^{\alpha,\nu}$ displayed in Fig. 3.11 (and Fig. B.7) to the bound at 95% confidence level (c.l.) on (the absolute value of) the non-oscillatory, simple power-law PNG amplitude $|f_{\text{NL}}^{\alpha,\psi}|$ that we had studied in detail in [193].²⁶ We observe that the oscillatory bounds are weaker

²⁶While we analyzed the BOSS DR12 dataset for $\alpha \in [0, 2]$ in [193], we performed here an equivalent reanalysis

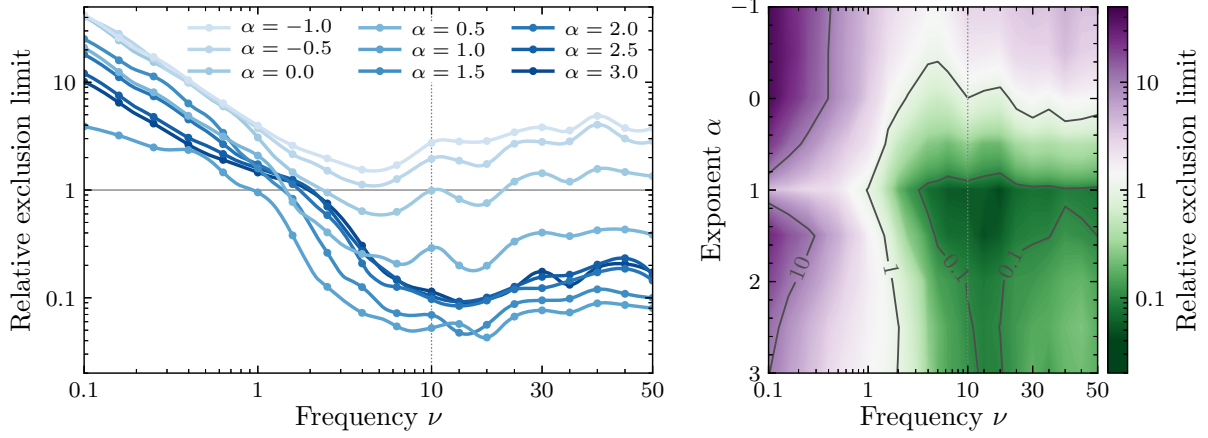


Figure 3.12. Exclusion limits at 95% c.l. on the amplitude of the oscillatory primordial non-Gaussianity, $Af_{\text{NL}}^{\alpha,\nu}$, from BOSS DR12 relative to the 95% limit on the absolute value of the amplitude of the non-oscillatory, simple power-law non-Gaussianity, $|f_{\text{NL}}^{\alpha,\psi}|$. The relative limits are shown as a function of the frequency ν for the nine scaling exponents α (*left*), and as a function of both α and ν (*right*). The axes for the exclusion limit, frequency and exponent are on a logarithmic, logarithmic-linear and linear scale, respectively. The oscillatory limits are better than those for the power-law signal for $\alpha \gtrsim 0$ and $\nu \gtrsim 1$, while they are worse otherwise. The improvement reaches factors of five to 23 for $\alpha \geq 0.5$. The relative degradation at small frequencies originates from the phase marginalization and the related strong degeneracy between A_{cos} and A_{sin} as discussed in the main text. For $\alpha \lesssim 0$, the highly oscillatory limits do not improve over the non-oscillatory constraints since the non-Gaussian signal peaks at the smallest survey wavenumbers and may be aliased. On the other hand, the power of the oscillatory imprint to distinguish the PNG signal for $\alpha > 0$ from the smooth galaxy power spectrum is clearly evident.

than the non-oscillatory bounds for $\nu \lesssim 1 - 2$ depending on the scaling exponent α and further weakening towards smaller frequencies. In fact, even for larger frequencies, the oscillatory limits are generally weaker for $\alpha < 0$ and the ratio is about unity for $\alpha = 0$. On the other hand, the oscillatory amplitude is constrained better by a factor of up to five to 23 for $\alpha \geq 0.5$ and frequencies $\nu \gtrsim 10$ on our grid of parameter values.

We can understand the origin of the degraded constraints at small frequencies to come from the degeneracies of the oscillatory signal with the smooth broadband spectrum and of the amplitudes A_X , $X = \sin, \cos$, as discussed above. This also results in a relative limit that is not unity in the limit of very small frequencies ($\nu \rightarrow 0$) as one might have initially expected. Relatedly, we do not observe an improvement in the inferred limits compared to the non-oscillatory PNG results for $\alpha \lesssim 0$ since the signal is dominated by the largest scales of the survey where the non-Gaussian signal is generally not degenerate with the broadband power spectrum. On the other hand, we see a significant improvement for positive exponents α and larger frequencies since the constraining power is coming from smaller and smaller scales where the additional oscillations break the larger and larger degeneracy of the smooth power-law non-Gaussian signal with the Gaussian broadband power spectrum.

Comparing these results inferred from the BOSS data to our Fisher forecasts described in §3.4.3, we can clearly see the qualitative similarities as a function of both scaling exponent α and oscillation frequency ν . While the precise quantitative limits somewhat differ between the forecasts and data analysis, these may be due to the differences in the underlying modeling and specifications as in [193], especially regarding the specific scale cuts and biases (with larger α being particularly sensitive). We in particular note that the bumps for $\nu \approx 1 - 2$ and $\alpha \gtrsim 0$ are as present in the forecasts as the seemingly unexpected behavior for smaller frequencies and $\alpha \gtrsim 1$ in Fig. 3.12. At the same time, the forecasted and inferred limits are also similar for larger frequencies including their roughly logarithmically equidistant values. We additionally

for $\alpha \in [-1, 3]$ (see Appendix B.2).

observe the same degeneracies and behavior in the two-dimensional $(A_{\cos}, A_{\sin})$ contours. We therefore conclude that our 95% exclusion limits on the oscillatory PNG amplitude $A f_{\text{NL}}^{\alpha, \nu}$ (and other results) inferred from the BOSS DR12 data in this section are broadly consistent with our BOSS forecasts of §3.4.3, with the agreement between the two providing a mutual validation of the forecasting and analysis pipelines.

To summarize, we have inferred the first constraints from data on the class of oscillatory primordial non-Gaussianity described by the scale-dependent bias (3.42) in the galaxy power spectrum for the wide ranges of $\alpha \in [-1, 3]$ and $\nu \in [0, 50]$. The power law governed by the scaling exponent α and the oscillations with frequency ν decisively change the sensitivity of BOSS to the PNG signal compared to the non-oscillatory or standard shapes. We explored this entire and previously unexplored parameter space and explained the origin of the constraining power in the different regimes. This will therefore make it straightforward to augment analyses of DESI, Euclid, SPHEREx and other future datasets targeting standard local and equilateral non-Gaussianity (such as the recent [266] for DESI) to also probe this promising class of oscillatory primordial non-Gaussianity.

3.6 Conclusions and Outlook

Primordial non-Gaussianity offers a unique opportunity to search for new particles and interactions during inflation, and has the potential to probe energy scales many orders of magnitude higher than any terrestrial experiment. Accessing these signatures of ultra-high-energy physics in cosmological data typically requires new templates and dedicated search strategies since the relevant signals may be hidden in more conventional analyses. Given the unparalleled reach of current and near-term surveys, it is timely and important to thoroughly explore the space of models and theoretically motivated signals to the greatest extent possible.

In this paper, we studied a new theoretically motivated class of models, which was first introduced in [24] and exhibits non-trivial behavior in the squeezed limit of the primordial

bispectrum, and their observational imprints in galaxy surveys. These models include a regime with both oscillatory and power-law behavior which gives rise to a hybrid signal between the traditional light- and heavy-field signals that have previously been explored.

The main phenomenological implication of this primordial physics for observations of the large-scale structure of the Universe is that the scale-dependent bias of the galaxy spectra, especially the power spectrum, is modified to include a non-trivial power law and logarithmic oscillations with a wide range of both scaling exponents and frequencies, respectively. We showed that this oscillatory component results in comparatively easy searches in observations and a more straightforward separation of primordial physics from late-time astrophysics. We explicitly demonstrated this in forecasts for current and future galaxy surveys, and in an analysis of the BOSS DR12 dataset. In this way, we also highlighted the significant discovery potential in this extended parameter space and, since our results are consistent with no such primordial non-Gaussianity, inferred the first constraints on these phenomenological non-Gaussian signals.

At the same time, it is also this oscillatory nature that typically suppresses the signal beyond the level that intrinsically arises from the inflationary mechanism itself when the primordial bispectrum is translated to the scale-dependent bias in late-time observables. The reason is that this transfer of power involves an average over small-scale (oscillatory) fluctuations. This suppression, which we absorbed into the inferred PNG amplitude, becomes increasingly significant at high frequencies. In turn, this suggests that a bispectrum or trispectrum analysis may provide more powerful constraints on the underlying microphysics of these models at least in parts of their parameter space. In fact, the first cosmological analyses for the conventional cosmological collider (i.e. with $\alpha = 3/2$) were only recently performed in [194–198]. While these measurements offered the first direct constraints on heavy inflationary fields, they are currently much less sensitive than indirect constraints from the equilateral bispectrum [16]. This therefore also means that there is still much space to explore not only phenomenologically in extensions of the cosmological collider, but also observationally in pipelines extending our

analysis to all observed wavenumbers and involving higher-point statistics in both the cosmic microwave background and the large-scale structure.²⁷

More broadly, new signals continue to appear that shift our perspective on what is both allowed and observable, even as techniques ranging from the effective field theory of inflation [279] to the cosmological bootstrap [280] have simplified and expanded the exploration of theory space. The signals described in this paper are one such example, but also only represent a small step away from existing searches. In fact, it is important to recognize that the signatures that may be easiest to detect in observations are not necessarily those that may be considered the most generic. This in particular includes the realizations that signals are not limited to low-point statistics [93] and can manifest themselves directly as features in maps of the primordial fluctuations [161]. A common aspect of these models, including the ones studied in this paper, is that their novel signatures can make it easier to isolate in the data than many of the (smooth) conventional shapes of primordial non-Gaussianity. In this sense, there is real discovery potential available from theoretical exploration alone, and the forthcoming wave of cosmological data provides a timely opportunity to convert this potential into new signature-driven insights into the primordial universe.

Acknowledgments

We are grateful to Daniel Baumann, Priyesh Chakraborty, Xingang Chen, Tim Cohen, Kshitij Gupta, Yiwen Huang, Oksana Iarygina, Alec Ridgway, Chia-Hsien Shen, Raman Sundrum, Xi Tong and Zhong-Zhi Xianyu for helpful discussions. D. G. and J. H. acknowledge support from the US Department of Energy under Grant DE-SC0009919. B. W. was supported by the Swedish Research Council under Contract No. 638-2013-8993. Nordita was supported in part by NordForsk. B. W. is grateful to INFN Ferrara, Scuola Normale Superiore di Pisa, University of California San Diego and University of Illinois Urbana-Champaign for their hospitality. Part

²⁷Our analysis can be extended to all observed wavenumbers for the oscillatory part of the scale-dependent bias similar to [18]. Apart from higher-point functions of the primary cosmic microwave background and maps of galaxy clustering, which can be jointly analyzed, statistically combined and/or studied at the level of the underlying fields or maps (cf. e.g. [143, 145, 162, 163, 165, 267–274]), other interesting observables in this context include CMB secondaries, cosmic shear, line intensity mapping and various cross-correlations (see e.g. [115–120, 158–160, 178, 275–278]).

of this work is based on observations obtained by the Sloan Digital Sky Survey III (SDSS-III, <http://www.sdss3.org/>). Funding for SDSS-III has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation and the US Department of Energy Office of Science. Parts of this work were performed using computing resources provided by the National Academic Infrastructure for Supercomputing in Sweden (NAISS) under Projects 2023/3-21, 2023/6-297, 2024/5-666 and 2024/6-339, which is partially funded by the Swedish Research Council through Grant 2022-06725. We acknowledge the use of CAMB [82], CLASS [128], CLASS-PT [127], FAST-PT [179], IPython [180] and MontePython [136, 137], and the Python packages Matplotlib [181], NumPy [182] and SciPy [183].

Chapter 3, in full, is a reprint of the material as it appears in Extending the Cosmological Collider: New Scaling Regimes and Constraints from BOSS, Green, Daniel; Han, Jiashu; Walisch, Benjamin, arXiv:2602.12232. The dissertation author was one of the primary investigators and authors of this paper.

Chapter 4

Conclusions

Throughout this dissertation, the main question has been whether the two-point statistics of the large-scale structure, and more specifically, the scale-dependent bias feature in the galaxy power spectrum, can be used to robustly probe the field content during inflation in the near-term. In Chapters 2 and 3, we answer this question through forecasts of current and near-future galaxy surveys as well as constraints derived from the BOSS DR12 dataset for two complementary regimes: light fields that generate a power-law scale dependence in the galaxy power spectrum, and heavier fields that leave behind oscillatory signatures. These two analyses together cover a wide range of masses and interaction couplings accessible to the galaxy power spectrum through the scale-dependent bias. In this chapter, instead of restating the findings of Chapters 2 and 3 in detail, we integrate the results and discuss the state and future of the search for primordial non-Gaussianity through large-scale structure measurements.

In Chapter 2, we studied the scale-dependent signature of light fields during inflation in the galaxy power spectrum by considering a minimally model-dependent framework that captures a continuum of behavior between the local and equilateral limits of primordial non-Gaussianity. We used forecasts based on the Fisher information formalism to show that the sensitivity of a galaxy survey to this signal depends strongly on its design parameters, such as the bias of the survey population, galaxy number densities, and the use of multiple tracers. In particular, we demonstrated that a multi-tracer survey configuration can substantially improve the achievable

constraints on the non-Gaussian amplitude f_{NL}^Δ , and in some cases this can compensate for the otherwise less favorable survey configurations. We then applied this framework to the BOSS DR12 dataset and obtained constraints on f_{NL}^Δ for a general scaling exponent Δ . We found the results to be consistent with both the previous measurements in the massless limit ($\Delta = 0$) and the forecasted sensitivities. This consistency gives us confidence that our modeling of the galaxy power spectrum and the forecasted sensitivities based on it reflect what is achievable with real survey data.

In Chapter 3, we extended the analysis to include oscillatory features in the scale-dependent bias, which arise from heavier fields during inflation. We described this qualitatively different class of signals by a scaling exponent α and a (logarithmic) oscillation frequency ν . Despite the fact that the amplitudes of these oscillatory signals suffer from the inherent suppression due to their production mechanism and the transfer mechanism to late-time observables, we showed that these oscillations, if present in the survey data, can be more easily isolated from the galaxy power spectrum than the power-law signatures considered in Chapter 2 due to their distinct shape. We forecasted the sensitivities for current and future galaxy surveys across the (α, ν) parameter space, and obtained (to our knowledge) the first observational constraints on this class of primordial non-Gaussianity using the BOSS DR12 dataset. As in Chapter 2, the BOSS constraints are consistent with our forecasted sensitivities, again validating the theoretical pipeline we used to connect primordial physics to late-time galaxy clustering.

Beyond the individual results presented in these two chapters, the recurring forecast-data agreement for two qualitatively different classes of signals (power-law behavior on the large scales and oscillations on the small scales) suggests that the modeling choices common to both analyses (the parameterization of the scale-dependent bias, the handling of degeneracies with the non-primordial galaxy biases, and the scale cuts) are applicable to a wide range of signals. Therefore, our work provides a tested and reasonably general pipeline for connecting a broad range of inflationary dynamics to the galaxy power spectrum. Additionally, it is worth pointing out that while higher- N -point statistics and map-level analyses are in principle capable

of extracting more primordial information than the power spectrum, their reach for non-local shapes of non-Gaussianity has historically been limited by observational and computational obstacles (see the related discussions in Chapters 2 and 3). Given these limitations, our results in Chapters 2 and 3 show that despite being the most well-understood observable of the large-scale structure, the galaxy power spectrum remains an under-exploited channel for constraining inflationary physics today.

Looking ahead, the findings of this work point in several directions. The most immediate next step is to apply the tested pipeline to current and next-generation galaxy surveys, including DESI, Euclid, SPHEREx, and LSST, whose larger volumes, higher number densities, and in some cases multi-tracer capabilities were already shown in our forecasts to substantially improve their sensitivity to primordial non-Gaussianity over BOSS. On the theoretical side, the models considered in this work constitute only a small set of a large space of possibilities. Extending the analysis of PNG beyond the squeezed limit to other bispectrum configurations, and to a broader range of spins, multi-field scenarios, and mechanisms such as chemical potentials (see the discussion in Chapter 2), would allow future work to probe an even wider range of inflationary models.

Furthermore, the suppression of the oscillatory signal arising from the averaging over small-scale features in the transfer of primordial fluctuations to the power spectrum, identified in Chapter 3, provides motivation for higher- N -point statistics and map-level analyses, which may be less susceptible to this suppression due to retaining more direct information about the underlying mode-coupling. There are also opportunities for complementary measurements from cross-correlations of the galaxy distribution with other observables such as CMB secondary anisotropies, weak lensing, and line intensity mapping (see the outlook discussion in Chapter 3). Finally, many practical challenges identified in both chapters, such as the identification and mitigation of large-angle systematics in future photometric surveys, and the modeling of the halo formation process based on universality relations, remain unresolved and will need to be addressed as survey precision continues to improve.

As the next generation of galaxy surveys becomes available, the large-scale structure will provide an increasingly precise laboratory for physics at energy scales far beyond the reach of any terrestrial experiment. This work forms part of a broader effort, spanning advances in both theoretical understanding and observational strategies, to extract the maximum possible information about the earliest moments of the Universe. Our results show that searches for scale-dependent and oscillatory signatures in the galaxy power spectrum can already make meaningful progress with existing methods and data, while providing a foundation for the substantially greater precision of upcoming surveys.

Appendix A

Appendix for Chapter 2

A.1 Forecasting Details

In this appendix, we collect additional information on our LSS Fisher forecasts. We first provide the full set of definitions underlying our model for the galaxy power spectrum and then describe the experimental specifications of all galaxy surveys employed in the main text.

A.1.1 Galaxy Power Spectrum Model

We model the galaxy power spectrum at one-loop order in standard Eulerian perturbation theory and in the bias expansion where the galaxy overdensity δ_g is a function of terms up to third order in the linear matter density contrast δ_m [cf. (2.8)]. The theoretical galaxy power spectrum for two biased tracers A and B is then given by (2.11). The various loop contributions

in that equation are defined as follows:

$$\sigma^4 = \int_{\vec{q}} P_{\text{lin}}^2(q), \quad (\text{A.1})$$

$$P_{22}(k) = 2 \int_{\vec{q}} \left[F_2^{(s)}(\vec{q}, \vec{k} - \vec{q}) \right]^2 P_{\text{lin}}(q) P_{\text{lin}}(|\vec{k} - \vec{q}|), \quad (\text{A.2})$$

$$P_{13}(k) = 6 P_{\text{lin}}(k) \int_{\vec{q}} F_3^{(s)}(\vec{k}, \vec{q}, -\vec{q}) P_{\text{lin}}(q), \quad (\text{A.3})$$

$$P_{\delta^2}(k) = 2 \int_{\vec{q}} F_2^{(s)}(\vec{q}, \vec{k} - \vec{q}) P_{\text{lin}}(q) P_{\text{lin}}(|\vec{k} - \vec{q}|), \quad (\text{A.4})$$

$$P_{s^2}(k) = 2 \int_{\vec{q}} F_2^{(s)}(\vec{q}, \vec{k} - \vec{q}) \left(\mu_-^2 - \frac{1}{3} \right) P_{\text{lin}}(q) P_{\text{lin}}(|\vec{k} - \vec{q}|), \quad (\text{A.5})$$

$$P_{\delta^2 \delta^2}(k) = 2 \int_{\vec{q}} P_{\text{lin}}(q) P_{\text{lin}}(|\vec{k} - \vec{q}|), \quad (\text{A.6})$$

$$P_{\delta^2 s^2}(k) = 2 \int_{\vec{q}} \left(\mu_-^2 - \frac{1}{3} \right) P_{\text{lin}}(q) P_{\text{lin}}(|\vec{k} - \vec{q}|), \quad (\text{A.7})$$

$$P_{s^2 s^2}(k) = 2 \int_{\vec{q}} \left(\mu_-^2 - \frac{1}{3} \right)^2 P_{\text{lin}}(q) P_{\text{lin}}(|\vec{k} - \vec{q}|), \quad (\text{A.8})$$

$$P_{\text{III}[2]}(k) = 2 P_{\text{lin}}(k) \int_{\vec{q}} \left\{ \frac{2}{7} \left[\left(\frac{\vec{k} \cdot \vec{q}}{k q} \right)^2 - 1 \right] \frac{2}{3} \mu_-^2 + \frac{8}{63} \right\} P_{\text{lin}}(q), \quad (\text{A.9})$$

with $\int_{\vec{q}} \equiv \int \frac{d^3 q}{(2\pi)^3}$ and $\mu_- \equiv \frac{\vec{q} \cdot (\vec{k} - \vec{q})}{q |\vec{k} - \vec{q}|}$. The functions $F_2^{(s)}(\vec{k}_1, \vec{k}_2)$ and $F_3^{(s)}(\vec{k}_1, \vec{k}_2, \vec{k}_3)$ are the symmetric, second- and third-order kernels of δ_m in standard perturbation theory [15],

$$F_2^{(s)}(\vec{k}_1, \vec{k}_2) = \frac{5}{7} + \frac{2}{7} \left(\frac{\vec{k}_1 \cdot \vec{k}_2}{k_1 k_2} \right)^2 + \frac{1}{2} \frac{\vec{k}_1 \cdot \vec{k}_2}{k_1 k_2} \left(\frac{k_1}{k_2} + \frac{k_2}{k_1} \right), \quad (\text{A.10})$$

$$\begin{aligned} F_3^{(s)}(\vec{k}_1, \vec{k}_2, \vec{k}_3) = & \frac{k^2}{27} \left[\frac{\vec{k}_1 \cdot (\vec{k}_2 + \vec{k}_3)}{k_1^2 (\vec{k}_2 + \vec{k}_3)^2} G_2^{(s)}(\vec{k}_2, \vec{k}_3) + 2 \text{ cyclic} \right] \\ & + \frac{7}{54} \vec{k} \cdot \left[\frac{\vec{k}_1 + \vec{k}_2}{(\vec{k}_1 + \vec{k}_2)^2} G_2^{(s)}(\vec{k}_1, \vec{k}_2) + 2 \text{ cyclic} \right] \\ & + \frac{7}{54} \vec{k} \cdot \left[\frac{\vec{k}_1}{k_1^2} F_2^{(s)}(\vec{k}_2, \vec{k}_3) + 2 \text{ cyclic} \right], \end{aligned} \quad (\text{A.11})$$

where $\vec{k} \equiv \vec{k}_1 + \vec{k}_2 + \vec{k}_3$ and $G_2^{(s)}(\vec{k}_1, \vec{k}_2)$ is the symmetric, second-order kernel of the velocity divergence θ_m ,

$$G_2^{(s)}(\vec{k}_1, \vec{k}_2) = \frac{3}{7} + \frac{4}{7} \left(\frac{\vec{k}_1 \cdot \vec{k}_2}{k_1 k_2} \right)^2 + \frac{1}{2} \frac{\vec{k}_1 \cdot \vec{k}_2}{k_1 k_2} \left(\frac{k_1}{k_2} + \frac{k_2}{k_1} \right). \quad (\text{A.12})$$

A.1.2 Survey Specifications

In the following, we provide detailed information about the spectroscopic and photometric redshift surveys that we employ in our Fisher forecasts, including the assumed redshift distribution in bias and number density.

Spectroscopic Surveys

The experimental specifications of the spectroscopic redshift surveys BOSS, DESI, Euclid, SPHEREx, MegaMapper and the billion-object survey are provided in Tables A.1 to A.6. We provide the linear bias b_1 and number density $\bar{n}_g$ for each sample and redshift bin with mean redshift $\bar{z}$ and spherical volume V . For DESI, different types of tracers were combined into one galaxy sample with a single effective number density and bias following [18, 111]. Since the number density of DESI is not in the regime where sample-variance cancellation is most effective (see Fig. 2.4), our results should be relatively insensitive to this choice. For all surveys, we treat each redshift bin as being independent so that the Fisher matrix for the entire survey is the sum of the Fisher matrices associated with each separate redshift bin.

Table A.1. Basic specifications for a BOSS-like survey [18] (inspired by [281] as detailed in [111]), covering a sky area of 10252 deg^2 with a total of about 1.2×10^6 objects in a volume of roughly $6.3 h^{-3} \text{ Gpc}^3$.

$\bar{z}$	0.35	0.625
b_1	1.634	1.877
$10^3 \bar{n}_g [h^3 \text{ Mpc}^{-3}]$	0.275	0.142
$V [h^{-3} \text{ Gpc}^3]$	2.18	4.15

Photometric Surveys

To forecast Vera Rubin Observatory’s LSST, which is a photometric redshift survey, we use the “gold” sample defined in their science book [107], which includes more than four billion galaxies over 20000 deg^2 . The number of objects in each redshift bin is calculated according to

$$N_g(z_{\min}, z_{\max}) = \frac{N_{\text{tot}} \Omega_{\text{sky}}}{2z_0^2} (2z_0^2 + 2z_0 z + z^2) e^{-z/z_0} \bigg|_{z=z_{\max}}^{z=z_{\min}}, \quad (\text{A.13})$$

where $N_{\text{tot}} = 55 \text{ arcmin}^{-2}$, $\Omega_{\text{sky}} = 20000 \text{ deg}^2$ and $z_0 = 0.3$. The bias is taken to be $b_1(z) = 0.95/D(z)$ and we assume a photometric redshift error of $\sigma_{z_0} = 0.05$. Table A.7 explicitly provides the derived experimental specifications for this survey with $z_{\max} = 3$ and twelve redshift bins with width $\Delta z = 0.25$.

Table A.2. Basic specifications for DESI [18] (derived from [104] as explained in [111]), covering a sky area of 14000 deg^2 with a total of about 2.7×10^7 objects in a volume of roughly $58 h^{-3} \text{ Gpc}^3$.

$\bar{z}$	0.05	0.15	0.25	0.35	0.45	0.65
b_1	1.40	1.48	1.55	1.61	1.67	2.05
$10^3 \bar{n}_g [h^3 \text{ Mpc}^{-3}]$	38.8	15.7	3.96	0.883	0.0992	0.591
$V [h^{-3} \text{ Gpc}^3]$	0.0356	0.229	0.560	0.979	1.44	2.39
$\bar{z}$	0.75	0.85	0.95	1.05	1.15	1.25
b_1	1.71	1.71	1.53	1.45	1.48	1.47
$10^3 \bar{n}_g [h^3 \text{ Mpc}^{-3}]$	1.31	0.920	0.779	0.466	0.398	0.387
$V [h^{-3} \text{ Gpc}^3]$	2.83	3.24	3.61	3.94	4.24	4.49
$\bar{z}$	1.35	1.45	1.55	1.65	1.75	1.85
b_1	1.47	1.69	1.68	2.27	2.45	2.47
$10^3 \bar{n}_g [h^3 \text{ Mpc}^{-3}]$	0.180	0.133	0.110	0.0387	0.0197	0.0208
$V [h^{-3} \text{ Gpc}^3]$	4.71	4.90	5.05	5.18	5.29	5.37

Table A.3. Basic specifications for the Euclid survey [111] (derived from [282]), covering a sky area of 15000 deg^2 with a total of about 4.9×10^7 objects in a volume of roughly $71 h^{-3} \text{ Gpc}^3$.

$\bar{z}$	0.65	0.75	0.85	0.95	1.05	1.15	1.25	1.35
b_1	1.06	1.11	1.16	1.21	1.27	1.33	1.38	1.44
$10^3 \bar{n}_g [h^3 \text{ Mpc}^{-3}]$	0.637	1.46	1.63	1.50	1.33	1.14	1.00	0.837
$V [h^{-3} \text{ Gpc}^3]$	2.56	3.03	3.47	3.87	4.23	4.54	4.81	5.05
$\bar{z}$	1.45	1.55	1.65	1.75	1.85	1.95	2.05	
b_1	1.51	1.54	1.63	1.70	1.85	1.90	1.26	
$10^3 \bar{n}_g [h^3 \text{ Mpc}^{-3}]$	0.652	0.512	0.357	0.246	0.149	0.0904	0.0721	
$V [h^{-3} \text{ Gpc}^3]$	5.25	5.41	5.55	5.67	5.76	5.83	5.88	

Table A.4. Basic specifications for SPHEREx [283], covering a sky fraction $f_{\text{sky}} = 0.75$ with a total of about 7.0×10^8 objects in a volume of roughly $450 h^{-3} \text{Gpc}^3$. SPHEREx is a spectro-photometric survey and the observed objects are divided into five samples based on their photometric redshift uncertainty bin, with respective maximum error $\sigma_{z0} = \{0.003, 0.01, 0.03, 0.1, 0.2\}$. The volume V and the galaxy number density $\bar{n}_g$ are given in units of $h^{-3} \text{Gpc}^3$ and $h^3 \text{Mpc}^{-3}$, respectively.

$\bar{z}$	0.1	0.3	0.5	0.7	0.9	1.3
V	0.584	3.40	7.43	11.5	15.1	60.4
$b_1^{(1)}$	1.3	1.5	1.8	2.3	2.1	2.7
$10^5 \bar{n}_g^{(1)}$	997	411	50.1	7.05	3.16	1.64
$b_1^{(2)}$	1.2	1.4	1.6	1.9	2.3	2.6
$10^5 \bar{n}_g^{(2)}$	1230	856	282	93.7	43.0	5.00
$b_1^{(3)}$	1.0	1.3	1.5	1.7	1.9	2.6
$10^5 \bar{n}_g^{(3)}$	1340	857	362	294	204	21.2
$b_1^{(4)}$	0.98	1.3	1.4	1.5	1.7	2.2
$10^5 \bar{n}_g^{(4)}$	2290	1290	535	495	415	79.6
$b_1^{(5)}$	0.83	1.2	1.3	1.4	1.6	2.1
$10^5 \bar{n}_g^{(5)}$	1490	752	327	250	183	73.4
$\bar{z}$	1.9	2.5	3.1	3.7	4.3	
V	71.4	73.6	71.8	68.3	64.2	
$b_1^{(1)}$	3.6	2.3	3.2	2.7	3.8	
$10^5 \bar{n}_g^{(1)}$	0.359	0.0807	0.184	0.150	0.113	
$b_1^{(2)}$	3.4	4.2	4.3	3.7	4.6	
$10^5 \bar{n}_g^{(2)}$	0.803	0.383	0.328	0.107	0.0679	
$b_1^{(3)}$	3.0	3.2	3.5	4.1	5.0	
$10^5 \bar{n}_g^{(3)}$	0.697	0.202	0.143	0.193	0.0679	
$b_1^{(4)}$	3.6	3.7	2.7	2.9	5.0	
$10^5 \bar{n}_g^{(4)}$	7.75	0.787	0.246	0.193	0.136	
$b_1^{(5)}$	3.2	4.2	4.1	4.5	5.0	
$10^5 \bar{n}_g^{(5)}$	25.3	5.41	2.99	0.941	0.204	

Table A.5. Basic specifications for MegaMapper [108], covering a sky area of 14000 deg^2 with a total of about 6.6×10^7 objects in a volume of roughly $200 h^{-3} \text{ Gpc}^3$.

$\bar{z}$	2.0	3.0	4.0	5.0
b_1	2.5	4.0	3.5	5.5
$10^3 \bar{n}_g [h^3 \text{ Mpc}^{-3}]$	0.98	0.12	0.10	0.040
$V [h^{-3} \text{ Gpc}^3]$	54.0	54.4	50.0	44.8

Table A.6. Basic specifications for our (futuristic) spectroscopic billion-object survey, covering a sky fraction $f_{\text{sky}} = 0.5$ with a total of 10^9 objects in a volume of roughly $330 h^{-3} \text{ Gpc}^3$. The observed objects are divided into two samples based on their linear bias, $b_1(z=0) = 2.0$ and 1.2 .

$\bar{z}$	0.25	0.75	1.25	1.75	2.25
$V [h^{-3} \text{ Gpc}^3]$	4.78	20.6	32.8	38.8	40.8
$b_1^{(1)}$	2.28	2.94	3.66	4.41	5.17
$10^3 \bar{n}_g^{(1)} [h^3 \text{ Mpc}^{-3}]$	1.54	1.54	1.54	1.54	1.54
$b_1^{(2)}$	1.37	1.76	2.19	2.64	3.10
$10^3 \bar{n}_g^{(2)} [h^3 \text{ Mpc}^{-3}]$	1.54	1.54	1.54	1.54	1.54
$\bar{z}$	2.75	3.25	3.75	4.25	4.75
$V [h^{-3} \text{ Gpc}^3]$	40.7	39.4	37.7	35.9	34.0
$b_1^{(1)}$	5.94	6.72	7.51	8.29	9.07
$10^3 \bar{n}_g^{(1)} [h^3 \text{ Mpc}^{-3}]$	1.54	1.54	1.54	1.54	1.54
$b_1^{(2)}$	3.57	4.03	4.50	4.97	5.45
$10^3 \bar{n}_g^{(2)} [h^3 \text{ Mpc}^{-3}]$	1.54	1.54	1.54	1.54	1.54

Table A.7. Basic specifications for Vera Rubin Observatory's LSST [107], covering a sky area of 20000 deg^2 with a total of about 4×10^9 objects in a volume of roughly $170 h^{-3} \text{ Gpc}^3$. The photometric redshift error is $\sigma_{z0} = 0.05$.

$\bar{z}$	0.125	0.375	0.625	0.875	1.125	1.375
b_1	1.01	1.16	1.31	1.48	1.65	1.82
$10^3 \bar{n}_g [h^3 \text{ Mpc}^{-3}]$	292	183	108	63.6	36.8	21.0
$V [h^{-3} \text{ Gpc}^3]$	0.709	3.93	8.10	11.9	14.9	17.0
$\bar{z}$	1.625	1.875	2.125	2.375	2.625	2.875
b_1	2.00	2.18	2.36	2.54	2.73	2.91
$10^3 \bar{n}_g [h^3 \text{ Mpc}^{-3}]$	11.9	6.46	3.55	1.94	0.969	0.646
$V [h^{-3} \text{ Gpc}^3]$	18.4	19.2	19.7	19.9	19.8	19.6

A.2 Measuring the Scaling Behavior

We have assumed for all forecasts throughout the paper that the fiducial value of the non-Gaussian amplitude f_{NL}^Δ vanishes. This is of course motivated by the fact that we have only been placing observational bounds on PNG that are consistent with $f_{\text{NL}}^\Delta = 0$, including in Section 2.5. We additionally always considered the scaling exponent Δ to be fixed. In this appendix, we will consider nonzero fiducial PNG amplitudes and clarify to what degree measurements of the galaxy power spectrum are sensitive to the precise value of Δ .

We defined the correlation between the non-Gaussian signals of two different values of Δ in (2.33) in terms of the respective Fisher matrix elements (after marginalization) for a fiducial value of $\bar{f}_{\text{NL}}^{\Delta_{1,2}} = 0$. For BOSS, we can observe in Fig 2.10 that this correlation is close to unity, even for widely separated values of Δ . We see in Fig. A.1 that this statement still holds for the billion-object survey. In fact, the range of Δ which are highly correlated with the local shape further expanded. Naively, this suggests there is little information about the shape of the signal, described by Δ , encoded in these observables.

To first approximation, the fiducial value of $\bar{f}_{\text{NL}}^\Delta = 0$ is the reason why we see little ability to distinguish different values of Δ . Since there is no signal to measure, it is not surprising that the precise shape is not very important. If we were instead to detect f_{NL}^Δ , we would expect that the signal to noise would be a fairly sensitive function of Δ and we could therefore measure the true value of Δ precisely. This is exactly what is found in Fig. A.2, in which we show the forecasted constraints for measuring Δ , $\sigma(\Delta)$, assuming that f_{NL}^Δ is detected at a significance of 5σ in the billion-object survey. The key result is that if we detect f_{NL}^Δ while holding Δ fixed, we expect to have enough sensitivity to measure the scaling exponent with $\sigma(\Delta) \sim 0.05 - 0.5$. In this sense, the strategy of searching for f_{NL}^Δ at fixed Δ is a reliable strategy to search for new physics, but would yield additional information about the origin of the signal even at the threshold of a detection. When we instead marginalize over Δ , we require a significantly larger non-Gaussian amplitude f_{NL}^Δ to be detected. In this case, the corresponding uncertainties on Δ

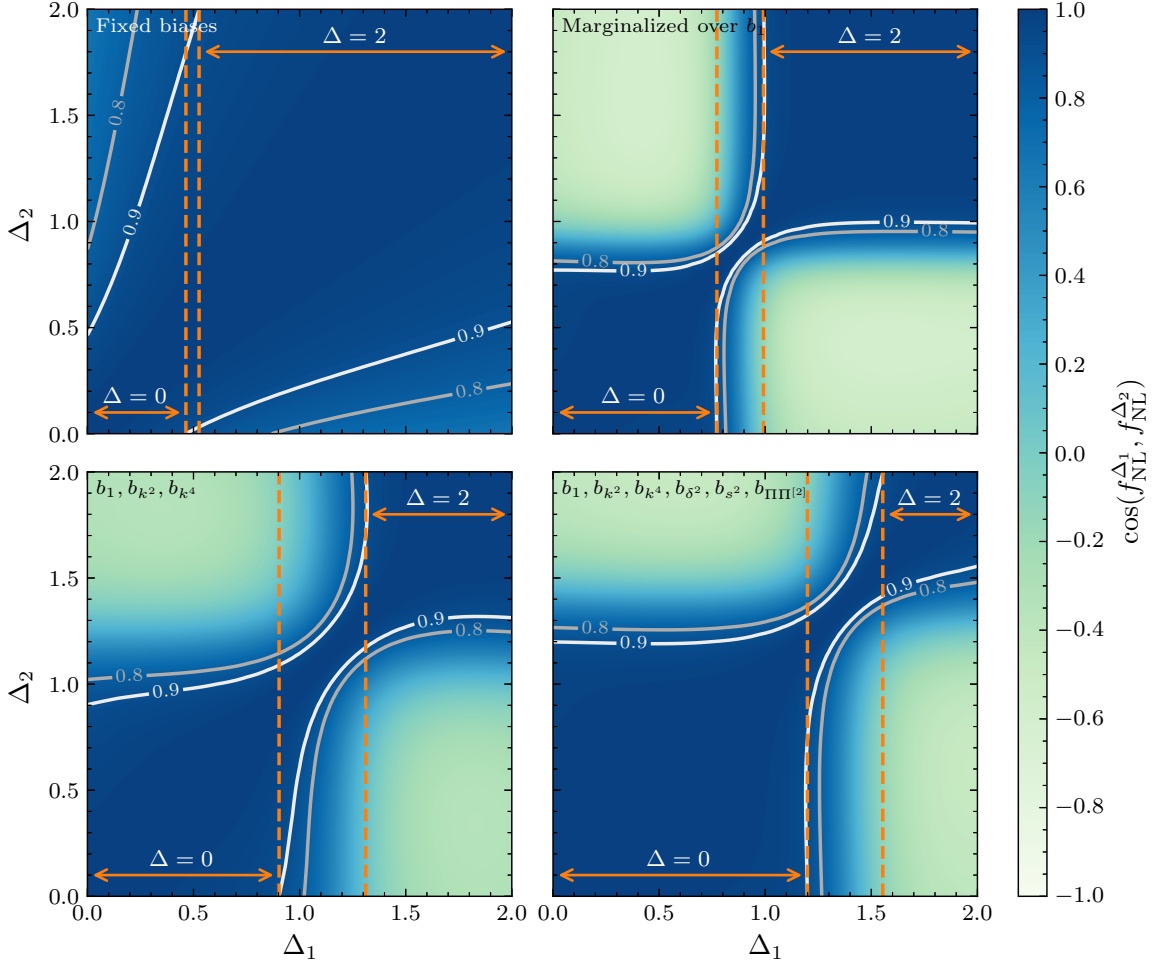


Figure A.1. Correlation matrix for measurements of galaxy power spectra with different values of the non-Gaussian scaling exponent Δ as defined in (2.33) for the billion-object survey and four biasing models. The lower right panel is equivalent to Fig. 2.10, which shows the correlations for BOSS. As in that figure, we hold the Λ CDM parameters fixed, take the fiducial non-Gaussian amplitudes to be zero, $\tilde{f}_{\text{NL}}^{\Delta_i} = 0$, and indicate the coverage of the local and equilateral templates with a correlation coefficient of at least 0.9 by the orange lines. We can clearly observe the impact of the biasing model and the different experimental specifications compared to BOSS.

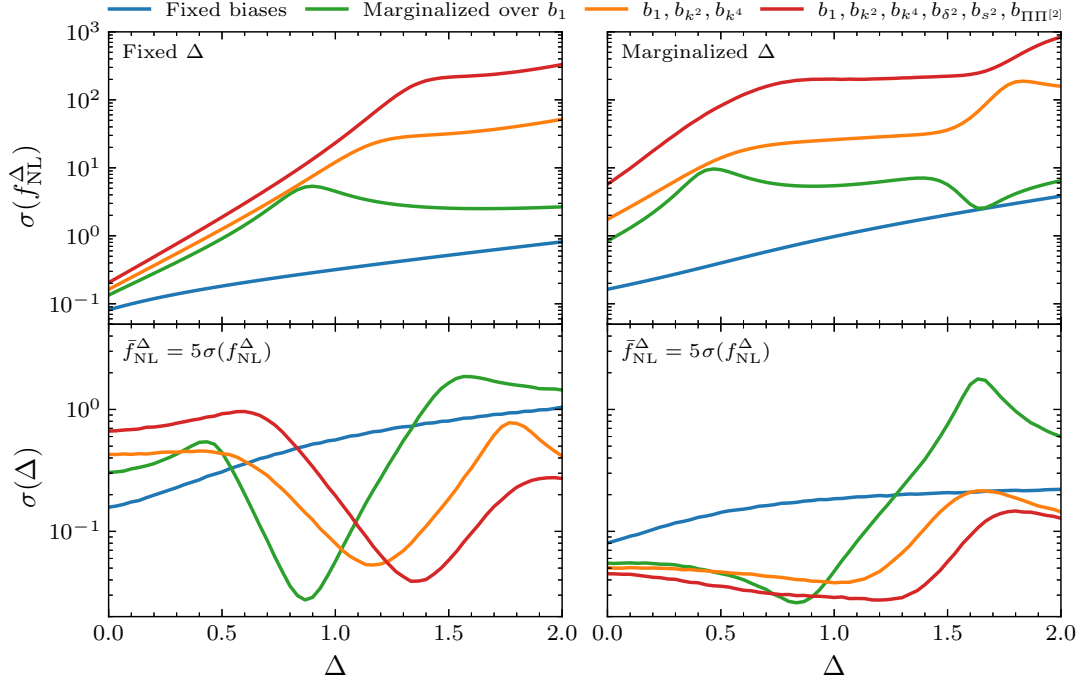


Figure A.2. Forecasts for $\sigma(f_{\text{NL}}^\Delta)$ (*top*) and $\sigma(\Delta)$ (*bottom*) as a function of Δ if the scaling exponent is either fixed (*left*) or varied (*right*) for the billion-object survey. To compute the constraint on Δ , we take the fiducial value of the non-Gaussian amplitude, $\bar{f}_{\text{NL}}^\Delta$, to be the value necessary to achieve a 5σ detection in the respective forecast. We fix the Λ CDM parameters and marginalize over four different biasing models.

would however be generally smaller, with $\sigma(\Delta) < 0.1$.

At first sight, it might seem surprising that the curves of $\sigma(\Delta)$ in Fig. A.2 have minima at the values of Δ where the curves of $\sigma(f_{\text{NL}}^\Delta)$ have maxima or downward bends. By comparing the curves for different biasing models, we can however see that these features are the result of marginalizing over the bias parameters. For example, we noted in the main text that the maximum in $\sigma(f_{\text{NL}}^\Delta)$ near $\Delta = 1$ when marginalizing over b_1 arises because the linear bias b_1 and $f_{\text{NL}}^{\Delta=1}$ are degenerate at high k . The fact that the measurement of f_{NL}^Δ is less constraining at this point then also implies that we are very sensitive to the exact value of the scaling exponent Δ . When we hold $f_{\text{NL}}^\Delta/\sigma(f_{\text{NL}}^\Delta)$ fixed, we should therefore expect that $\sigma(\Delta)$ is smaller in regions of Δ where f_{NL}^Δ is degenerate with bias parameters.

Let us finally return to the correlation coefficients between scale-dependent biases with different scaling exponent Δ . In order to be consistent, the impact of marginalization over the bias parameters must also appear in $\cos(f_{\text{NL}}^{\Delta_1}, f_{\text{NL}}^{\Delta_2})$. Specifically, the impact of the marginalization has a large impact on $\sigma(f_{\text{NL}}^\Delta)$ for larger values of Δ . For the inferred constraints to be consistent with our forecasts, this has to imply that the correlation between large and small Δ decreases by a similar factor. If this was not the case, our constraints inferred from $f_{\text{NL}}^{\text{loc}}$ at large Δ would be stronger than directly measuring f_{NL}^Δ .

This expectation is precisely what occurs, as shown in Fig. A.1. The correlation coefficients for larger Δ are affected by the number of bias parameters that are marginalized over. With more marginalized bias parameters, the range over which $\Delta < 2$ and $\Delta = 2$ are degenerate decreases significantly. This is expected since the (equilateral) signal for $\Delta = 2$ gets most of its information from large wavenumbers, where the biases play a significant role (cf. §2.4.2). In contrast, the increase of the range of strong overlap with $\Delta = 0$ is likely due to the increasing difficulty to distinguish similar shapes of Δ with the inclusion of more bias parameters. Finally, it is also worth noticing that the non-Gaussian signal with $\Delta = 0$ (local PNG) appears to be anti-correlated with the scale-dependent bias signal for $\Delta \sim 2$ in the last three panels of Fig. A.1

in which we marginalize over the biasing model.

Appendix B

Appendix for Chapter 3

B.1 Details for the Fisher Forecasts

The forecasts presented in Section 3.4 were obtained following standard Fisher methodology. Many of the specific choices and experimental specifications are described in detail in [193]. In this appendix, we summarize the implemented changes to address the subtleties induced by high-frequency oscillations that are unique to the models discussed in this paper. We also include additional forecasting results.

Bandpowers

A lot of useful intuition for measuring primordial oscillations can be derived from time-series analysis. We measure the power spectrum at a finite number of wavenumbers k_i and can therefore only extract a finite list of frequencies. Unlike most time series, however, we do not actually measure the power spectrum at a unique set of points. Instead, different Fourier modes $\vec{k}$ are binned according to their length k . As the oscillatory frequencies approach the bin size (or spacing), this detail can significantly impact the results.

In order to model the binning procedure, we computed the Fisher matrix by summing over discrete wavenumbers k_i based on the minimum binning width Δk , which is conservatively set by the volume V of a given redshift bin, and the minimum accessible wavenumber $k_{\min} = \Delta k$. For a sufficiently small Δk , the power spectrum is smooth and we can approximate it in a

given band $[k_i - \Delta/2, k_i + \Delta/2]$ by taking the midpoint value $P_i = P_g(k_i)$. For high-frequency oscillations, there will however be significant variations within each band and our ability to resolve these features will be limited, as depicted in Fig. B.1. If we do not take this into account, our results would suffer from aliasing. This is why we compute the bandpowers P_i as the average power spectrum in a given bin centered around k_i ,

$$P_i = \frac{1}{\Delta k} \int_{k_i - \Delta k/2}^{k_i + \Delta k/2} dk P_g(k). \quad (\text{B.1})$$

When numerically implementing this, it is important to sufficiently sample the theoretical galaxy power spectrum P_g to resolve the oscillations which requires a sampling step of $\Delta k' < \Delta k$, with $\Delta k' \leq \delta k(k_{\min}, \nu_{\max})$, where $\nu_{\max}$ is the largest logarithmic frequency considered. We introduced here $\delta k(k, \nu)$ as the separation of the zeroes of the oscillatory term $O(k)$ defined in (3.46) which depends on both the wavenumbers k and the logarithmic frequency ν , as illustrated in Fig. B.1. To use our intuition from time series and linear features, we can convert δk to an (effective) linear

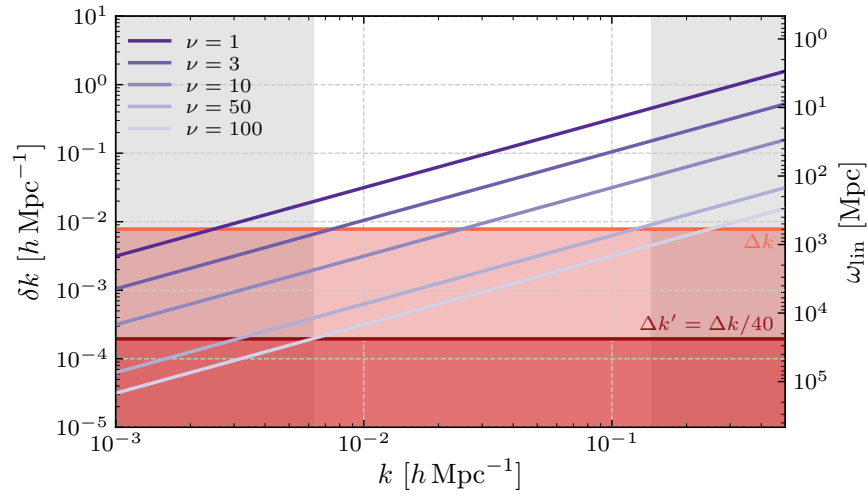


Figure B.1. Separation of zeroes δk of the isolated oscillatory signal (or, alternatively, its effective linear frequency ω_{lin}) for five values of the logarithmic frequency ν as a function of wavenumber k . The gray-shaded regions indicate the wavenumbers that are inaccessible in our BOSS forecasts below $k_{\min}$ and above $k_{\max}$, respectively, and the red-shaded regions are those in which the survey is partially (light red) and completely (dark red) insufficient to resolve the oscillations at $k_{\min}$, highlighting the need to use bandpowers in our forecasts.

frequency ω_{lin} which naturally decreases with increasing wavenumber for a given logarithmic frequency.

To illustrate the ability to resolve the logarithmic oscillations over the range of wavenumbers used in our forecasts for BOSS, we also highlight their values of k_{min} , k_{max} and bandwidth Δk (see also footnote 22). Signals in the white region of Fig. B.1 ($\delta k > \Delta k$) are fully resolved by the survey, those in the light-shaded region ($\Delta k' < \delta k < \Delta k$) are partially affected by the bandpass filtering and those in the dark-shaded region ($\delta k \lesssim \Delta k' = \Delta k/40$) would be entirely affected for frequencies $\nu > 100$. This also explains and the figure visualizes the maximum frequency $\nu_{\text{max}} = 55$ for which we forecast the constraining power of BOSS in §3.4.3. As a point of comparison, we found that the BOSS forecasts using the midpoint approximation lead to approximately the same constraints as those including the bandpowers if $\nu \lesssim 20$. We finally note that our bandpass filtering is additionally serving as an approximate way to include the effects of the survey window function on the oscillations as shown in [18].

Phase Dependence

The purpose of forecasting is to anticipate the behavior of our analysis of observed data and estimate the sensitivity of future surveys. One of the downsides of a Fisher approach however is that the uncertainties are effectively implemented as Gaussians around a fiducial value. This is important to keep in mind and one of the reasons why we employ the amplitude instead of the phase parametrization as mentioned in the main text. We explicitly show in Fig. B.2, how the forecasted uncertainty for the amplitude(s) is affected by the fiducial value of the phase when characterizing the logarithmic oscillations in terms of an amplitude and a phase. Specifically, we display the two-dimensional contours in the $(A_{\text{cos}}, A_{\text{sin}})$ plane for $\alpha = 0$ and $\nu = 1$, as in Fig. 3.8, but for five different values of the fiducial phase. The forecasts in this case are computed by setting the fiducial values $(A_{\text{cos}}^{\text{fid}}, A_{\text{sin}}^{\text{fid}})$ as given by the fiducial phase $\phi_{\text{fid}} = \arctan(-A_{\text{sin}}^{\text{fid}}/A_{\text{cos}}^{\text{fid}})$ and overall amplitude. We then calculate the posterior distribution around that (necessarily non-Gaussian) fiducial point in the $(A_{\text{cos}}, A_{\text{sin}})$ plane and convert the

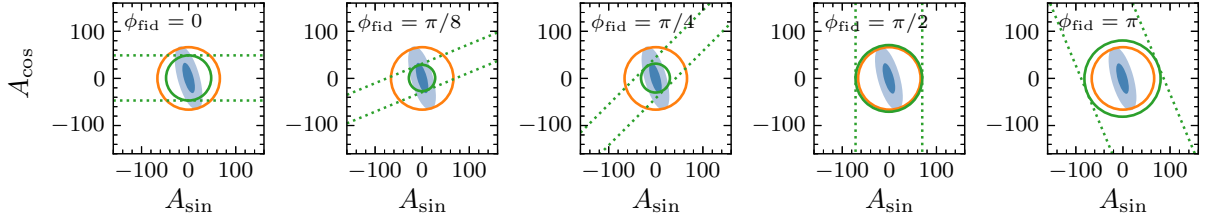


Figure B.2. Forecasted posterior distributions for $(A_{\cos}, A_{\sin})$ and the derived limits on $Af_{\text{NL}}^{\alpha, \nu}$ for different fiducial phases $\phi_{\text{fid}} = \arctan(-A_{\sin}^{\text{fid}}/A_{\cos}^{\text{fid}})$ if we had used the phase instead of the amplitude parametrization to model the logarithmic oscillations. The scaling dimension is set to $\alpha = 0$ and $\nu = 1$, and the phase is varied by setting $A_{\cos}^{\text{fid}} = \cos \phi_{\text{fid}}$ and changing $A_{\sin}^{\text{fid}}$ accordingly. The light and dark blue contours represent the two-dimensional posterior distributions at 1σ and 2σ , respectively, while the orange circle encloses 95% of the probability centered around the origin to indicate the 95% upper limit on $Af_{\text{NL}}^{\alpha, \nu}$. The green dotted lines visualize the constraints projected along the direction specified by the fiducial phase which then translates to the green circle displaying the corresponding limit on $Af_{\text{NL}}^{\alpha, \nu}$. We can clearly observe the dependence of these projected limits on the fiducial phase due to the highly degenerate nature of $A_{\cos}$ and $A_{\sin}$ for the considered scaling dimension, showing why the 95% upper limit (orange circle) is a more robust indicator of the survey sensitivity to the overall non-Gaussian amplitude.

two-dimensional distribution to a 95% exclusion limit on $Af_{\text{NL}}^{\alpha, \nu}$ by taking the half-width of the 2σ contour along the direction specified by the fiducial phase. This is displayed by the green dotted lines and their corresponding circles. While the forecasted two-dimensional posterior for $(A_{\cos}, A_{\sin})$ is unaffected, the inferred 95% exclusion limits on the overall amplitude clearly vary with the fiducial phase and can therefore obfuscate the survey sensitivity by directly depending on its value. We consequently followed the method developed in [18] to work in the amplitude parametrization and implicitly marginalize over the phase when deriving the 95% upper limits on the non-Gaussian amplitude $Af_{\text{NL}}^{\alpha, \nu}$ (orange circles). As an additional advantage, this also allows us to fiducially pick the Gaussian cosmology with $A_{\cos} = A_{\sin} = 0$ which would not have been possible with the phase parametrization.

Cosmology Dependence

Our results depend on cosmology (and astrophysics) directly through the signal and noise that enter the forecasts, and indirectly through choices in our analyses and parameters we marginalize over. The direct cosmological impact is often self-evident, while the indirect effects

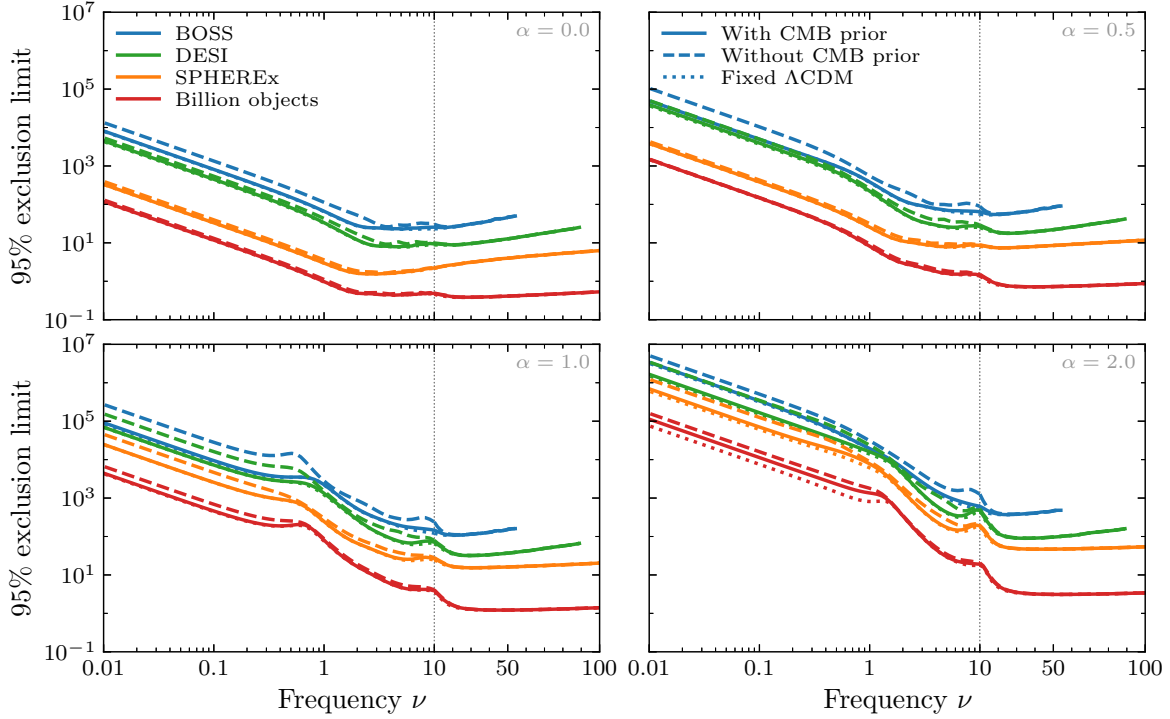


Figure B.3. Effect of marginalizing over the cosmological parameters and imposing a CMB prior on the forecasted 95% exclusion limits on the non-Gaussian amplitude $Af_{\text{NL}}^{\alpha, \nu}$ for four values of α as a function of ν . The solid lines, which show the results when marginalizing over the ΛCDM parameters with a Planck CMB prior, are the same as in Fig. 3.6. The dashed and dotted lines display the limits without imposing this prior and when fixing the cosmological parameters to their fiducial values while still marginalizing over the bias parameters. The impact of the CMB prior on the forecasts amounts to factors of $O(1)$ for small frequencies ($\nu \lesssim 10$) and percent-level to negligible differences at high frequencies ($\nu \gtrsim 10$) as expected. The difference between imposing a CMB prior and fixing the cosmological parameters is very small, except for the billion-object survey at large α and small ν .

may go unnoticed. Given that we scan over a wide range of frequencies, with the signal moving in k with the scaling exponents, it may be unclear where the changes in sensitivity originate from.

We illustrate the effects that marginalizing over the cosmological parameters and imposing priors from a Planck-like experiments have on the forecasted non-Gaussian limits in Fig. B.3. We clearly see that forecasts for $\nu > 10$ are unaffected by marginalizing over ΛCDM without any priors. This quantifies our expectation that the non-Gaussian signal at high frequencies is not degenerate with the smoother changes of the galaxy power spectrum due to cosmological pa-

rameters. Perhaps more surprising is that the impact of the Planck prior for $\nu < 1$ is not dramatic although visible up to factors of $O(1)$. An important take-away for the analysis of BOSS data in Section 3.5 is that fixing the Λ CDM parameters to their fiducial values leads to essentially the same results as varying them with the mentioned CMB prior. Given the goals of this work and the dramatic speed up in the statistical analysis of the data, we indeed adopted this approach as detailed in the main text. In fact, fixing the cosmological parameters and marginalizing over them with a Planck prior only lead to changes at the level of a few to 20% depending on α even for future surveys (especially around $\nu = 10$ due to the discussed degeneracy with the scale of baryon acoustic oscillations), except for the billion-object survey at large α and small ν for which the forecasted limits are more visibly affected.

In Fig. B.4, we investigate the impact of a variety of cosmological and astrophysical factors on our forecasts for the billion-object survey through the lens of the employed maximum wavenumber k_{max} by fixing and varying the cosmological and bias parameters, respectively. If our forecasts are limited only by cosmic variance, then increasing the number of modes should dramatically improve the sensitivity. This relative improvement is seen consistently for $\nu > 30$ for all values of α , regardless of the details of the analysis. This is consistent with our expectation that the high-frequency regimes are robust to most degeneracies. In contrast, we notice a significant feature around $\nu = O(10)$ for the full forecasts that is absent when the Λ CDM parameters are held fixed. This feature is consistent with being the result of the degeneracy between ν and the scale of baryon acoustic oscillations which corresponds to a linear frequency of 150 Mpc (cf. Fig. B.1). We also see many features around $\nu = O(1)$ due to degeneracies with both biasing and cosmological parameters. Overall, we clearly see that the high-frequency regime is very robust due to its oscillatory nature breaking degeneracies with the other effects that all result in smooth changes to the galaxy power spectrum, as is also evident at small frequencies.

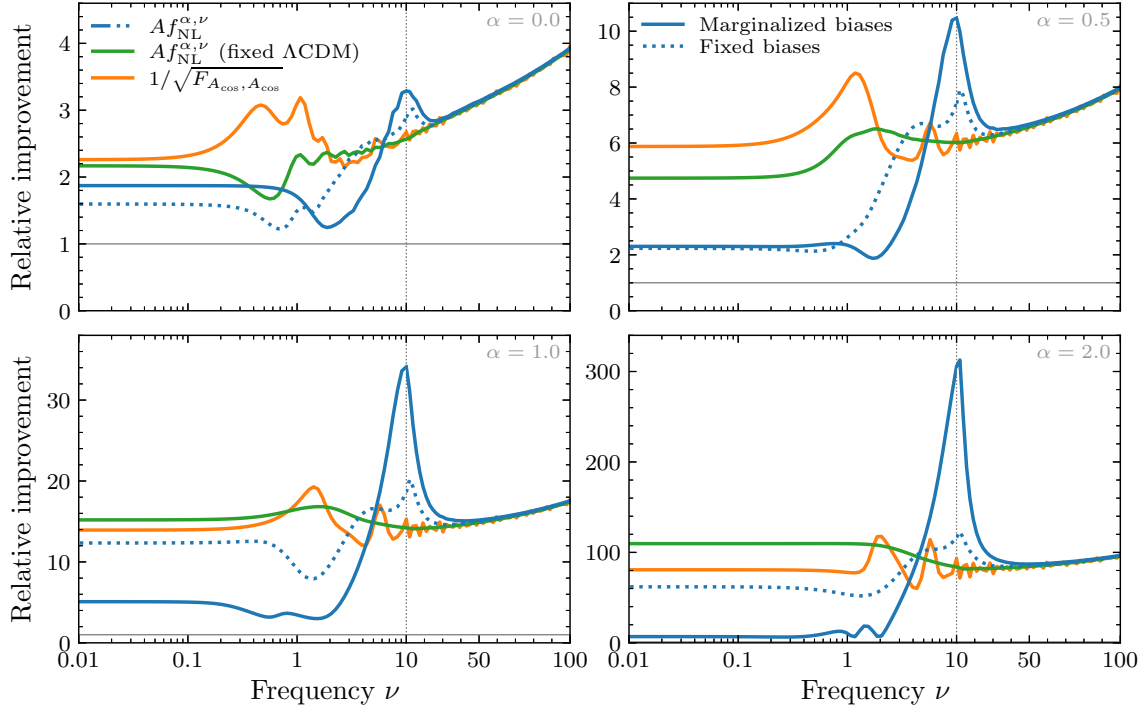


Figure B.4. Relative improvement of the 95% upper limit on $Af_{\text{NL}}^{\alpha, \nu}$ for the billion-object survey when increasing the maximum wavenumber from our fiducial and conservative choice of $k_{\text{max}} = k_{\text{halo}}$ to the more optimistic $k_{\text{max}} = k_{\text{NL}}(z)$ as a function of frequency ν for four scaling exponents. The solid blue line is the ratio of the lines displayed in Fig. 3.9, while the dotted blue line fixes instead of marginalizes over the bias parameters. For the green and orange curves, the biases were marginalized, but the ΛCDM parameters and all parameters except for A_{cos} fixed, respectively. We can observe that the relative improvement is robust to these variations for $\nu \gtrsim 30$, shows the largest value for $\nu = O(10)$ due to the degeneracy with the baryon acoustic oscillations and exhibits the underlying degeneracies in the small-frequency regime, especially for increasing scaling exponent α .

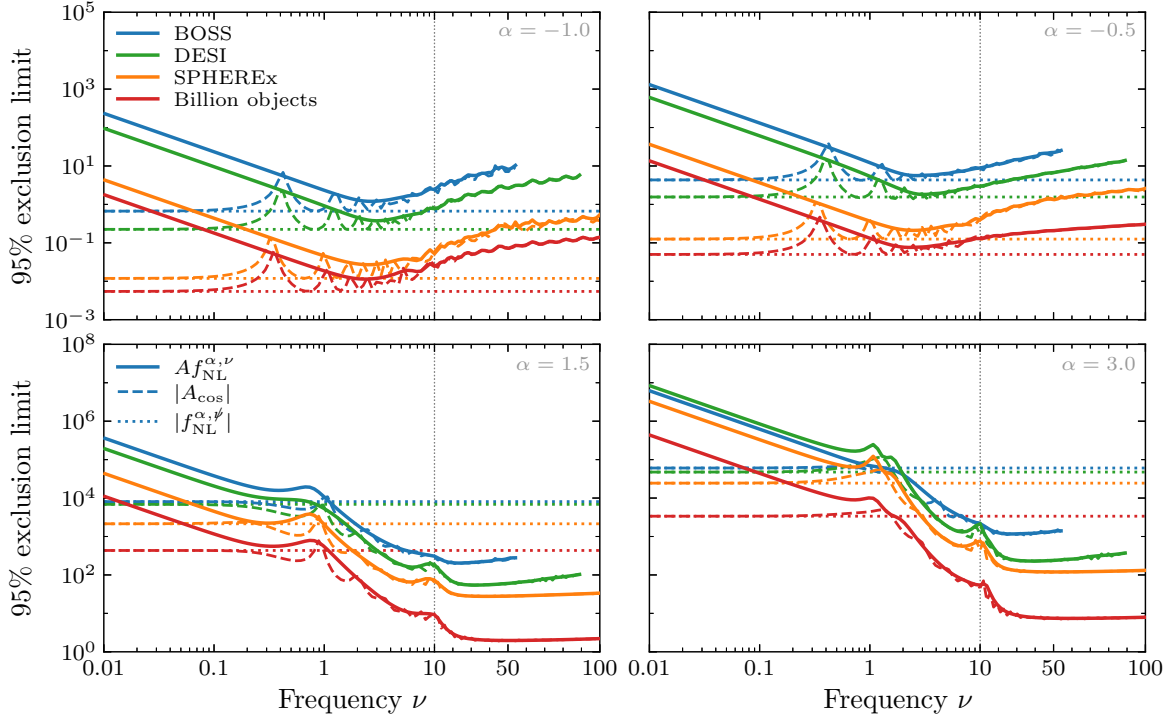


Figure B.5. Forecasted constraints on the amplitude of primordial non-Gaussianity in terms of its 95% exclusion limit as a function of the oscillation frequency ν for four additional values of the scaling exponent α than in Fig. 3.6. We refer to the caption in the main text for the description and explanation of the displayed curves, but note that the range of limits differs here between the top and bottom panels by three orders of magnitude. These results extend the values of α to the entire range investigated in this work and additionally complement the relative exclusion limits of the oscillatory with respect to the non-oscillatory signals presented in Fig. 3.7.

Additional Forecasts

To conclude this appendix, we show the forecasted upper limits on $Af_{\text{NL}}^{\alpha,\nu}$ for four additional values of the scaling exponent α in Fig. B.5 to complement the results displayed in Figures 3.6 and 3.7 in the main text. The negative values of $\alpha = -1$ and -0.5 further highlight that the forecasts are generally worse for $\nu \neq 0$ than for $\nu = 0$ when the scale-dependent bias signal is dominant at small wavenumbers. This is compatible with the idea that the low- k information is lost by marginalizing over the phase at small ν without sufficient gains from the high- k modes for which the signal vanishes, even at large ν . At the same time, the overall constraint on the non-Gaussian amplitude is tighter than for positive α , as expected. The

cases of $\alpha = 1.5$ and 3 have signals that peak at larger k where they can be distinguished from nonlinearities when $\nu > 10$. The case $\alpha = 1.5$ is of particular note here because it corresponds to the classical cosmological-collider signal of a coupling to a single massive particle. Finally, we see that the improvement in forecasts from BOSS through a billion-object survey are largely independent of the scaling exponent α . For $\nu > 10$, the improvement by roughly an order of magnitude is effectively independent of α , showing that the improvement in the sensitivity is not strongly affected by the nature of the signal.

B.2 Details for the BOSS Analysis

In this appendix, we collect additional information from our data analysis of the BOSS DR12 dataset presented in Section 3.5. We present three additional figures to illustrate the posterior distributions of the oscillatory amplitudes and our inference of the 95% limits, the obtained exclusion limits on the oscillatory PNG amplitude, and those on the non-oscillatory PNG amplitude.

To provide additional insight into our parameter inference and our procedure for inferring the 95% exclusion limits on the overall PNG amplitude $Af_{\text{NL}}^{\alpha, \nu}$, Fig. B.6 shows all posterior distributions directly based on the MCMC samples for the scaling exponent $\alpha = 0$ in the plane of the oscillatory amplitudes A_{cos} and A_{sin} . We observe the same rotating degeneracy line with increasing frequency ν as predicted in our forecasts of Fig. 3.8. For frequencies $\nu \gtrsim 1$, this degeneracy is effectively absent and the posterior is close to isotropic, while for $\nu \lesssim 1$ the degeneracy becomes very pronounced. As a result, the orange circle, whose radius is the 95% exclusion limit on $Af_{\text{NL}}^{\alpha, \nu}$ since it encloses 95% of the Monte Carlo samples (see §3.5.1 for details), provides an accurate summary of the statistical sensitivity of the survey to our PNG signal at large frequencies, but underestimates the constraining power for small ν in the form of weaker reported limits. We also note that the other eight values of α exhibit qualitatively the same patterns and behaviors as shown here for $\alpha = 0$, with some scaling dimensions exhibiting non-Gaussian posteriors in $(A_{\text{cos}}, A_{\text{sin}})$.

We presented the 95% exclusion limits on the oscillatory PNG amplitude $Af_{\text{NL}}^{\alpha, \nu}$ in Fig. 3.11 as a function of the frequency ν for the nine sampled exponents α . In Figure B.7, we display these limits as a function of both ν and α providing an additional way of interpreting especially their qualitative behavior. The degradation of the inferred limits with decreasing frequency and increasing exponent is clearly visible both in the color gradient and in the contours of constant exclusion limit. This behavior reflects the interplay between the effective wavenumber range probed by the signal at a given signal-to-noise ratio of the survey and the ability of the

non-Gaussian oscillations to break degeneracies with smooth bias contributions, as discussed in detail in the main text in the context of both the forecasts (§3.4.3) and the data analysis (§3.5.2).

Finally, we show the 95% exclusion limits on the absolute value of the non-oscillatory PNG amplitude $|f_{\text{NL}}^{\alpha,\psi}|$ as a function of the exponent α in Fig. B.8. Since we extended the range of scaling exponents studied in this work compared to [193] from $[0, 2]$ to $[-1, 3]$ (see the main text for the phenomenological motivation), we reran the analysis of [193] in order

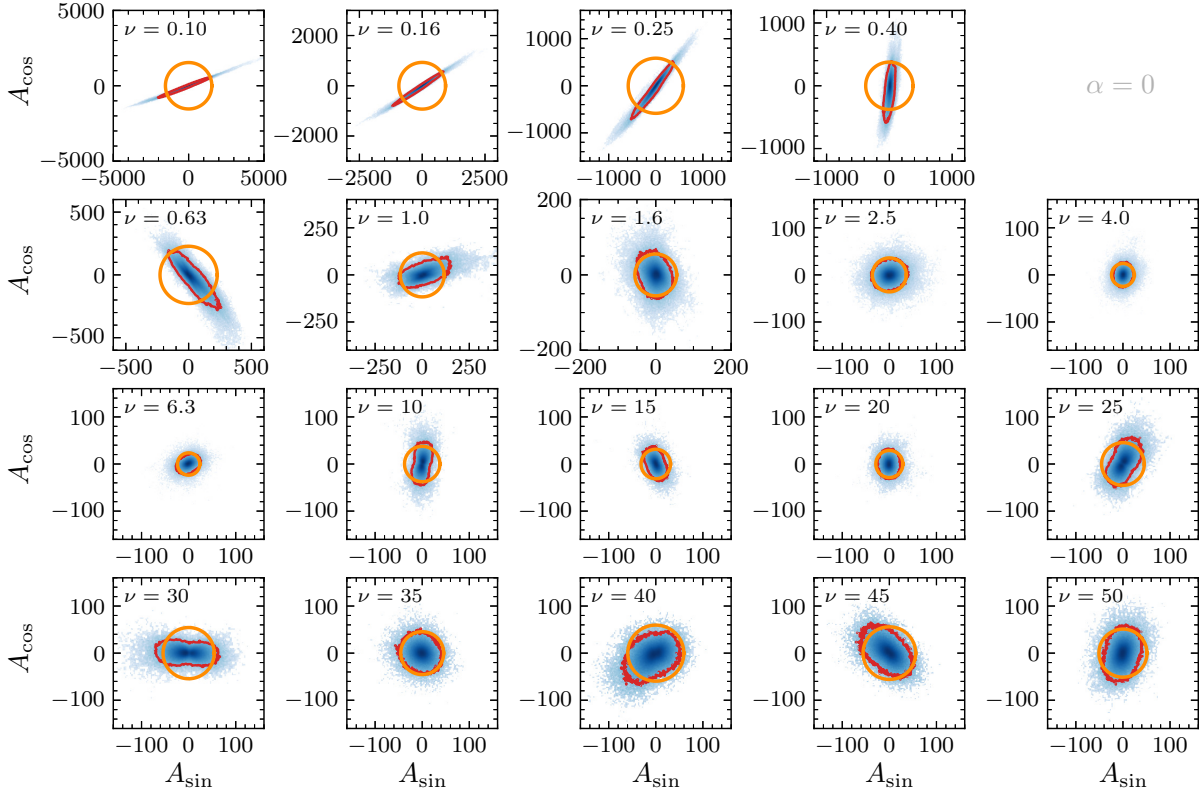


Figure B.6. Posterior distributions of $(A_{\text{cos}}, A_{\text{sin}})$ for a scaling exponent of $\alpha = 0$ and the 19 frequencies ν sampled in our analysis of BOSS DR12 data. The panels show the pixelated posteriors based on the MCMC samples marginalized over the bias and EFT parameters, with their density increasing from light to dark blue. The red contours enclose the pixelated 95% confidence regions around the respective maximum posterior points while the orange circles contain 95% of the total probability around the origin $A f_{\text{NL}}^{\alpha,\nu} = A_{\text{cos}} = A_{\text{sin}} = 0$ as obtained from the shown pixelated posteriors. As explained in §3.5.1, the radius of the latter therefore is the 95% exclusion limit of $A f_{\text{NL}}^{\alpha,\nu}$ which we employ to quantify and compare our results. We see that this method and our inferred limits well represent the posteriors for frequencies $\nu \gtrsim 1$ when the oscillatory amplitudes A_{cos} and A_{sin} are (effectively) uncorrelated, but may underestimate the actual statistical power of BOSS for smaller frequencies for which this correlation is strong.

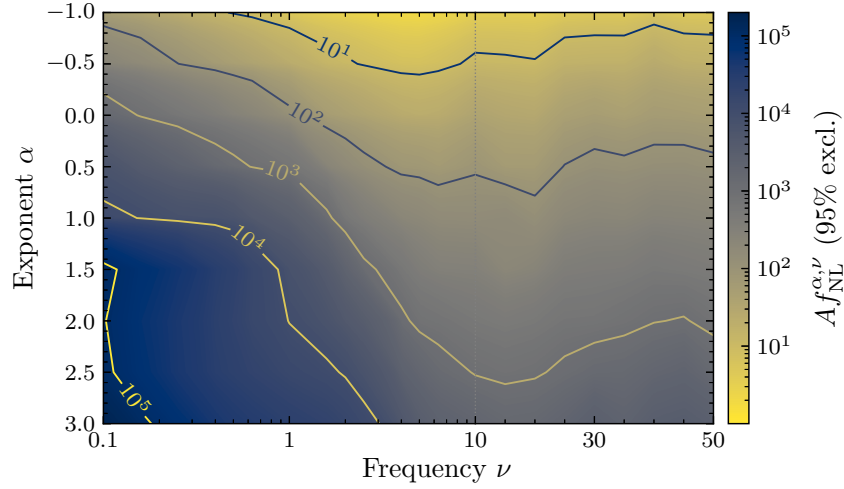


Figure B.7. Exclusion limits at 95% c.l. on the non-Gaussian oscillatory amplitude $Af_{\text{NL}}^{\alpha,\nu}$ inferred from the BOSS DR12 dataset as in Fig. 3.11, but shown as a function of both the scaling exponent α and oscillatory frequency ν of the non-Gaussian signature. The axes for α and ν , and the colorbar for $Af_{\text{NL}}^{\alpha,\nu}$ are displayed on a linear, logarithmic-linear and logarithmic scale, respectively. We also show the contours of constant exclusion limit for every decade in the size of the amplitude. We can clearly observe how the constraining power of BOSS decreases with increasing α while larger oscillation frequencies help to distinguish and constrain the PNG amplitude.

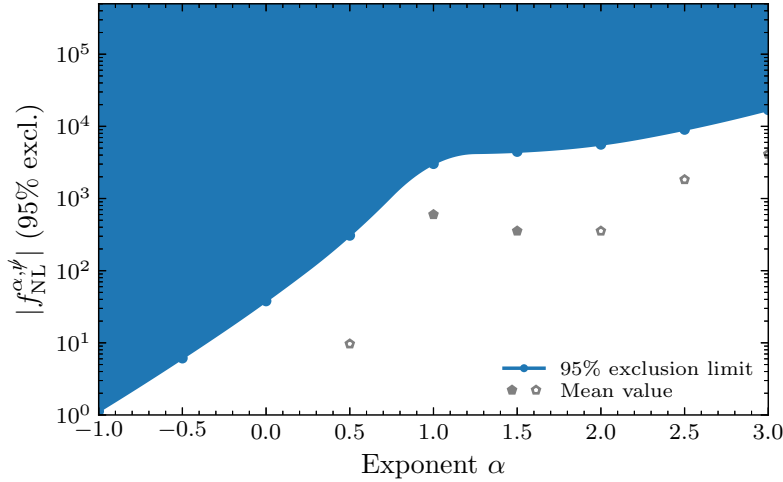


Figure B.8. Exclusion limit at 95% c.l. on the absolute value of the non-Gaussian non-oscillatory amplitude $|f_{\text{NL}}^{\alpha,\psi}|$ inferred from the BOSS DR12 dataset as a function of the scaling exponent α . The filled (unfilled) pentagons display the inferred positive (negative) mean values of $f_{\text{NL}}^{\alpha,\psi}$. This extends the range of scaling exponents from $[0, 2]$ as studied in detail in [193] to the range $[-1, 3]$ considered in this work. For ease of comparison, the displayed range of values for $f_{\text{NL}}^{\alpha,\psi}$ matches that of Fig. 3.11, with the mean values for $\alpha = -1.0, -0.5$ and 0.0 being factors of approximately 100, 110 and 220 smaller than their respective 95% exclusion limits.

to compare the inferred limits for these two classes of primordial non-Gaussianity across the same range of exponents in Fig. 3.12. To be precise, we computed the upper limits on $|f_{\text{NL}}^{\alpha,\psi}|$ at 95% c.l. so that they can be directly compared to the limits on the positive semi-definite oscillatory amplitude $Af_{\text{NL}}^{\alpha,\psi}$. These limits are essentially the same as the upper and lower 2σ constraints on $f_{\text{NL}}^{\alpha,\psi}$ reported in [193] (with the exception of $\alpha = 3$, for which the difference is approximately 25%). Across the extended regions $\alpha < 0$ and $\alpha > 2$, the inferred limits and mean values behave as expected from the more detailed and more densely sampled results of [193], with no indication of qualitatively new or unexpected behavior. This in particular includes a large correlation length in the scaling exponent α and no detection of a nonzero PNG signal.

Bibliography

- [1] N. Aghanim *et al.* (Planck Collaboration), “Planck 2018 Results. VI. Cosmological Parameters”, *Astron. Astrophys.* **641** (2020) A6, arXiv:1807.06209 [astro-ph.CO].
- [2] A. A. Starobinsky, “A New Type of Isotropic Cosmological Models Without Singularity”, *Phys. Lett. B* **91** (1980) 99–102.
- [3] A. H. Guth, “The Inflationary Universe: A Possible Solution to the Horizon and Flatness Problems”, *Phys. Rev. D* **23** (1981) 347–356.
- [4] A. D. Linde, “A New Inflationary Universe Scenario: A Possible Solution of the Horizon, Flatness, Homogeneity, Isotropy and Primordial Monopole Problems”, *Phys. Lett. B* **108** (1982) 389–393.
- [5] A. Albrecht and P. J. Steinhardt, “Cosmology for Grand Unified Theories with Radiatively Induced Symmetry Breaking”, *Phys. Rev. Lett.* **48** (1982) 1220–1223.
- [6] L. Kofman, A. D. Linde, and A. A. Starobinsky, “Towards the theory of reheating after inflation”, *Phys. Rev. D* **56** (1997) 3258–3295, arXiv:hep-ph/9704452.
- [7] V. F. Mukhanov and G. V. Chibisov, “Quantum Fluctuations and a Nonsingular Universe”, *JETP Lett.* **33** (1981) 532–535.
- [8] X. Chen, “Primordial Non-Gaussianities from Inflation Models”, *Adv. Astron.* **2010** (2010) 638979, arXiv:1002.1416 [astro-ph.CO].
- [9] P. D. Meerburg *et al.*, “Primordial Non-Gaussianity”, *Bull. Am. Astron. Soc.* **51** (2019) 107, arXiv:1903.04409 [astro-ph.CO].
- [10] N. Dalal, O. Doré, D. Huterer, and A. Shirokov, “The Imprints of Primordial Non-Gaussianities on Large-Scale Structure: Scale-Dependent Bias and Abundance of Virialized Objects”, *Phys. Rev. D* **77** (2008) 123514, arXiv:0710.4560 [astro-ph].
- [11] X. Chen and Y. Wang, “Large Non-Gaussianities with Intermediate Shapes from Quasi-Single-Field Inflation”, *Phys. Rev. D* **81** (2010) 063511, arXiv:0909.0496

- [astro-ph.CO].
- [12] X. Chen and Y. Wang, “Quasi-Single-Field Inflation and Non-Gaussianities”, *JCAP* **04** (2010) 027, arXiv:0911.3380 [hep-th].
 - [13] D. Baumann and D. Green, “Signatures of Supersymmetry from the Early Universe”, *Phys. Rev. D* **85** (2012) 103520, arXiv:1109.0292 [hep-th].
 - [14] N. Arkani-Hamed and J. Maldacena, “Cosmological Collider Physics”, arXiv:1503.08043 [hep-th].
 - [15] F. Bernardeau, S. Colombi, E. Gaztanaga, and R. Scoccimarro, “Large-Scale Structure of the Universe and Cosmological Perturbation Theory”, *Phys. Rept.* **367** (2002) 1, arXiv:astro-ph/0112551.
 - [16] Y. Akrami *et al.* (Planck Collaboration), “Planck 2018 Results. IX. Constraints on Primordial Non-Gaussianity”, *Astron. Astrophys.* **641** (2020) A9, arXiv:1905.05697 [astro-ph.CO].
 - [17] A. Achúcarro *et al.*, “Inflation: Theory and Observations”, arXiv:2203.08128 [astro-ph.CO].
 - [18] F. Beutler, M. Biagetti, D. Green, A. Slosar, and B. Wallisch, “Primordial Features from Linear to Nonlinear Scales”, *Phys. Rev. Res.* **1** (2019) 033209, arXiv:1906.08758 [astro-ph.CO].
 - [19] P. Creminelli and M. Zaldarriaga, “Single-Field Consistency Relation for the Three-Point Function”, *JCAP* **10** (2004) 006, arXiv:astro-ph/0407059.
 - [20] P. Creminelli, J. Noreña, M. Simonović, and F. Vernizzi, “Single-Field Consistency Relations of Large-Scale Structure”, *JCAP* **12** (2013) 025, arXiv:1309.3557 [astro-ph.CO].
 - [21] P. Creminelli, J. Gleyzes, M. Simonović, and F. Vernizzi, “Single-Field Consistency Relations of Large-Scale Structure. Part II: Resummation and Redshift Space”, *JCAP* **02** (2014) 051, arXiv:1311.0290 [astro-ph.CO].
 - [22] P. Creminelli, J. Gleyzes, L. Hui, M. Simonović, and F. Vernizzi, “Single-Field Consistency Relations of Large-Scale Structure. Part III: Test of the Equivalence Principle”, *JCAP* **06** (2014) 009, arXiv:1312.6074 [astro-ph.CO].
 - [23] D. Green, Y. Huang, C.-H. Shen, and D. Baumann, “Positivity from Cosmological Correlators”, *JHEP* **04** (2024) 034, arXiv:2310.02490 [hep-th].

- [24] M. McAneny and A. Ridgway, “New Shapes of Primordial Non-Gaussianity from Quasi-Single-Field Inflation with Multiple Isocurvatons”, *Phys. Rev. D* **100** (2019) 043534, arXiv:1903.11607 [astro-ph.CO].
- [25] V. Desjacques, D. Jeong, and F. Schmidt, “Large-Scale Galaxy Bias”, *Phys. Rept.* **733** (2018) 1, arXiv:1611.09787 [astro-ph.CO].
- [26] T. Baldauf, M. Mirbabayi, M. Simonović, and M. Zaldarriaga, “LSS Constraints with Controlled Theoretical Uncertainties”, arXiv:1602.00674 [astro-ph.CO].
- [27] G. Cabass, M. Ivanov, M. Lewandowski, M. Mirbabayi, and M. Simonović, “Snowmass White Paper: Effective Field Theories in Cosmology”, *Phys. Dark Univ.* **40** (2023) 101193, arXiv:2203.08232 [astro-ph.CO].
- [28] G. Cabass, M. Ivanov, O. Philcox, M. Simonović, and M. Zaldarriaga, “Constraining Single-Field Inflation with MegaMapper”, *Phys. Lett. B* **841** (2023) 137912, arXiv:2211.14899 [astro-ph.CO].
- [29] D. Green *et al.*, “Snowmass Theory Frontier: Astrophysics and Cosmology”, arXiv:2209.06854 [hep-ph].
- [30] D. Baumann, “Inflation”, in *Theoretical Advanced Study Institute in Elementary Particle Physics: Physics of the Large and the Small*, pp. 523–686. 2011. arXiv:0907.5424 [hep-th].
- [31] D. H. Weinberg, M. J. Mortonson, D. J. Eisenstein, C. Hirata, A. G. Riess, and E. Rozo, “Observational Probes of Cosmic Acceleration”, *Phys. Rept.* **530** (2013) 87–255, arXiv:1201.2434 [astro-ph.CO].
- [32] R. Holman and A. J. Tolley, “Enhanced Non-Gaussianity from Excited Initial States”, *JCAP* **05** (2008) 001, arXiv:0710.1302 [hep-th].
- [33] X. Chen, M.-x. Huang, S. Kachru, and G. Shiu, “Observational signatures and non-Gaussianities of general single field inflation”, *JCAP* **01** (2007) 002, arXiv:hep-th/0605045.
- [34] J. Maldacena, “Non-Gaussian Features of Primordial Fluctuations in Single-Field Inflationary Models”, *JHEP* **05** (2003) 013, arXiv:astro-ph/0210603.
- [35] M. Sasaki and E. D. Stewart, “A General analytic formula for the spectral index of the density perturbations produced during inflation”, *Prog. Theor. Phys.* **95** (1996) 71–78, arXiv:astro-ph/9507001.
- [36] D. H. Lyth and Y. Rodriguez, “The Inflationary prediction for primordial

- non-Gaussianity”, *Phys. Rev. Lett.* **95** (2005) 121302, arXiv:astro-ph/0504045.
- [37] E. Pajer, F. Schmidt, and M. Zaldarriaga, “The Observed Squeezed Limit of Cosmological Three-Point Functions”, *Phys. Rev. D* **88** (2013) 083502, arXiv:1305.0824 [astro-ph.CO].
 - [38] P. Creminelli, J. Noreña, and M. Simonović, “Conformal Consistency Relations for Single-Field Inflation”, *JCAP* **07** (2012) 052, arXiv:1203.4595 [hep-th].
 - [39] V. Assassi, D. Baumann, and F. Schmidt, “Galaxy Bias and Primordial Non-Gaussianity”, *JCAP* **12** (2015) 043, arXiv:1510.03723 [astro-ph.CO].
 - [40] J. Gleyzes, R. de Putter, D. Green, and O. Doré, “Biasing and the Search for Primordial Non-Gaussianity Beyond the Local Type”, *JCAP* **04** (2017) 002, arXiv:1612.06366 [astro-ph.CO].
 - [41] S. Matarrese and L. Verde, “The Effect of Primordial Non-Gaussianity on Halo Bias”, *Astrophys. J. Lett.* **677** (2008) L77, arXiv:0801.4826 [astro-ph].
 - [42] A. Slosar, C. Hirata, U. Seljak, S. Ho, and N. Padmanabhan, “Constraints on Local Primordial Non-Gaussianity from Large-Scale Structure”, *JCAP* **08** (2008) 031, arXiv:0805.3580 [astro-ph].
 - [43] N. Kaiser, “Clustering in Real Space and in Redshift Space”, *Mon. Not. Roy. Astron. Soc.* **227** (1987) 1.
 - [44] C. Alcock and B. Paczynski, “An Evolution-Free Test for Non-Zero Cosmological Constant”, *Nature* **281** (1979) 358.
 - [45] D. J. Eisenstein, H.-j. Seo, and M. J. White, “On the Robustness of the Acoustic Scale in the Low-Redshift Clustering of Matter”, *Astrophys. J.* **664** (2007) 660, arXiv:astro-ph/0604361.
 - [46] D. Eisenstein, H.-j. Seo, E. Sirko, and D. Spergel, “Improving Cosmological Distance Measurements by Reconstruction of the Baryon Acoustic Peak”, *Astrophys. J.* **664** (2007) 675, arXiv:astro-ph/0604362.
 - [47] M. Crocce and R. Scoccimarro, “Nonlinear Evolution of Baryon Acoustic Oscillations”, *Phys. Rev. D* **77** (2008) 023533, arXiv:0704.2783 [astro-ph].
 - [48] C. Chang *et al.*, “Report of the Topical Group on Cosmic Frontier 5 Dark Energy and Cosmic Acceleration: Cosmic Dawn and Before for Snowmass 2021”, arXiv:2209.08265 [hep-ex].

- [49] D. Green, “TASI Lectures on Cosmic Signals of Fundamental Physics”, arXiv:2212.08685 [hep-ph].
- [50] R. Flauger, V. Gorbenko, A. Joyce, L. McAllister, G. Shiu, and E. Silverstein, “Snowmass White Paper: Cosmology at the Theory Frontier”, arXiv:2203.07629 [hep-th].
- [51] M. Biagetti, “The Hunt for Primordial Interactions in the Large-Scale Structure of the Universe”, *Galaxies* **7** (2019) 71, arXiv:1906.12244 [astro-ph.CO].
- [52] G. Cabass, E. Pajer, and F. Schmidt, “How Gaussian Can Our Universe Be?”, *JCAP* **01** (2017) 003, arXiv:1612.00033 [hep-th].
- [53] V. Assassi, D. Baumann, D. Green, and L. McAllister, “Planck-Suppressed Operators”, *JCAP* **01** (2014) 033, arXiv:1304.5226 [hep-th].
- [54] D. Schlegel *et al.* (DESI Collaboration), “A Spectroscopic Road Map for Cosmic Frontier: DESI, DESI-II, Stage-5”, arXiv:2209.03585 [astro-ph.CO].
- [55] D. Schlegel *et al.*, “The MegaMapper: A Stage-5 Spectroscopic Instrument Concept for the Study of Inflation and Dark Energy”, arXiv:2209.04322 [astro-ph.IM].
- [56] J. Annis, J. Newman, and A. Slosar, “Snowmass2021 Cosmic Frontier: Report of the CF04 Topical Group on Dark Energy and Cosmic Acceleration in the Modern Universe”, arXiv:2209.08049 [astro-ph.CO].
- [57] M. Alvarez *et al.*, “Testing Inflation with Large-Scale Structure: Connecting Hopes with Reality”, arXiv:1412.4671 [astro-ph.CO].
- [58] D. Lyth, C. Ungarelli, and D. Wands, “The Primordial Density Perturbation in the Curvaton Scenario”, *Phys. Rev. D* **67** (2003) 023503, arXiv:astro-ph/0208055.
- [59] M. Zaldarriaga, “Non-Gaussianities in Models with a Varying Inflaton Decay Rate”, *Phys. Rev. D* **69** (2004) 043508, arXiv:astro-ph/0306006.
- [60] M. Sasaki, J. Valiviita, and D. Wands, “Non-Gaussianity of the Primordial Perturbation in the Curvaton Model”, *Phys. Rev. D* **74** (2006) 103003, arXiv:astro-ph/0607627.
- [61] A. Ross *et al.* (BOSS Collaboration), “The Clustering of Galaxies in SDSS-III DR9 Baryon Oscillation Spectroscopic Survey: Constraints on Primordial Non-Gaussianity”, *Mon. Not. Roy. Astron. Soc.* **428** (2013) 1116, arXiv:1208.1491 [astro-ph.CO].
- [62] B. Leistedt, H. Peiris, and N. Roth, “Constraints on Primordial Non-Gaussianity from 800000 Photometric Quasars”, *Phys. Rev. Lett.* **113** (2014) 221301, arXiv:1405.4315 [astro-ph.CO].

- [63] E. Castorina *et al.* (eBOSS Collaboration), “Redshift-Weighted Constraints on Primordial Non-Gaussianity from the Clustering of the eBOSS DR14 Quasars in Fourier Space”, *JCAP* **09** (2019) 010, arXiv:1904.08859 [astro-ph.CO].
- [64] E.-M. Mueller *et al.* (eBOSS Collaboration), “Primordial Non-Gaussianity from the Completed SDSS-IV Extended Baryon Oscillation Spectroscopic Survey II: Measurements in Fourier Space with Optimal Weights”, *Mon. Not. Roy. Astron. Soc.* **514** (2022) 3396, arXiv:2106.13725 [astro-ph.CO].
- [65] G. D’Amico, M. Lewandowski, L. Senatore, and P. Zhang, “Limits on Primordial Non-Gaussianities from BOSS Galaxy Clustering Data”, arXiv:2201.11518 [astro-ph.CO].
- [66] G. Cabass, M. Ivanov, O. Philcox, M. Simonović, and M. Zaldarriaga, “Constraints on Multi-Field Inflation from the BOSS Galaxy Survey”, *Phys. Rev. D* **106** (2022) 043506, arXiv:2204.01781 [astro-ph.CO].
- [67] M. Rezaie *et al.*, “Local Primordial Non-Gaussianity from the Large-Scale Clustering of Photometric DESI Luminous Red Galaxies”, arXiv:2307.01753 [astro-ph.CO].
- [68] A. Krolewski *et al.* (DESI Collaboration), “Constraining Primordial Non-Gaussianity from DESI Quasar Targets and Planck CMB Lensing”, *JCAP* **03** (2024) 021, arXiv:2305.07650 [astro-ph.CO].
- [69] M. Cagliari, E. Castorina, M. Bonici, and D. Bianchi, “Optimal Constraints on Primordial Non-Gaussianity with the eBOSS DR16 Quasars in Fourier Space”, arXiv:2309.15814 [astro-ph.CO].
- [70] E. Komatsu and D. Spergel, “Acoustic Signatures in the Primary Microwave Background Bispectrum”, *Phys. Rev. D* **63** (2001) 063002, arXiv:astro-ph/0005036.
- [71] O. Doré *et al.* (SPHEREx Collaboration), “Cosmology with the SPHEREX All-Sky Spectral Survey”, arXiv:1412.4872 [astro-ph.CO].
- [72] R. de Putter, J. Gleyzes, and O. Doré, “Next Non-Gaussianity Frontier: What Can a Measurement with $\sigma(f_{\text{NL}}) \lesssim 1$ Tell Us about Multi-Field Inflation?”, *Phys. Rev. D* **95** (2017) 123507, arXiv:1612.05248 [astro-ph.CO].
- [73] R. de Putter and O. Doré, “Designing an Inflation Galaxy Survey: How to Measure $\sigma(f_{\text{NL}}) \sim 1$ using Scale-Dependent Galaxy Bias”, *Phys. Rev. D* **95** (2017) 123513, arXiv:1412.3854 [astro-ph.CO].
- [74] E.-M. Mueller, W. Percival, and R. Ruggeri, “Optimizing Primordial Non-Gaussianity Measurements from Galaxy Surveys”, *Mon. Not. Roy. Astron. Soc.* **485** (2019) 4160,

arXiv:1702.05088 [astro-ph.CO].

- [75] E. Sefusatti, J. Fergusson, X. Chen, and E. P. S. Shellard, “Effects and Detectability of Quasi-Single-Field Inflation in the Large-Scale Structure and Cosmic Microwave Background”, *JCAP* **08** (2012) 033, arXiv:1204.6318 [astro-ph.CO].
- [76] J. Noreña, L. Verde, G. Barenboim, and C. Bosch, “Prospects for Constraining the Shape of Non-Gaussianity with the Scale-Dependent Bias”, *JCAP* **08** (2012) 019, arXiv:1204.6324 [astro-ph.CO].
- [77] S. Alam *et al.* (BOSS Collaboration), “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** (2017) 2617, arXiv:1607.03155 [astro-ph.CO].
- [78] N. Agarwal, S. Ho, and S. Shandera, “Constraining the Initial Conditions of the Universe using Large-Scale Structure”, *JCAP* **02** (2014) 038, arXiv:1311.2606 [astro-ph.CO].
- [79] M. Schmittfull, M. Simonović, V. Assassi, and M. Zaldarriaga, “Modeling Biased Tracers at the Field Level”, *Phys. Rev. D* **100** (2019) 043514, arXiv:1811.10640 [astro-ph.CO].
- [80] P. McDonald and A. Roy, “Clustering of Dark Matter Tracers: Generalizing Bias for the Coming Era of Precision LSS”, *JCAP* **08** (2009) 020, arXiv:0902.0991 [astro-ph.CO].
- [81] T. Lazeyras, C. Wagner, T. Baldauf, and F. Schmidt, “Precision Measurement of the Local Bias of Dark Matter Halos”, *JCAP* **02** (2016) 018, arXiv:1511.01096 [astro-ph.CO].
- [82] A. Lewis, A. Challinor, and A. Lasenby, “Efficient Computation of CMB Anisotropies in Closed FRW Models”, *Astrophys. J.* **538** (2000) 473, arXiv:astro-ph/9911177.
- [83] H.-J. Seo and D. Eisenstein, “Probing Dark Energy with Baryon Acoustic Oscillations from Future Large Galaxy-Redshift Surveys”, *Astrophys. J.* **598** (2003) 720, arXiv:astro-ph/0307460.
- [84] H. Zhan and L. Knox, “Baryon Oscillations and Consistency Tests for Photometrically Determined Redshifts of Very Faint Galaxies”, *Astrophys. J.* **644** (2006) 663, arXiv:astro-ph/0509260.
- [85] M. LoVerde, A. Miller, S. Shandera, and L. Verde, “Effects of Scale-Dependent Non-Gaussianity on Cosmological Structures”, *JCAP* **04** (2008) 014, arXiv:0711.4126

[astro-ph].

- [86] D. Baumann, S. Ferraro, D. Green, and K. Smith, “Stochastic Bias from Non-Gaussian Initial Conditions”, *JCAP* **05** (2013) 001, arXiv:1209.2173 [astro-ph.CO].
- [87] R. de Putter, O. Doré, and D. Green, “Is There Scale-Dependent Bias in Single-Field Inflation?”, *JCAP* **10** (2015) 024, arXiv:1504.05935 [astro-ph.CO].
- [88] P. Creminelli, A. Perko, L. Senatore, M. Simonović, and G. Trevisan, “The Physical Squeezed Limit: Consistency Relations at Order q^2 ”, *JCAP* **11** (2013) 015, arXiv:1307.0503 [astro-ph.CO].
- [89] A. Testa and M. Wise, “Impact of Transforming to Conformal Fermi Coordinates on Quasi-Single-Field Non-Gaussianity”, *Phys. Rev. D* **102** (2020) 023533, arXiv:2004.06126 [astro-ph.CO].
- [90] D. Green, M. Lewandowski, L. Senatore, E. Silverstein, and M. Zaldarriaga, “Anomalous Dimensions and Non-Gaussianity”, *JHEP* **10** (2013) 171, arXiv:1301.2630 [hep-th].
- [91] H. An, M. Wise, and Z. Zhang, “De Sitter Quantum Loops as the Origin of Primordial Non-Gaussianities”, *Phys. Rev. D* **99** (2019) 056007, arXiv:1806.05194 [hep-ph].
- [92] T. Noumi, M. Yamaguchi, and D. Yokoyama, “Effective Field Theory Approach to Quasi-Single-Field Inflation and Effects of Heavy Fields”, *JHEP* **06** (2013) 051, arXiv:1211.1624 [hep-th].
- [93] R. Flauger, M. Mirbabayi, L. Senatore, and E. Silverstein, “Productive Interactions: Heavy Particles and Non-Gaussianity”, *JCAP* **10** (2017) 058, arXiv:1606.00513 [hep-th].
- [94] F. Schmidt and M. Kamionkowski, “Halo Clustering with Non-Local Non-Gaussianity”, *Phys. Rev. D* **82** (2010) 103002, arXiv:1008.0638 [astro-ph.CO].
- [95] V. Desjacques, D. Jeong, and F. Schmidt, “Accurate Predictions for the Scale-Dependent Galaxy Bias from Primordial Non-Gaussianity”, *Phys. Rev. D* **84** (2011) 061301, arXiv:1105.3476 [astro-ph.CO].
- [96] V. Desjacques, D. Jeong, and F. Schmidt, “Non-Gaussian Halo Bias Re-Examined: Mass-Dependent Amplitude from the Peak-Background Split and Thresholding”, *Phys. Rev. D* **84** (2011) 063512, arXiv:1105.3628 [astro-ph.CO].
- [97] T. Giannantonio, C. Porciani, J. Carron, A. Amara, and A. Pillepich, “Constraining Primordial Non-Gaussianity with Future Galaxy Surveys”, *Mon. Not. Roy. Astron. Soc.* **422** (2012) 2854, arXiv:1109.0958 [astro-ph.CO].

- [98] A. Barreira, “On the Impact of Galaxy Bias Uncertainties on Primordial Non-Gaussianity Constraints”, *JCAP* **12** (2020) 031, arXiv:2009.06622 [astro-ph.CO].
- [99] A. Barreira, “Predictions for Local PNG Bias in the Galaxy Power Spectrum and Bispectrum and the Consequences for f_{NL} Constraints”, *JCAP* **01** (2022) 033, arXiv:2107.06887 [astro-ph.CO].
- [100] A. Barreira, T. Lazeyras, and F. Schmidt, “Galaxy Bias from Forward Models: Linear and Second-Order Bias of IllustrisTNG Galaxies”, *JCAP* **08** (2021) 029, arXiv:2105.02876 [astro-ph.CO].
- [101] A. Barreira, “Can We Actually Constrain f_{NL} using the Scale-Dependent Bias Effect? An Illustration of the Impact of Galaxy Bias Uncertainties using the BOSS DR12 Galaxy Power Spectrum”, *JCAP* **11** (2022) 013, arXiv:2205.05673 [astro-ph.CO].
- [102] A. Chudaykin, M. Ivanov, and M. Simonović, “Optimizing Large-Scale Structure Data Analysis with the Theoretical Error Likelihood”, *Phys. Rev. D* **103** (2021) 043525, arXiv:2009.10724 [astro-ph.CO].
- [103] M. Moreira, F. Andrade-Oliveira, X. Fang, H.-J. Huang, E. Krause, V. Miranda, R. Rosenfeld, and M. Simonović, “Mitigating Baryonic Effects with a Theoretical Error Covariance”, *Mon. Not. Roy. Astron. Soc.* **507** (2021) 5592, arXiv:2104.01397 [astro-ph.CO].
- [104] A. Aghamousa *et al.* (DESI Collaboration), “The DESI Experiment Part I: Science, Targeting and Survey Design”, arXiv:1611.00036 [astro-ph.IM].
- [105] R. Laureijs *et al.* (Euclid Collaboration), “Euclid Definition Study Report”, arXiv:1110.3193 [astro-ph.CO].
- [106] R. Scaramella *et al.* (Euclid Collaboration), “Euclid Preparation – I. The Euclid Wide Survey”, *Astron. Astrophys.* **662** (2022) A112, arXiv:2108.01201 [astro-ph.CO].
- [107] P. Abell *et al.* (LSST Science & LSST Project Collaborations), “LSST Science Book, Version 2.0”, arXiv:0912.0201 [astro-ph.IM].
- [108] S. Ferraro *et al.*, “Inflation and Dark Energy from Spectroscopy at $z > 2$ ”, *Bull. Am. Astron. Soc.* **51** (2019) 72, arXiv:1903.09208 [astro-ph.CO].
- [109] K. Abazajian *et al.* (CMB-S4 Collaboration), “CMB-S4 Science Case, Reference Design, and Project Plan”, arXiv:1907.04473 [astro-ph.IM].
- [110] K. Abazajian *et al.* (CMB-S4 Collaboration), “Snowmass 2021 CMB-S4 White Paper”, arXiv:2203.08024 [astro-ph.CO].

- [111] D. Baumann, D. Green, and B. Wallisch, “Searching for Light Relics with Large-Scale Structure”, *JCAP* **08** (2018) 029, arXiv:1712.08067 [astro-ph.CO].
- [112] R. Allison, P. Caucal, E. Calabrese, J. Dunkley, and T. Louis, “Towards a Cosmological Neutrino Mass Detection”, *Phys. Rev. D* **92** (2015) 123535, arXiv:1509.07471 [astro-ph.CO].
- [113] U. Seljak, “Extracting Primordial Non-Gaussianity Without Cosmic Variance”, *Phys. Rev. Lett.* **102** (2009) 021302, arXiv:0807.1770 [astro-ph].
- [114] C. Orban and D. Weinberg, “Self-Similar Bumps and Wiggles: Isolating the Evolution of the BAO Peak with Power-Law Initial Conditions”, *Phys. Rev. D* **84** (2011) 063501, arXiv:1101.1523 [astro-ph.CO].
- [115] S. Camera, M. Santos, and R. Maartens, “Probing Primordial Non-Gaussianity with SKA Galaxy Redshift Surveys: A Fully Relativistic Analysis”, *Mon. Not. Roy. Astron. Soc.* **448** (2015) 1035, arXiv:1409.8286 [astro-ph.CO].
- [116] Y.-C. Li and Y.-Z. Ma, “Constraints on Primordial Non-Gaussianity from Future HI Intensity Mapping Experiments”, *Phys. Rev. D* **96** (2017) 063525, arXiv:1701.00221 [astro-ph.CO].
- [117] A. Moradinezhad Dizgah and G. Keating, “Line Intensity Mapping with [CII] and CO(1-0) as Probes of Primordial Non-Gaussianity”, *Astrophys. J.* **872** (2019) 126, arXiv:1810.02850 [astro-ph.CO].
- [118] R. Ansari *et al.* (Cosmic Visions 21 cm Collaboration), “Inflation and Early Dark Energy with a Stage-II Hydrogen Intensity Mapping Experiment”, arXiv:1810.09572 [astro-ph.CO].
- [119] A. Slosar *et al.* (PUMA Collaboration), “Packed Ultra-wideband Mapping Array (PUMA): A Radio Telescope for Cosmology and Transients”, *Bull. Am. Astron. Soc.* **51** (2019) 53, arXiv:1907.12559 [astro-ph.IM].
- [120] D. Karagiannis, J. Fonseca, R. Maartens, and S. Camera, “Probing Primordial Non-Gaussianity with the Power Spectrum and Bispectrum of Future 21 cm Intensity Maps”, *Phys. Dark Univ.* **32** (2021) 100821, arXiv:2010.07034 [astro-ph.CO].
- [121] A. Pullen and C. Hirata, “Systematic Effects in Large-Scale Angular Power Spectra of Photometric Quasars and Implications for Constraining Primordial Non-Gaussianity”, *Publ. Astron. Soc. Pac.* **125** (2013) 705, arXiv:1212.4500 [astro-ph.CO].
- [122] D. Babich, P. Creminelli, and M. Zaldarriaga, “The Shape of Non-Gaussianities”, *JCAP* **08** (2004) 009, arXiv:astro-ph/0405356.

- [123] O. Philcox and M. Ivanov, “BOSS DR12 Full-Shape Cosmology: Λ CDM Constraints from the Large-Scale Galaxy Power Spectrum and Bispectrum Monopole”, *Phys. Rev. D* **105** (2022) 043517, arXiv:2112.04515 [astro-ph.CO].
- [124] G. Cabass, M. Ivanov, O. Philcox, M. Simonović, and M. Zaldarriaga, “Constraints on Single-Field Inflation from the BOSS Galaxy Survey”, *Phys. Rev. Lett.* **129** (2022) 021301, arXiv:2201.07238 [astro-ph.CO].
- [125] O. Philcox, “Cosmology Without Window Functions: Quadratic Estimators for the Galaxy Power Spectrum”, *Phys. Rev. D* **103** (2021) 103504, arXiv:2012.09389 [astro-ph.CO].
- [126] F.-S. Kitaura *et al.*, “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** (2016) 4156, arXiv:1509.06400 [astro-ph.CO].
- [127] A. Chudaykin, M. Ivanov, O. Philcox, and M. Simonović, “Nonlinear Perturbation Theory Extension of the Boltzmann Code CLASS”, *Phys. Rev. D* **102** (2020) 063533, arXiv:2004.10607 [astro-ph.CO].
- [128] D. Blas, J. Lesgourgues, and T. Tram, “The Cosmic Linear Anisotropy Solving System (CLASS) II: Approximation Schemes”, *JCAP* **07** (2011) 034, arXiv:1104.2933 [astro-ph.CO].
- [129] M. Ivanov, M. Simonović, and M. Zaldarriaga, “Cosmological Parameters from the BOSS Galaxy Power Spectrum”, *JCAP* **05** (2020) 042, arXiv:1909.05277 [astro-ph.CO].
- [130] G. D’Amico, J. Gleyzes, N. Kokron, K. Markovic, L. Senatore, P. Zhang, F. Beutler, and H. Gil-Marín, “The Cosmological Analysis of the SDSS/BOSS Data from the Effective Field Theory of Large-Scale Structure”, *JCAP* **05** (2020) 005, arXiv:1909.05271 [astro-ph.CO].
- [131] D. Baumann, A. Nicolis, L. Senatore, and M. Zaldarriaga, “Cosmological Nonlinearities as an Effective Fluid”, *JCAP* **07** (2012) 051, arXiv:1004.2488 [astro-ph.CO].
- [132] J. J. Carrasco, M. Hertzberg, and L. Senatore, “The Effective Field Theory of Cosmological Large-Scale Structure”, *JHEP* **09** (2012) 082, arXiv:1206.2926 [astro-ph.CO].
- [133] E. Pajer and M. Zaldarriaga, “On the Renormalization of the Effective Field Theory of Large-Scale Structure”, *JCAP* **08** (2013) 037, arXiv:1301.7182 [astro-ph.CO].
- [134] J. J. Carrasco, S. Foreman, D. Green, and L. Senatore, “The Effective Field Theory of

- Large-Scale Structure at Two Loops”, *JCAP* **07** (2014) 057, arXiv:1310.0464 [astro-ph.CO].
- [135] M. Ivanov, “Effective Field Theory for Large-Scale Structure”, arXiv:2212.08488 [astro-ph.CO].
- [136] B. Audren, J. Lesgourgues, K. Benabed, and S. Prunet, “Conservative Constraints on Early Cosmology: An Illustration of the MontePython Cosmological Parameter Inference Code”, *JCAP* **02** (2013) 001, arXiv:1210.7183 [astro-ph.CO].
- [137] T. Brinckmann and J. Lesgourgues, “MontePython 3: Boosted MCMC Sampler and Other Features”, *Phys. Dark Univ.* **24** (2019) 100260, arXiv:1804.07261 [astro-ph.CO].
- [138] V. Assassi, D. Baumann, D. Green, and M. Zaldarriaga, “Renormalized Halo Bias”, *JCAP* **08** (2014) 056, arXiv:1402.5916 [astro-ph.CO].
- [139] L. Senatore, “Bias in the Effective Field Theory of Large-Scale Structure”, *JCAP* **11** (2015) 007, arXiv:1406.7843 [astro-ph.CO].
- [140] O. Philcox, M. Ivanov, M. Zaldarriaga, M. Simonović, and M. Schmittfull, “Fewer Mocks and Less Noise: Reducing the Dimensionality of Cosmological Observables with Subspace Projections”, *Phys. Rev. D* **103** (2021) 043508, arXiv:2009.03311 [astro-ph.CO].
- [141] D. Green, J. Han, and B. Wallisch, *work in progress*.
- [142] P. A. R. Ade *et al.* (Planck Collaboration), “Planck 2013 Results. XXIV. Constraints on Primordial Non-Gaussianity”, *Astron. Astrophys.* **571** (2014) A24, arXiv:1303.5084 [astro-ph.CO].
- [143] R. de Putter, “Primordial Physics from Large-Scale Structure Beyond the Power Spectrum”, arXiv:1802.06762 [astro-ph.CO].
- [144] A. Barreira and E. Krause, “Towards Optimal and Robust f_{NL} Constraints with Multi-Tracer Analyses”, *JCAP* **10** (2023) 044, arXiv:2302.09066 [astro-ph.CO].
- [145] D. Karagiannis, R. Maartens, J. Fonseca, S. Camera, and C. Clarkson, “Multi-Tracer Power Spectra and Bispectra: Formalism”, *JCAP* **03** (2024) 034, 2305.04028.
- [146] V. Desjacques, U. Seljak, and I. Iliev, “Scale-Dependent Bias Induced by Local Non-Gaussianity: A Comparison to N -Body Simulations”, *Mon. Not. Roy. Astron. Soc.* **396** (2009) 85, arXiv:0811.2748 [astro-ph].
- [147] A. Pillepich, C. Porciani, and O. Hahn, “Universal Halo Mass Function and

- Scale-Dependent Bias from N -Body Simulations with Non-Gaussian Initial Conditions”, *Mon. Not. Roy. Astron. Soc.* **402** (2010) 191, arXiv:0811.4176 [astro-ph].
- [148] M. Grossi, L. Verde, C. Carbone, K. Dolag, E. Branchini, F. Iannuzzi, S. Matarrese, and L. Moscardini, “Large-Scale Non-Gaussian Mass Function and Halo Bias: Tests on N -Body Simulations”, *Mon. Not. Roy. Astron. Soc.* **398** (2009) 321, arXiv:0902.2013 [astro-ph.CO].
 - [149] B. Reid, L. Verde, K. Dolag, S. Matarrese, and L. Moscardini, “Non-Gaussian Halo Assembly Bias”, *JCAP* **07** (2010) 013, arXiv:1004.1637 [astro-ph.CO].
 - [150] C. Wagner and L. Verde, “ N -Body Simulations with Generic Non-Gaussian Initial Conditions II: Halo Bias”, *JCAP* **03** (2012) 002, arXiv:1102.3229 [astro-ph.CO].
 - [151] N. Hamaus, U. Seljak, and V. Desjacques, “Optimal Constraints on Local Primordial Non-Gaussianity from the Two-Point Statistics of Large-Scale Structure”, *Phys. Rev. D* **84** (2011) 083509, arXiv:1104.2321 [astro-ph.CO].
 - [152] R. Scoccimarro, L. Hui, M. Manera, and K. C. Chan, “Large-Scale Bias and Efficient Generation of Initial Conditions for Non-Local Primordial Non-Gaussianity”, *Phys. Rev. D* **85** (2012) 083002, arXiv:1108.5512 [astro-ph.CO].
 - [153] T. Baldauf, U. Seljak, L. Senatore, and M. Zaldarriaga, “Linear Response to Long-Wavelength Fluctuations using Curvature Simulations”, *JCAP* **09** (2016) 007, arXiv:1511.01465 [astro-ph.CO].
 - [154] M. Biagetti, T. Lazeyras, T. Baldauf, V. Desjacques, and F. Schmidt, “Verifying the Consistency Relation for the Scale-Dependent Bias from Local Primordial Non-Gaussianity”, *Mon. Not. Roy. Astron. Soc.* **468** (2017) 3277, arXiv:1611.04901 [astro-ph.CO].
 - [155] A. Barreira, G. Cabass, F. Schmidt, A. Pillepich, and D. Nelson, “Galaxy Bias and Primordial Non-Gaussianity: Insights from Galaxy Formation Simulations with IllustrisTNG”, *JCAP* **12** (2020) 013, arXiv:2006.09368 [astro-ph.CO].
 - [156] A. Moradinezhad Dizgah and C. Dvorkin, “Scale-Dependent Galaxy Bias from Massive Particles with Spin during Inflation”, *JCAP* **01** (2018) 010, arXiv:1708.06473 [astro-ph.CO].
 - [157] D. Anbajagane, C. Chang, H. Lee, and M. Gatti, “Primordial Non-Gaussianities with Weak Lensing: Information on Nonlinear Scales in the Ulagam Full-Sky Simulations”, *JCAP* **03** (2024) 062, arXiv:2310.02349 [astro-ph.CO].
 - [158] M. Schmittfull and U. Seljak, “Parameter Constraints from Cross-Correlation of

- CMB Lensing with Galaxy Clustering”, *Phys. Rev. D* **97** (2018) 123540, arXiv:1710.09465 [astro-ph.CO].
- [159] M. Münchmeyer, M. Madhavacheril, S. Ferraro, M. Johnson, and K. Smith, “Constraining Local Non-Gaussianities with Kinetic Sunyaev-Zel’dovich Tomography”, *Phys. Rev. D* **100** (2019) 083508, arXiv:1810.13424 [astro-ph.CO].
 - [160] S.-F. Chen, H. Lee, and C. Dvorkin, “Precise and Accurate Cosmology with CMB×LSS Power Spectra and Bispectra”, *JCAP* **05** (2021) 030, arXiv:2103.01229 [astro-ph.CO].
 - [161] M. Münchmeyer and K. Smith, “Higher N -Point Function Data Analysis Techniques for Heavy Particle Production and WMAP Results”, *Phys. Rev. D* **100** (2019) 123511, arXiv:1910.00596 [astro-ph.CO].
 - [162] M. Biagetti, A. Cole, and G. Shiu, “The Persistence of Large-Scale Structures I: Primordial Non-Gaussianity”, *JCAP* **04** (2021) 061, arXiv:2009.04819 [astro-ph.CO].
 - [163] D. Baumann and D. Green, “The Power of Locality: Primordial Non-Gaussianity at the Map Level”, *JCAP* **08** (2022) 061, arXiv:2112.14645 [astro-ph.CO].
 - [164] M. Biagetti, J. Calles, L. Castiblanco, A. Cole, and J. Noreña, “Fisher Forecasts for Primordial Non-Gaussianity from Persistent Homology”, *JCAP* **10** (2022) 002, arXiv:2203.08262 [astro-ph.CO].
 - [165] A. Andrews, J. Jasche, G. Lavaux, and F. Schmidt, “Bayesian Field-Level Inference of Primordial Non-Gaussianity using Next-Generation Galaxy Surveys”, *Mon. Not. Roy. Astron. Soc.* **520** (2023) 5746, arXiv:2203.08838 [astro-ph.CO].
 - [166] G. Jung, D. Karagiannis, M. Liguori, M. Baldi, W. Coulton, D. Jamieson, L. Verde, F. Villaescusa-Navarro, and B. Wandelt, “Quijote-PNG: Quasi-Maximum Likelihood Estimation of Primordial Non-Gaussianity in the Nonlinear Dark Matter Density Field”, *Astrophys. J.* **940** (2022) 71, arXiv:2206.01624 [astro-ph.CO].
 - [167] M. LoVerde, “Spherical Collapse in Λ CDM”, *Phys. Rev. D* **90** (2014) 083518, arXiv:1405.4858 [astro-ph.CO].
 - [168] S. Vagnozzi, T. Brinckmann, M. Archidiacono, K. Freese, M. Gerbino, J. Lesgourgues, and T. Sprenger, “Bias Due to Neutrinos Must Not Uncorrect’d Go”, *JCAP* **09** (2018) 001, arXiv:1807.04672 [astro-ph.CO].
 - [169] C. Shiveshwarkar, T. Brinckmann, M. Loverde, and M. McQuinn, “Post-Inflationary Contamination of Local Primordial Non-Gaussianity in Galaxy Power Spectra”, *Phys.*

- Rev. D* **108** (2023) 103538, arXiv:2306.07517 [astro-ph.CO].
- [170] D. Alonso, P. Bull, P. Ferreira, R. Maartens, and M. Santos, “Ultra-Large-Scale Cosmology in Next-Generation Experiments with Single Tracers”, *Astrophys. J.* **814** (2015) 145, arXiv:1505.07596 [astro-ph.CO].
 - [171] E. Castorina and A. Moradinezhad Dizgah, “Local Primordial Non-Gaussianities and Super-Sample Variance”, *JCAP* **10** (2020) 007, arXiv:2005.14677 [astro-ph.CO].
 - [172] E. Castorina and E. di Dio, “The Observed Galaxy Power Spectrum in General Relativity”, *JCAP* **01** (2022) 061, arXiv:2106.08857 [astro-ph.CO].
 - [173] R. Martinez-Carrillo, J. Hidalgo, K. Malik, and A. Pourtsidou, “Contributions from Primordial Non-Gaussianity and General Relativity to the Galaxy Power Spectrum”, *JCAP* **12** (2021) 025, arXiv:2107.10815 [astro-ph.CO].
 - [174] M. Foglieni, M. Pantiri, E. Di Dio, and E. Castorina, “Large-Scale Limit of the Observed Galaxy Power Spectrum”, *Phys. Rev. Lett.* **131** (2023) 111201, arXiv:2303.03142 [astro-ph.CO].
 - [175] H. Lee, D. Baumann, and G. Pimentel, “Non-Gaussianity as a Particle Detector”, *JHEP* **12** (2016) 040, arXiv:1607.03735 [hep-th].
 - [176] S. Behbahani and D. Green, “Collective Symmetry Breaking and Resonant Non-Gaussianity”, *JCAP* **11** (2012) 056, arXiv:1207.2779 [hep-th].
 - [177] A. Bodas, S. Kumar, and R. Sundrum, “The Scalar Chemical Potential in Cosmological Collider Physics”, *JHEP* **02** (2021) 079, arXiv:2010.04727 [hep-ph].
 - [178] P. D. Meerburg, M. Münchmeyer, J. Muñoz, and X. Chen, “Prospects for Cosmological Collider Physics”, *JCAP* **03** (2017) 050, arXiv:1610.06559 [astro-ph.CO].
 - [179] J. McEwen, X. Fang, C. Hirata, and J. Blazek, “FAST-PT: A Novel Algorithm to Calculate Convolution Integrals in Cosmological Perturbation Theory”, *JCAP* **09** (2016) 015, arXiv:1603.04826 [astro-ph.CO].
 - [180] F. Pérez and B. Granger, “IPython: A System for Interactive Scientific Computing”, *Comput. Sci. Eng.* **9** (2007) 21.
 - [181] J. Hunter, “Matplotlib: A 2D Graphics Environment”, *Comput. Sci. Eng.* **9** (2007) 90.
 - [182] C. Harris *et al.*, “Array Programming with NumPy”, *Nature* **585** (2020) 3572, arXiv:2006.10256 [cs.MS].

- [183] P. Virtanen *et al.*, “SciPy 1.0 – Fundamental Algorithms for Scientific Computing in Python”, *Nat. Methods* **17** (2020) 261, arXiv:1907.10121 [cs.MS].
- [184] V. Assassi, D. Baumann, and D. Green, “On Soft Limits of Inflationary Correlation Functions”, *JCAP* **11** (2012) 047, arXiv:1204.4207 [hep-th].
- [185] X. Chen, Y. Wang, and Z.-Z. Xianyu, “Standard Model Background of the Cosmological Collider”, *Phys. Rev. Lett.* **118** (2017) 261302, arXiv:1610.06597 [hep-th].
- [186] S. Kumar and R. Sundrum, “Cosmological Collider Physics and the Curvaton”, *JHEP* **04** (2020) 077, arXiv:1908.11378 [hep-ph].
- [187] Q. Lu, M. Reece, and Z.-Z. Xianyu, “Missing Scalars at the Cosmological Collider”, *JHEP* **12** (2021) 098, arXiv:2108.11385 [hep-ph].
- [188] P. Chakraborty and J. Stout, “Light Scalars at the Cosmological Collider”, *JHEP* **02** (2024) 021, arXiv:2310.01494 [hep-th].
- [189] P. Chakraborty and J. Stout, “Compact Scalars at the Cosmological Collider”, *JHEP* **03** (2024) 149, arXiv:2311.09219 [hep-th].
- [190] S. Goldstein, O. Philcox, J. C. Hill, and L. Hui, “Intermediate Mass-Range Particles from Small Scales: Nonperturbative Techniques for Cosmological Collider Physics from Large-Scale Structure Surveys”, *Phys. Rev. D* **110** (2024) 083516, arXiv:2407.08731 [astro-ph.CO].
- [191] P. Chakraborty, “Primordial Non-Gaussianity from Light Compact Scalars”, *JHEP* **11** (2025) 023, arXiv:2501.07672 [hep-th].
- [192] S. Goldstein, O. Philcox, E. Fondi, and W. Coulton, “Wonderings on Wiggly Bispectra: Nonlinear Evolution and Reconstruction of Oscillations in the Squeezed Bispectrum”, *Phys. Rev. D* **112** (2025) 083503, arXiv:2505.13443 [astro-ph.CO].
- [193] D. Green, Y. Guo, J. Han, and B. Wallisch, “Light Fields during Inflation from BOSS and Future Galaxy Surveys”, *JCAP* **05** (2024) 090, arXiv:2311.04882 [astro-ph.CO].
- [194] G. Cabass, O. Philcox, M. Ivanov, K. Akitsu, S.-F. Chen, M. Simonović, and M. Zaldarriaga, “BOSS Constraints on Massive Particles during Inflation: The Cosmological Collider in Action”, *Phys. Rev. D* **111** (2025) 063510, arXiv:2404.01894 [astro-ph.CO].
- [195] W. Sohn, D.-G. Wang, J. Fergusson, and E. P. S. Shellard, “Searching for Cosmological Collider in the Planck CMB Data”, *JCAP* **09** (2024) 016, arXiv:2404.07203 [astro-ph.CO].

- [196] P. Suman, D.-G. Wang, W. Sohn, J. Fergusson, and E. P. S. Shellard, “How Significant Are Cosmological Collider Signals in the Planck Data?”, *arXiv:2511.17500* [astro-ph.CO].
- [197] O. Philcox, K. Zhong, and S. S. Sirletti, “Separating the Inseparable: Constraining Arbitrary Primordial Bispectra with Cosmic Microwave Background Data”, *arXiv:2511.19179* [astro-ph.CO].
- [198] P. Suman, D.-G. Wang, W. Sohn, J. Fergusson, and E. P. S. Shellard, “Searching for Cosmological Collider in the Planck CMB Data II: Collider Templates and Modal Analysis”, *arXiv:2512.22085* [astro-ph.CO].
- [199] Y. Akrami *et al.* (Planck Collaboration), “Planck 2018 Results. X. Constraints on Inflation”, *Astron. Astrophys.* **641** (2020) A10, *arXiv:1807.06211* [astro-ph.CO].
- [200] T. Louis *et al.* (ACT Collaboration), “The Atacama Cosmology Telescope: DR6 Power Spectra, Likelihoods and Λ CDM Parameters”, *JCAP* **11** (2025) 062, *arXiv:2503.14452* [astro-ph.CO].
- [201] E. Camphuis *et al.* (SPT-3G Collaboration), “SPT-3G D1: CMB Temperature and Polarization Power Spectra and Cosmology from 2019 and 2020 Observations of the SPT-3G Main Field”, *arXiv:2506.20707* [astro-ph.CO].
- [202] K. Smith and M. Zaldarriaga, “Algorithms for Bispectra: Forecasting, Optimal Analysis and Simulation”, *Mon. Not. Roy. Astron. Soc.* **417** (2011) 2, *arXiv:astro-ph/0612571*.
- [203] L. Senatore, K. Smith, and M. Zaldarriaga, “Non-Gaussianities in Single-Field Inflation and Their Optimal Limits from the WMAP 5-Year Data”, *JCAP* **01** (2010) 028, *arXiv:0905.3746* [astro-ph.CO].
- [204] R. Flauger, L. McAllister, E. Pajer, A. Westphal, and G. Xu, “Oscillations in the CMB from Axion Monodromy Inflation”, *JCAP* **06** (2010) 009, *arXiv:0907.2916* [hep-th].
- [205] R. Flauger and E. Pajer, “Resonant Non-Gaussianity”, *JCAP* **01** (2011) 017, *arXiv:1002.0833* [hep-th].
- [206] K. Smith, L. Senatore, and M. Zaldarriaga, “Optimal Analysis of the CMB Trispectrum”, *arXiv:1502.00635* [astro-ph.CO].
- [207] O. Philcox, “Searching for Inflationary Physics with the CMB Trispectrum. III. Constraints from Planck”, *Phys. Rev. D* **111** (2025) 123534, *arXiv:2502.06931* [astro-ph.CO].
- [208] S. Chang, M. Kleban, and T. Levi, “Watching Worlds Collide: Effects on the CMB from

- Cosmological Bubble Collisions”, *JCAP* **04** (2009) 025, arXiv:0810.5128 [hep-th].
- [209] S. Feeney, M. Johnson, D. Mortlock, and H. Peiris, “First Observational Tests of Eternal Inflation: Analysis Methods and WMAP 7-Year Results”, *Phys. Rev. D* **84** (2011) 043507, arXiv:1012.3667 [astro-ph.CO].
 - [210] T. Kim, J. H. Kim, S. Kumar, A. Martin, M. Münchmeyer, and Y. Tsai, “Probing Cosmological Particle Production and Pairwise Hotspots with Deep Neural Networks”, *Phys. Rev. D* **108** (2023) 043525, arXiv:2303.08869 [hep-ph].
 - [211] O. Philcox, S. Kumar, and J. C. Hill, “Searching for Inflationary Particle Production in Planck Data”, *Phys. Rev. D* **111** (2025) 103523, arXiv:2405.03738 [astro-ph.CO].
 - [212] J. Jasche and B. Wandelt, “Bayesian Physical Reconstruction of Initial Conditions from Large-Scale-Structure Surveys”, *Mon. Not. Roy. Astron. Soc.* **432** (2013) 894, arXiv:1203.3639 [astro-ph.CO].
 - [213] D. K. Ramanah, G. Lavaux, J. Jasche, and B. Wandelt, “Cosmological Inference from Bayesian Forward Modelling of Deep Galaxy Redshift Surveys”, *Astron. Astrophys.* **621** (2019) A69, arXiv:1808.07496 [astro-ph.CO].
 - [214] F. Villaescusa-Navarro *et al.*, “Robust Marginalization of Baryonic Effects for Cosmological Inference at the Field Level”, arXiv:2109.10360 [astro-ph.CO].
 - [215] S. Stopyra, H. Peiris, A. Pontzen, J. Jasche, and G. Lavaux, “Towards Accurate Field-Level Inference of Massive Cosmic Structures”, *Mon. Not. Roy. Astron. Soc.* **527** (2023) 1244, arXiv:2304.09193 [astro-ph.CO].
 - [216] L. Doeser, D. Jamieson, S. Stopyra, G. Lavaux, F. Leclercq, and J. Jasche, “Bayesian Inference of Initial Conditions from Nonlinear Cosmic Structures using Field-Level Emulators”, *Mon. Not. Roy. Astron. Soc.* **535** (2024) 1258, arXiv:2312.09271 [astro-ph.CO].
 - [217] A. Bairagi, B. Wandelt, and F. Villaescusa-Navarro, “The BIG SOBEL SEQUENCE: How Many Simulations Do We Need for Simulation-Based Inference in Cosmology?”, *Astron. Astrophys.* **703** (2025) A301, arXiv:2503.13755 [astro-ph.CO].
 - [218] S. McAlpine, J. Jasche, M. Ata, G. Lavaux, R. Stiskalek, C. Frenk, and A. Jenkins, “The Manticore Project I: A Digital Twin of Our Cosmic Neighbourhood from Bayesian Field-Level Analysis”, *Mon. Not. Roy. Astron. Soc.* **540** (2025) 716, arXiv:2505.10682 [astro-ph.CO].
 - [219] F. Schmidt, F. Elsner, J. Jasche, N. M. Nguyen, and G. Lavaux, “A Rigorous EFT-Based Forward Model for Large-Scale Structure”, *JCAP* **01** (2019) 042, arXiv:1808.02002

- [astro-ph.CO].
- [220] N. Kokron, J. DeRose, S.-F. Chen, M. White, and R. Wechsler, “The Cosmology Dependence of Galaxy Clustering and Lensing from a Hybrid N -Body-Perturbation-Theory Model”, *Mon. Not. Roy. Astron. Soc.* **505** (2021) 1422, arXiv:2101.11014 [astro-ph.CO].
 - [221] M. Pellejero-Ibanez, R. Angulo, M. Zennaro, J. Stuecker, S. Contreras, G. Arico, and F. Maion, “The BACCO Simulation Project: BACCO Hybrid Lagrangian Bias Expansion Model in Redshift Space”, *Mon. Not. Roy. Astron. Soc.* **520** (2023) 3725, arXiv:2207.06437 [astro-ph.CO].
 - [222] A. Nicola *et al.* (LSST Dark Energy Science Collaboration), “Galaxy Bias in the Era of LSST: Perturbative Bias Expansions”, *JCAP* **02** (2024) 015, arXiv:2307.03226 [astro-ph.CO].
 - [223] M. Ivanov, A. Obuljen, C. Cuesta-Lazaro, and M. Toomey, “Full-Shape Analysis with Simulation-Based Priors: Cosmological Parameters and the Structure-Growth Anomaly”, *Phys. Rev. D* **111** (2025) 063548, arXiv:2409.10609 [astro-ph.CO].
 - [224] C. Heinrich, O. Dore, and E. Krause, “Measuring $f_{\text{NL}}^{\text{loc}}$ with the SPHEREx Multitracer Redshift-Space Bispectrum”, *Phys. Rev. D* **109** (2024) 123511, arXiv:2311.13082 [astro-ph.CO].
 - [225] C. Shiveshwarkar, T. Brinckmann, and M. Loverde, “Constraining Multi-Field Inflation using the SPHEREx All-Sky Survey Power Spectra”, *JCAP* **05** (2024) 094, arXiv:2312.15038 [astro-ph.CO].
 - [226] X. Chen, R. Easther, and E. Lim, “Large Non-Gaussianities in Single-Field Inflation”, *JCAP* **06** (2007) 023, arXiv:astro-ph/0611645.
 - [227] A. Slosar *et al.*, “Scratches from the Past: Inflationary Archaeology through Features in the Power Spectrum of Primordial Fluctuations”, *Bull. Am. Astron. Soc.* **51** (2019) 98, arXiv:1903.09883 [astro-ph.CO].
 - [228] Z. Vlah, U. Seljak, M. Y. Chu, and Y. Feng, “Perturbation Theory, Effective Field Theory and Oscillations in the Power Spectrum”, *JCAP* **03** (2016) 057, arXiv:1509.02120 [astro-ph.CO].
 - [229] A. Vasudevan, M. Ivanov, S. Sibiryakov, and J. Lesgourgues, “Time-Sliced Perturbation Theory with Primordial Non-Gaussianity and Effects of Large Bulk Flows on Inflationary Oscillating Features”, *JCAP* **09** (2019) 037, arXiv:1906.08697 [astro-ph.CO].
 - [230] Y. Li, H.-M. Zhu, and B. Li, “Nonlinear Reconstruction of Features in the Primordial

- Power Spectrum from Large-Scale Structure”, *Mon. Not. Roy. Astron. Soc.* **514** (2022) 4363, arXiv:2102.09007 [astro-ph.CO].
- [231] S.-F. Chen, Z. Vlah, and M. White, “Modeling Features in the Redshift-Space Halo Power Spectrum with Perturbation Theory”, *JCAP* **11** (2020) 035, arXiv:2007.00704 [astro-ph.CO].
 - [232] M. Ballardini and N. Barbieri, “Refining the Nonlinear Modelling of Primordial Oscillatory Features”, *JCAP* **05** (2025) 059, arXiv:2411.02261 [astro-ph.CO].
 - [233] M. Ballardini, F. Finelli, F. Marulli, L. Moscardini, and A. Veropalumbo, “New Constraints on Primordial Features from the Galaxy Two-Point Correlation Function”, *Phys. Rev. D* **107** (2023) 043532, arXiv:2202.08819 [astro-ph.CO].
 - [234] T. Mergulhão, F. Beutler, and J. Peacock, “Primordial Feature Constraints from BOSS + eBOSS”, *JCAP* **08** (2023) 012, arXiv:2303.13946 [astro-ph.CO].
 - [235] R. Calderon, T. Simon, A. Shafieloo, and D. K. Hazra, “Primordial Features in Light of the Effective Field Theory of Large-Scale Structure”, *JCAP* **01** (2026) 057, arXiv:2504.06183 [astro-ph.CO].
 - [236] W. Sohn and J. Fergusson, “CMB-S4 Forecast on the Primordial Non-Gaussianity Parameter of Feature Models”, *Phys. Rev. D* **100** (2019) 063536, arXiv:1902.01142 [astro-ph.CO].
 - [237] M. Ballardini *et al.* (Euclid Collaboration), “Euclid: The Search for Primordial Features”, *Astron. Astrophys.* **683** (2024) A220, arXiv:2309.17287 [astro-ph.CO].
 - [238] D. Marolf and I. Morrison, “The IR Stability of de Sitter: Loop Corrections to Scalar Propagators”, *Phys. Rev. D* **82** (2010) 105032, arXiv:1006.0035 [gr-qc].
 - [239] M. Hogervorst, J. Penedones, and K. S. Vaziri, “Towards the Non-Perturbative Cosmological Bootstrap”, *JHEP* **02** (2023) 162, arXiv:2107.13871 [hep-th].
 - [240] L. Di Pietro, V. Gorbenko, and S. Komatsu, “Analyticity and Unitarity for Cosmological Correlators”, *JHEP* **03** (2022) 023, arXiv:2108.01695 [hep-th].
 - [241] M. Loparco, J. Penedones, K. Salehi Vaziri, and Z. Sun, “The Källén-Lehmann Representation in de Sitter Spacetime”, *JHEP* **12** (2023) 159, arXiv:2306.00090 [hep-th].
 - [242] T. Cohen, D. Green, and Y. Huang, “Operator Origin of Anomalous Dimensions in de Sitter Space”, *Phys. Rev. D* **111** (2025) 103513, arXiv:2407.08581 [hep-th].

- [243] G. Pimentel and C. Yang, “Strongly Coupled Sectors in Inflation: Gapless Theories and Unparticles”, arXiv:2503.17840 [hep-th].
- [244] Y. Jiang, G. Pimentel, and C. Yang, “Strongly Coupled Sectors in Inflation: Gapped Theories of Unparticles”, arXiv:2512.23796 [hep-th].
- [245] C. McCulloch, E. Pajer, and X. Tong, “A Cosmological Tachyon Collider: Enhancing the Long-Short-Scale Coupling”, *JHEP* **05** (2024) 262, arXiv:2401.11009 [hep-th].
- [246] M. Mirbabayi and M. Simonović, “Effective Theory of Squeezed Correlation Functions”, *JCAP* **03** (2016) 056, arXiv:1507.04755 [hep-th].
- [247] X. Chen and Y. Wang, “Quasi-Single Field Inflation with Large Mass”, *JCAP* **09** (2012) 021, arXiv:1205.0160 [hep-th].
- [248] D. Green and Y. Huang, “Flat Space Analog for the Quantum Origin of Structure”, *Phys. Rev. D* **106** (2022) 023531, arXiv:2203.10042 [hep-th].
- [249] K. Hinterbichler, L. Hui, and J. Khoury, “Conformal Symmetries of Adiabatic Modes in Cosmology”, *JCAP* **08** (2012) 017, arXiv:1203.6351 [hep-th].
- [250] W. Goldberger, L. Hui, and A. Nicolis, “One-Particle-Irreducible Consistency Relations for Cosmological Perturbations”, *Phys. Rev. D* **87** (2013) 103520, arXiv:1303.1193 [hep-th].
- [251] R. de Putter, O. Doré, D. Green, and J. Meyers, “Single-Field Inflation and the Local Ansatz: Distinguishability and Consistency”, *Phys. Rev. D* **95** (2017) 063501, arXiv:1610.00785 [hep-th].
- [252] L. Hui, A. Joyce, and S. Wong, “Inflationary Soft Theorems Revisited: A Generalized Consistency Relation”, *JCAP* **02** (2019) 060, arXiv:1811.05951 [hep-th].
- [253] B. Horn, L. Hui, and X. Xiao, “Soft-Pion Theorems for Large-Scale Structure”, *JCAP* **09** (2014) 044, arXiv:1406.0842 [hep-th].
- [254] A. Esposito, L. Hui, and R. Scoccimarro, “Non-Perturbative Test of Consistency Relations and Their Violation”, *Phys. Rev. D* **100** (2019) 043536, arXiv:1905.11423 [astro-ph.CO].
- [255] D. Baumann and D. Green, “Equilateral Non-Gaussianity and New Physics on the Horizon”, *JCAP* **09** (2011) 014, arXiv:1102.5343 [hep-th].
- [256] S. Behbahani, A. Dymarsky, M. Mirbabayi, and L. Senatore, “(Small) Resonant Non-Gaussianities: Signatures of a Discrete Shift Symmetry in the Effective Field Theory

- of Inflation”, *JCAP* **12** (2012) 036, arXiv:1111.3373 [hep-th].
- [257] S. Jazayeri, X. Tong, and Y. Zhu, “Every Wrinkle Carries A Memory: An Integro-Differential Bootstrap for Features in Cosmological Correlators”, arXiv:2511.00152 [hep-th].
 - [258] W. Press and P. Schechter, “Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation”, *Astrophys. J.* **187** (1974) 425.
 - [259] W. Coulton, F. Villaescusa-Navarro, D. Jamieson, M. Baldi, G. Jung, D. Karagiannis, M. Liguori, L. Verde, and B. Wandelt, “Quijote-PNG: The Information Content of the Halo Power Spectrum and Bispectrum”, *Astrophys. J.* **943** (2023) 178, arXiv:2206.15450 [astro-ph.CO].
 - [260] J. Sullivan, T. Prijon, and U. Seljak, “Learning to Concentrate: Multi-Tracer Forecasts on Local Primordial Non-Gaussianity with Machine-Learned Bias”, *JCAP* **08** (2023) 004, arXiv:2303.08901 [astro-ph.CO].
 - [261] B. Hadzhiyska and S. Ferraro, “Refining Local-Type Primordial Non-Gaussianity: Sharpened b_ϕ Constraints through Bias Expansion”, *Phys. Rev. D* **111** (2025) 103521, arXiv:2501.14873 [astro-ph.CO].
 - [262] D. Sharma, J. Sullivan, K. Akitsu, and M. Ivanov, “Equilateral Non-Gaussian Bias at the Field Level”, arXiv:2511.14677 [astro-ph.CO].
 - [263] L. Perez, S. Genel, E. Krause, and R. Somerville, “The Impact of Galaxy Formation on Galaxy Biasing and Implications for Primordial Non-Gaussianity Constraints”, arXiv:2602.04987 [astro-ph.CO].
 - [264] C. Zhao *et al.*, “MUltiplexed Survey Telescope: Perspectives for Large-Scale Structure Cosmology in the Era of Stage-V Spectroscopic Survey”, arXiv:2411.07970 [astro-ph.CO].
 - [265] E. Castorina *et al.*, “Packed Ultra-wideband Mapping Array (PUMA): Astro2020 RFI Response”, arXiv:2002.05072 [astro-ph.IM].
 - [266] A. Chudaykin, M. Ivanov, and O. Philcox, “Reanalyzing DESI DR1: 3. Constraints on Inflation from Galaxy Power Spectra and Bispectra”, arXiv:2512.04266 [astro-ph.CO].
 - [267] J. Fergusson, H. Gruetjen, E. P. S. Shellard, and M. Liguori, “Combining Power Spectrum and Bispectrum Measurements to Detect Oscillatory Features”, *Phys. Rev. D* **91** (2015) 023502, arXiv:1410.5114 [astro-ph.CO].

- [268] J. Fergusson, H. Gruetjen, E. P. S. Shellard, and B. Wallisch, “Polyspectra Searches for Sharp Oscillatory Features in Cosmic Microwave Sky Data”, *Phys. Rev. D* **91** (2015) 123506, arXiv:1412.6152 [astro-ph.CO].
- [269] P. D. Meerburg, M. Münchmeyer, and B. Wandelt, “Joint Resonant CMB Power Spectrum and Bispectrum Estimation”, *Phys. Rev. D* **93** (2016) 043536, arXiv:1510.01756 [astro-ph.CO].
- [270] A. Moradinezhad Dizgah, M. Biagetti, E. Sefusatti, V. Desjacques, and J. Noreña, “Primordial Non-Gaussianity from Biased Tracers: Likelihood Analysis of Real-Space Power Spectrum and Bispectrum”, *JCAP* **05** (2021) 015, arXiv:2010.14523 [astro-ph.CO].
- [271] W. Coulton, F. Villaescusa-Navarro, D. Jamieson, M. Baldi, G. Jung, D. Karagiannis, M. Liguori, L. Verde, and B. Wandelt, “Quijote-PNG: Simulations of Primordial Non-Gaussianity and the Information Content of the Matter Field Power Spectrum and Bispectrum”, *Astrophys. J.* **943** (2023) 64, arXiv:2206.01619 [astro-ph.CO].
- [272] A. Andrews *et al.* (Euclid Collaboration), “Euclid: Field-level inference of primordial non-Gaussianity and cosmic initial conditions”, arXiv:2412.11945 [astro-ph.CO].
- [273] J. Sullivan and S.-F. Chen, “Local Primordial Non-Gaussian Bias at the Field Level”, *JCAP* **03** (2025) 016, arXiv:2410.18039 [astro-ph.CO].
- [274] L. Doeser, M. Ata, and J. Jasche, “Learning the Universe: Learning to Optimize Cosmic Initial Conditions with Non-Differentiable Structure Formation Models”, *Mon. Not. Roy. Astron. Soc.* **1403** (2025) 1422, arXiv:2502.13243 [astro-ph.CO].
- [275] P. Adshead and A. Tishue, “Probing Beyond Local-Type Non-Gaussianity with Kinematic Sunyaev-Zel’dovich Tomography”, *Phys. Rev. D* **110** (2024) 103549, arXiv:2407.21094 [astro-ph.CO].
- [276] D. Anbajagane and H. Lee, “Primordial Physics in the Nonlinear Universe: Mapping Cosmological Collider Models to Weak-Lensing Observables”, arXiv:2509.02693 [astro-ph.CO].
- [277] D. Anbajagane and H. Lee, “Primordial Physics in the Nonlinear Universe: Signatures of Inflationary Resonances, Excitations and Scale Dependence”, arXiv:2509.02695 [astro-ph.CO].
- [278] S.-H. Zang and M. Münchmeyer, “Squeezed Limit Non-Gaussianity Estimation with Cosmic Shear”, arXiv:2512.07295 [astro-ph.CO].
- [279] C. Cheung, P. Creminelli, A. Fitzpatrick, J. Kaplan, and L. Senatore, “The Effective Field

- Theory of Inflation”, *JHEP* **03** (2008) 014, arXiv:0709.0293 [hep-th].
- [280] N. Arkani-Hamed, D. Baumann, H. Lee, and G. Pimentel, “The Cosmological Bootstrap: Inflationary Correlators from Symmetries and Singularities”, *JHEP* **04** (2020) 105, arXiv:1811.00024 [hep-th].
- [281] F. Beutler *et al.* (BOSS Collaboration), “The Clustering of Galaxies in the Completed SDSS-III Baryon Oscillation Spectroscopic Survey: Baryon Acoustic Oscillations in Fourier Space”, *Mon. Not. Roy. Astron. Soc.* **464** (2017) 3409, arXiv:1607.03149 [astro-ph.CO].
- [282] A. Font-Ribera, P. McDonald, N. Mostek, B. Reid, H.-J. Seo, and A. Slosar, “DESI and Other Dark Energy Experiments in the Era of Neutrino Mass Measurements”, *JCAP* **05** (2014) 023, arXiv:1308.4164 [astro-ph.CO].
- [283] O. Doré *et al.* (SPHEREx Collaboration), “SPHEREx Public Products Github Repository: Forecasted Galaxy Number Density and Bias for Fives Samples and Multiple Galaxy Redshift Bins”, 2020, https://github.com/SPHEREx/Public-products/blob/master/galaxy_density_v28_base_cbe.txt.