



Quijote-PNG: Simulations of Primordial Non-Gaussianity and the Information Content of the Matter Field Power Spectrum and Bispectrum

William R Coulton¹ , Francisco Villaescusa-Navarro^{1,2}, Drew Jamieson³ , Marco Baldi^{4,5,6} , Gabriel Jung^{7,8} , Dionysios Karagiannis⁹, Michele Liguori^{7,8}, Licia Verde^{10,11}, and Benjamin D. Wandelt^{1,12}

¹ Center for Computational Astrophysics, Flatiron Institute, 162 5th Avenue, New York, NY 10010, USA; wcoulton@flatironinstitute.org

² Department of Astrophysical Sciences, Princeton University, 4 Ivy Lane, Princeton, NJ 08544 USA

³ Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Straße 1, D-85748 Garching, Germany

⁴ Dipartimento di Fisica e Astronomia, Alma Mater Studiorum—University of Bologna, Via Piero Gobetti 93/2, I-40129 Bologna BO, Italy

⁵ INAF—Osservatorio Astronomico di Bologna, Via Piero Gobetti 93/3, I-40129 Bologna BO, Italy

⁶ INFN—Istituto Nazionale di Fisica Nucleare, Sezione di Bologna, Viale Bertè Pichat 6/2, I-40127 Bologna BO, Italy

⁷ Dipartimento di Fisica e Astronomia “G. Galilei,” Università degli Studi di Padova, via Marzolo 8, I-35131, Padova, Italy

⁸ INFN, Sezione di Padova, via Marzolo 8, I-35131, Padova, Italy

⁹ Department of Physics and Astronomy, University of the Western Cape, Cape Town 7535, South Africa

¹⁰ ICREA, Pg. Lluís Companys 23, Barcelona, E-08010, Spain

¹¹ ICC, University of Barcelona, IEEC-UB, Martí i Franquès, 1, E-08028 Barcelona, Spain

¹² Sorbonne Université, CNRS, UMR 7095, Institut d’Astrophysique de Paris, 98 bis bd Arago, F-75014 Paris, France

Received 2022 June 16; revised 2022 November 21; accepted 2022 December 2; published 2023 January 25

Abstract

Primordial non-Gaussianity (PNG) is one of the most powerful probes of the early universe, and measurements of the large-scale structure of the universe have the potential to transform our understanding of this area. However, relating measurements of the late-time universe to the primordial perturbations is challenging due to the nonlinear processes that govern the evolution of the universe. To help address this issue, we release a large suite of N -body simulations containing four types of PNG: QUIJOTE-PNG. These simulations were designed to augment the QUIJOTE suite of simulations that explored the impact of various cosmological parameters on large-scale structure observables. Using these simulations, we investigate how much information on PNG can be extracted by extending power spectrum and bispectrum measurements beyond the perturbative regime at $z = 0.0$. This is the first joint analysis of the PNG and cosmological information content accessible with power spectrum and bispectrum measurements of the nonlinear scales. We find that the constraining power improves significantly up to $k_{\max} \approx 0.3h \text{ Mpc}^{-1}$, with diminishing returns beyond as the statistical probes signal-to-noise ratios saturate. This saturation emphasizes the importance of accurately modeling all the contributions to the covariance matrix. Further, we find that combining the two probes is a powerful method of breaking the degeneracies with the Λ CDM parameters.

Unified Astronomy Thesaurus concepts: Non-Gaussianity (1116); Cosmology (343); Large-scale structure of the universe (902); Cosmological parameters from large-scale structure (340); Cosmological parameters (339)

1. Introduction

Measurements of the cosmic microwave background (CMB; Fixsen et al. 1997; Bennett et al. 2013; Planck Collaboration I 2020) and of the large-scale structure (LSS; Alam et al. 2017; d’Amico et al. 2020; Ivanov et al. 2020; Kobayashi et al. 2022) have revolutionized our understanding of the primordial universe, demonstrating that it was nearly homogeneous and isotropic with small, almost scale-invariant perturbations. However, wide classes of theoretical models, ranging from inflationary to ekpyrotic models (Lehners 2010; Martin 2016; Meerburg et al. 2019), can explain these observations. There are two proposed observational channels that have great power at discriminating between these theories: measurements of primordial gravitational waves and measurements of primordial non-Gaussianity (PNG; see, e.g., Achúcarro et al. 2022, for a recent overview of prospects for learning about inflation).

Primordial non-Gaussianity is interesting due to its strong sensitivity to the early universe physics: PNG can reveal

information about the field content of the universe, the strength of interactions, and particle content (Chen et al. 2007). Work by Green & Porto (2020) suggests that PNG measurements offer a method to verify the quantum nature of the primordial perturbations. The information encoded in PNG is complementary to other probes, such as primordial gravitational waves and the slope and running of the primordial power spectrum, that contain information about the duration and energy scale of inflation (see, e.g., Lyth & Riotto 1999, for a review).

Given the strong theoretical motivation, there have been many searches for primordial non-Gaussianity (Ferreira & Magueijo 1998; Komatsu et al. 2002, 2003; Creminelli et al. 2006; Slosar et al. 2008; Leistedt et al. 2014), with the leading constraints from measurements of the CMB by the Planck satellite (Planck Collaboration IX 2020). The CMB is an ideal probe for studying PNG, as CMB anisotropies are linearly related to the primordial perturbations and so CMB measurements can be straightforwardly related to the primordial universe. While upcoming CMB experiments will improve upon existing constraints (Abazajian et al. 2016; Ade et al. 2019), the challenges from foregrounds and the limitations imposed by silk-damping and the 2D nature of the CMB mean that improvements needed to reach the most interesting

theoretical levels will be highly challenging (Hill 2018; Jung et al. 2018; Coulton et al. 2020; Kalaja et al. 2021; Coulton et al. 2022a).

Constraints on PNG from measurements of the LSS promise, in principle, to far surpass existing constraints (Meerburg et al. 2017; Slosar et al. 2019). However, to date, the important measurements of PNG from LSS (Slosar et al. 2008; Ross et al. 2013; Giannantonio et al. 2014; Leistedt et al. 2014; Ho et al. 2015; Castorina et al. 2019; Mueller et al. 2021; Cabass et al. 2022a, 2022b; D’Amico et al. 2022) have not reached the level of the CMB.

This is due to complexities of large-scale foregrounds (Pullen & Hirata 2013; Rezaie et al. 2021), understanding the optimal statistics to use, and the challenge of modeling the LSS. Unlike the CMB, measurements of the late-time LSS are nonlinearly related to the primordial perturbations, which complicates inferences about the primordial universe. For the largest scales, powerful perturbation theory models have been developed and recently applied to data (Baldauf et al. 2011; Cabass et al. 2022a, 2022b; D’Amico et al. 2022). However, these methods will not be able to probe arbitrarily small scales, and thus it is unclear whether alternative approaches may be better suited to the problem. This is particularly interesting because recent work has suggested that information of PNG could be separated out from other small-scale effects by exploiting the locality of gravitational and baryonic processes (Baumann & Green 2022).

To aid our understanding of the relation between the primordial non-Gaussianity and late-time observables, we have run an extensive suite of simulations with PNG. This work builds on the results of numerous previous investigations: Dalal et al. (2008) and Desjacques et al. (2009) generated simulations containing *local* PNG and used these to discover and understand the impact of PNG on scale-dependent bias. These works were followed by those of Wagner et al. (2010) and Scoccimarro et al. (2012), who developed and implemented methods to generate simulations with PNG beyond the *local*. In this work, we use the methods developed in Scoccimarro et al. (2012) to generate a large ensemble of simulations with four different PNG shapes.

These simulations were designed to fit within the QUIJOTE suite of simulations (Villaescusa-Navarro et al. 2020); a large suite of simulations designed to quantify the information content on generic summary statistics and to provide enough training data for machine-learning algorithms. By making this choice, our PNG simulations can be used within a consistent framework to study how uncertainties in the standard cosmological parameters impact inferences of PNG. For example, while previous work has shown that the halo mass function (Lucchin & Matarrese 1988; LoVerde & Smith 2011), the matter probability density function (Valageas 2002; Uhlemann et al. 2018; Friedrich et al. 2020), topological measures (Biagetti et al. 2021, 2022), and field level analyses (Andrews et al. 2022) are potentially powerful probes of PNG, simulations are required to validate theoretical predictions or model these novel probes. An aim of this work is to help facilitate such analyses.

As both a validation of our simulations and a first application, we explore the properties of the matter power spectrum and bispectrum in these simulations. On the largest scales, the impact of PNG on the matter power spectrum and bispectrum is well-understood (e.g., Desjacques et al. 2009;

Baldauf et al. 2011; Karagiannis et al. 2018), and so these measurements can be used to validate our simulations. However, our simulations also allow us to push beyond the perturbative regime ($k \approx 0.1 h \text{ Mpc}^{-1}$ at $z = 0$) and explore the information content available on smaller scales. Recent work by Hahn et al. (2020) has shown that, for other cosmological parameters, small-scale bispectrum measurements of the halo and galaxy field potentially contain large amounts of information. Through the use of a simulation-based covariance matrix, we are able to assess the importance of all the contributions, including the supersample covariance, for parameter constraints.

This paper and our companion paper, Jung et al. (2022), are the first two in a series dedicated to investigating how we can learn more about primordial non-Gaussianity from upcoming measurements of the LSS.

In our companion paper, Jung et al. (2022), we explore the properties of matter power spectrum and bispectrum from a different and complementary perspective: we use the modal bispectrum (Fergusson et al. 2012; Schmittfull et al. 2013) to measure the matter bispectrum and then develop and validate a set of optimal compressed statistics, which enable the information in the matter power spectrum and bispectrum to be captured within a data vector of length the number of parameters. Additionally, the Jung et al. (2022) analysis focuses on measurements at $z = 1.0$, the redshift range relevant for upcoming surveys such as Euclid (Amendola et al. 2018), whereas here we examine $z = 0.0$. Through this complementarity, we can better understand the importance of small-scale information across redshifts. We stress that both of these works consider an idealized setup—the 3D matter field is not directly observable and we neglect all observational effects and their associated challenges, which will likely limit future analyses—and thus it is difficult to directly relate constraints reported here to future surveys. Further the unique scale-dependent bias feature that PNG introduces to biased tracers means that the measurements of biased tracers could contain more information than the matter field, on the same scales. However, the aim of these works is not to generate specific predictions but rather to assess the value of pushing to smaller scales and provide simulations that can be used as a sandbox to develop intuition into new estimators in a simplified environment. This represents a first step—and motivation—toward more complex and realistic analyses.

These two papers represent the first comprehensive investigation into what can be learned about PNG from small-scale measurements of the matter bispectrum and power spectrum. These papers are accompanied by the release the simulation data and our power spectrum and bispectrum measurements. In the next papers, we will consider the information content in the halos (Coulton et al. 2022b) and the information accessible from analyses of the field.

Our paper is structured as follows: in Section 2, we briefly overview the shapes of non-Gaussianity considered in this work. In Section 3, we describe our simulations of PNG. We describe our power spectrum and bispectrum estimators in Section 4, and then present our measurements in Section 5. In Section 6, we describe the details of how we compute our Fisher forecasts, and then we explore the constraining power of the bispectrum and power spectrum measurements in Section 7. We then summarize our key results in Section 8. In the three appendices, we describe the details of the generation of

non-Gaussian initial conditions (Appendix A), the details of the computation of the power spectrum and bispectrum covariance matrices (Appendix B), joints constraints of PNG and Λ CDM extensions, and finally the impact of a prior on the power spectrum convergence (Appendix D).

2. Primordial Non-Gaussianity

While primordial non-Gaussianity refers to any deviation from Gaussian initial conditions, in this work we restrict our analysis to the primordial bispectra, $B_\Phi(k_1, k_2, k_3)$, defined as

$$\begin{aligned} & (\Phi(\mathbf{k}_1)\Phi(\mathbf{k}_2)\Phi(\mathbf{k}_3)) \\ & = (2\pi)^3\delta^{(3)}(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3)B_\Phi(k_1, k_2, k_3), \end{aligned} \quad (1)$$

where $\Phi(\mathbf{k})$ is the primordial potential. We further restrict our analysis to four different primordial bispectra shapes, detailed below. These shapes are studied because they accurately approximate theoretically well-motivated shapes, provide insight into the primordial physics, and can be generated by a range of generic methods (see, e.g., Chen 2010; Achúcarro et al. 2022, for reviews of PNG).

The first shape we consider is the *local* shape, which has the primordial bispectrum

$$B_\Phi^{\text{local}}(k_1, k_2, k_3) = 2f_{\text{NL}}^{\text{local}} P_\Phi(k_1)P_\Phi(k_2) + 2 \text{ perm}, \quad (2)$$

where $P_\Phi(k_1)$ is the primordial power spectrum and $f_{\text{NL}}^{\text{local}}$ is the amplitude of this type of non-Gaussianity. *Local* non-Gaussianity is of observational interest because it is a powerful probe of the primordial field content. Maldacena (2003) and Creminelli & Zaldarriaga (2004) showed that, in single-field slow-roll inflation, the amplitude of the *local* bispectrum is $< O(\eta, \epsilon)$, where ϵ and η are the slow-roll parameters, and thus a measurement of larger *local* non-Gaussianity would present difficulties for slow-roll, single-field inflation. On the other hand, observable levels of *local* non-Gaussianity can be generated in multifield inflationary models, such as the curvaton or modulated reheating models (Lyth & Wands 2002; Dvali et al. 2004), and alternatives to inflation models (Lehners 2010). It should be noted that, while there is ongoing discussion (Pajer et al. 2013; Matarrese et al. 2021) as to whether the squeezed-limit bispectrum exactly vanishes for single-field models, this does not impact the power of a detection of *local* $\gg O(\epsilon, \eta)$ to rule out these models.

The second shape we consider is the *equilateral* non-Gaussianity, which has the following primordial bispectrum:

$$\begin{aligned} B_\Phi^{\text{equil.}}(k_1, k_2, k_3) & = 6f_{\text{NL}}^{\text{equil.}} [-P_\Phi(k_1)P_\Phi(k_2) + 2 \text{ perm.} \\ & - 2(P_\Phi(k_1)P_\Phi(k_2)P_\Phi(k_3))^{\frac{2}{3}} + P_\Phi(k_1)^{\frac{1}{3}}P_\Phi(k_2)^{\frac{2}{3}}P_\Phi(k_3) \\ & + 5 \text{ perm.}]. \end{aligned} \quad (3)$$

Equilateral non-Gaussianity well-approximates non-Gaussianities generated in Dirac—Born—Infeld inflation (Alishahiha et al. 2004) and ghost inflation (Arkani-Hamed et al. 2004), as well as generically in models with local, derivative interactions (Cheung et al. 2008a, 2008b).

The final shape that we consider is the *orthogonal* shape. The *orthogonal* bispectrum, along with the *equilateral* bispectrum, is used to span the types of non-Gaussianity within the effective field theory of inflation (EFTi; Cheung et al. 2008a, 2008b). Measurements of the *equilateral* and *orthogonal* bispectra can be used to

constrain the parameters of EFTi, including constraining the sound speed of primordial perturbations (Senatore et al. 2010; Planck Collaboration XXIV 2014). The full shape of the *orthogonal* bispectrum in EFTi is not separable (i.e., expressible as $F(k_1)G(k_2)H(k_3) + \text{perms.}$), and so it is challenging to simulate and analyze. Two approximations have been developed by Senatore et al. (2010). The first approximation was performed to approximate the *orthogonal* bispectrum in the CMB, and hence we refer to this as *orthogonal-CMB*. It has the bispectrum

$$\begin{aligned} B_\Phi^{\text{ortho-CMB}}(k_1, k_2, k_3) & = 6f_{\text{NL}}^{\text{ortho-CMB}} [-3P_\Phi(k_1)P_\Phi(k_2) \\ & + 2 \text{ perm.} - 8(P_\Phi(k_1)P_\Phi(k_2)P_\Phi(k_3))^{\frac{2}{3}} \\ & + 3P_\Phi(k_1)^{\frac{1}{3}}P_\Phi(k_2)^{\frac{2}{3}}P_\Phi(k_3) + 5 \text{ perm.}]. \end{aligned} \quad (4)$$

While this functional form is a good approximation for the CMB, which is a 2D projection of the primordial bispectrum, it is less accurate for measurements of the LSS (Senatore et al. 2010; Creminelli et al. 2011). In particular, its squeezed-limit divergence, which is suppressed for the 2D CMB, differs from the EFTi *orthogonal* shape. This different squeezed limit has important consequences for scale-dependent bias measurements and for the detectability. Despite these issues, we include it in our suite of simulations because the resulting scale-dependent bias ($\sim 1/k$) is different from the *local* case ($\sim 1/k^2$) and the *orthogonal-LSS* (see below) and *equilateral* shapes (~ 1), so the simulations help span possible PNG signatures (Schmidt & Kamionkowski 2010). Further *folded* bispectra, which can arise if the assumption of the Bunch–Davies vacuum is violated, is well-approximated by a linear sum of this shape with the *equilateral* shape (Chen et al. 2007; Meerburg et al. 2010). The impact of these subtleties on the halo field, where scale-dependent bias is important, is further explored in Coulton et al. (2022b).

The second approximation, also developed by Senatore et al. (2010) and hereafter referred to as *orthogonal-LSS*, has a bispectrum given by

$$\begin{aligned} B_\Phi^{\text{ortho-LSS}}(k_1, k_2, k_3) & = 6f_{\text{NL}}^{\text{ortho-LSS}} (P_\Phi(k_1)P_\Phi(k_2)P_\Phi(k_3))^{\frac{2}{3}} \\ & \left[- \left(1 + \frac{9p}{27} \right) \frac{k_3^2}{k_1k_2} + 2\text{perms} + \left(1 + \frac{15p}{27} \right) \frac{k_1}{k_3} \right. \\ & + 5\text{perms} - \left(2 + \frac{60p}{27} \right) \\ & + \frac{p}{27} \frac{k_1^4}{k_2^2k_3^2} + 2\text{perms} - \frac{20p}{27} \frac{k_1k_2}{k_3^2} + 2\text{perms} \\ & \left. - \frac{6p}{27} \frac{k_1^3}{k_2k_3^2} + 5\text{perms} + \frac{15p}{27} \frac{k_1^2}{k_3^2} + 5\text{perms} \right], \end{aligned} \quad (5)$$

where

$$p = \frac{27}{-21 + \frac{743}{7(20\pi^2 - 193)}}. \quad (6)$$

This shape is a better approximation to the EFTi *orthogonal* shape for the matter field and exhibits the correct squeezed-limit scaling. Note—as is discussed in Appendix A—that the *orthogonal-LSS* shape we consider is slightly modified to account for the fact that our simulations are not scale-invariant, i.e., $n_s \neq 1$.

Table 1
A Summary of the Key Properties of Our New Simulations

| Ω_m | Ω_Λ | Ω_b | σ_8 | h | n_s | Σm_ν (eV) | w | Box size (Mpc h ⁻¹) | $N_{\text{particles}}^{\frac{1}{3}}$ | # Realizations | ICs | Mass resolution ($M_\odot h^{-1}$) |
|------------|------------------|------------|------------|--------|--------|------------------------|-----|------------------------------------|--------------------------------------|----------------|---------|---|
| 0.3175 | 0.6825 | 0.049 | 0.834 | 0.6711 | 0.9624 | 0.0 | -1 | 1000 | 512 | 500 | 2LPTPNG | 6.56×10^{11} |

Note. The choices of cosmological parameters, random seeds, and simulation settings are otherwise chosen to match those used in Villaescusa-Navarro et al. (2020).

3. Simulations

In this work, we extend the QUIJOTE simulations by running a set of simulations at the QUIJOTE fiducial cosmology with primordial non-Gaussianity: QUIJOTE-PNG. These simulations have been run with the same settings (i.e., PM grid size, force resolution, etc.) as the original QUIJOTE simulations. Additionally, QUIJOTE-PNG uses simulations with matching random seeds to compute partial derivatives. Thus, these simulations are ideal for investigating the information in PNG jointly with other cosmological parameters. In Tables 1 and 2, we summarize the properties of the new simulations. In this section, we describe the details of the PNG initial conditions and the simulations. We refer the reader to Villaescusa-Navarro et al. (2020) for full details on the QUIJOTE simulations. All the PNG simulations are run with the Λ CDM cosmological parameters shown in Table 1. They follow the evolution of 512^3 dark matter particles from $z = 127$ down to $z = 0$. The following simulation products are publicly available: particle data at $z = 0.0, 0.5, 1.0, 2.0,$ and 3.0 , and the power spectrum and bispectrum measurements. See <https://quijote-simulations.readthedocs.io/en/latest/png.html> for more details.

3.1. Non-Gaussian Initial Conditions

To run non-Gaussian simulations, we use the method developed in Scoccimarro et al. (2012) to generate initial conditions that have primordial non-Gaussianity, and we implement several small changes to the public code released by Scoccimarro et al. (2012). Our updated code is available at <https://github.com/dsjamieson/2LPTPNG>. Here, we briefly review this method and refer the reader to Scoccimarro et al. (2012) for more details.

Generating initial conditions with *local* non-Gaussianity is straightforward. First, the modes of a Gaussian primordial potential field, $\Phi(\mathbf{k})$, are generated in Fourier space from the input power spectrum. This field is then inverse Fourier transformed to real space. The real space field is squared, mean-subtracted, scaled by the desired amplitude, $f_{\text{NL}}^{\text{local}}$, and then added back to the original potential. The resulting field is real Fourier transformed to obtain the modes of the primordial potential with *local* PnG.

In Fourier space, this corresponds to a convolution, so the process can be summarized as

$$\begin{aligned}
\Phi^{\text{local}}(\mathbf{k}) &= \Phi(\mathbf{k}) \\
&+ f_{\text{NL}}^{\text{local}} \int \frac{d^3k_1}{(2\pi)^3} \frac{d^3k_2}{(2\pi)^3} \Phi(\mathbf{k}_1) \Phi(\mathbf{k}_2) (2\pi)^3 \delta^{(3)}(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}) \\
&- f_{\text{NL}}^{\text{local}} \int \frac{d^3k}{(2\pi)^3} \Phi(\mathbf{k}) \Phi^*(\mathbf{k}),
\end{aligned} \tag{7}$$

Table 2

The Name Convention and f_{NL} Parameters Used for Each Simulation

| Name | $f_{\text{NL}}^{\text{local}}$ | $f_{\text{NL}}^{\text{equil.}}$ | $f_{\text{NL}}^{\text{ortho-CMB}}$ | $f_{\text{NL}}^{\text{ortho-LSS}}$ |
|----------|--------------------------------|---------------------------------|------------------------------------|------------------------------------|
| LC_p | 100 | 0 | 0 | 0 |
| LC_m | -100 | 0 | 0 | 0 |
| EQ_p | 0 | 100 | 0 | 0 |
| EQ_m | 0 | -100 | 0 | 0 |
| OR_p | 0 | 0 | 100 | 0 |
| OR_m | 0 | 0 | -100 | 0 |
| OR_LSS_p | 0 | 0 | 0 | 100 |
| OR_LSS_m | 0 | 0 | 0 | -100 |

where the second term removes the contribution from the mean, $\langle \Phi^2 \rangle$, which otherwise would contribute to background expansion.

Generating other types of primordial bispectra corresponds to adding a kernel to the convolution

$$\begin{aligned}
\Phi^X(\mathbf{k}) &= \Phi(\mathbf{k}) + f_{\text{NL}}^{\text{NG}} \int \frac{d^3k_1}{(2\pi)^3} \frac{d^3k_2}{(2\pi)^3} \\
&\times K(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}) \Phi(\mathbf{k}_1) \Phi(\mathbf{k}_2) (2\pi)^3 \delta^{(3)}(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}).
\end{aligned} \tag{8}$$

It should be noted that we impose $K(\mathbf{k}_1, \mathbf{k}_2, 0) = 0$, as this automatically accounts for the removal of the mean.

Crucially, for the *orthogonal-CMB*, *orthogonal-LSS*, and *equilateral* shapes, the kernels required can be written as a linear combination of separable terms, i.e., they can be written as sums over terms with the form $K(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}) = G(\mathbf{k}_1)H(\mathbf{k}_2)M(\mathbf{k})$. These can thus be generated in a similar manner to the *local*, with two modifications: first, the Gaussian potential modes are filtered before the inverse Fourier transform, and then they are filtered again after the final Fourier transform operation. The specific kernels we used and their coefficients are given in Appendix A.

The initial conditions are generated as follows: we generate the primordial anisotropies from a primordial power-law spectrum characterized by the amplitude, A_s , and tilt, n_s . These anisotropies are generated on a grid of size 1024^3 to minimize the impact of aliasing effects. We add non-Gaussianity via Equation (8). The perturbations are evolved to redshift $z = 0.0$ using linear transfer function, $T(k)$ from CAMB (Lewis et al. 2000). We then scale back its amplitude to $z = 127$ using the traditional growth factor, $D(z): P(k, z = 127) = D^2(z = 127)/D^2(z = 0)P(k, z = 0)$.¹³ The gridded field at $z = 127$ is then used with 2LPT to create the initial displacements of particles for our simulation. We note that we made a small change to the public code to generate initial conditions with $n_s \neq 1$ and with *orthogonal-LSS* non-Gaussianity.

¹³ Note that, in this work, the growth factor does not include relativistic species, and therefore the simulations are run without including radiation in the background. In Jung et al. (2022), we ran a suite of simulations including this effect and found that it makes a negligible difference to the observables considered here

3.2. N-body Evolution

After generating the initial conditions, we follow their gravitational evolution down to $z=0$ using the TreePM code GADGET-3 (Springel 2005). As stated above, we use the same parameter settings as the original QUIJOTE simulations. Halos are identified through the Friends-of-Friends (FoF) algorithm (Davis et al. 1985) with a value of the linking length equal to $b=0.2$.

3.3. Grid Assignment

To measure our statistics, we construct a mesh with $N_{\text{side}} = 512$ voxels for each simulation at the considered redshift. We use the Cloud-in-Cell (Hockney & Eastwood 1981) assignment scheme to distribute the particle positions to the grid. We account for this when computing the power spectra and bispectra by deconvolving the Cloud-in-Cell window function (Jing 2005). It should be noted that the large N_{side} used in our analysis means the effects of the window function and aliasing are negligible on the scales of interest: $k \leq 0.5 h \text{ Mpc}^{-1}$. This was verified by comparing our measurements to a set of grids generated using a fourth-order interpolation scheme with interlacing (see, e.g., Sefusatti et al. 2016, for a detailed discussion of these effects).

4. Statistical Probes

In this work, we examine two statistics of the field: the matter power spectrum and matter bispectrum.

4.1. Matter Power Spectrum

The matter power spectrum is defined as

$$\langle \delta^m(\mathbf{k}) \delta^{m*}(\mathbf{k}') \rangle = (2\pi)^3 \delta^{(3)}(\mathbf{k} - \mathbf{k}') P^{mm}(k). \quad (9)$$

We estimate the binned power spectrum, $\hat{P}(k_i)$ from simulations by computing

$$\hat{P}^{mm}(k_i) = \frac{1}{N_i} \sum_{k_i - \Delta k/2 < |\mathbf{k}| \leq k_i + \Delta k/2} \delta^m(\mathbf{k}) \delta^{m*}(\mathbf{k}), \quad (10)$$

where we sum over all modes that lay within a k -space bin of width Δk and N_i is the normalization that effectively counts the number of modes within our bin. We use the public Pylians³ code with bins of width equal to the fundamental mode, k_F , from k_F to $k_{\text{max}} = 0.5 h \text{ Mpc}^{-1}$.

4.2. Matter Bispectrum

The matter bispectrum is defined as

$$\begin{aligned} & \langle \delta^m(\mathbf{k}_1) \delta^m(\mathbf{k}_2) \delta^m(\mathbf{k}_3) \rangle \\ & = (2\pi)^3 \delta^{(3)}(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) B(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3). \end{aligned} \quad (11)$$

We estimate the bispectrum using the binned bispectrum estimator, in accordance with the implementation details given in Foreman et al. (2020)¹⁵ and using the methods developed in Scoccimarro (2000) and Tomlinson et al. (2019). The binned bispectrum estimator computes

$$\begin{aligned} \hat{B}(k_i, k_j, k_k) & = \frac{1}{N_{ijk}} \sum_{k_{i-1} < |\mathbf{k}_i| \leq k_i} \sum_{k_{j-1} < |\mathbf{k}_j| \leq k_j} \sum_{k_{k-1} < |\mathbf{k}_k| \leq k_k} \\ & \delta^h(\mathbf{k}_i) \delta^h(\mathbf{k}_j) \delta^h(\mathbf{k}_k) (2\pi)^3 \delta^{(3)}(\mathbf{k}_i + \mathbf{k}_j + \mathbf{k}_k), \end{aligned} \quad (12)$$

¹⁴ <https://github.com/franciscovillaescusa/Pylians3>

¹⁵ We do not use nbodykit, but rather a second code that was cross-validated during the work of Foreman et al. (2020).

which sums over all bispectra configurations that have wavenumbers that fall within three bins and is normalized by N_{ijk} , which counts the number of configurations in each bin. It should be noted that, as we sum over the discretized modes on the grid, the Dirac Delta function is really a Kronecker Delta for each of the three directions. The estimator is efficiently implemented by Kronecker delta

$$(2\pi)^3 \delta^{(3)}(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) = V \sum_{\mathbf{x}} e^{i\mathbf{x} \cdot (\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3)}, \quad (13)$$

where the summation is over the grid points with a volume factor, V , and then exchanging the order of the summations. The resulting expression can be rapidly evaluated by fast Fourier transforms. In this work, we use the same bins as in Hahn et al. (2020): we use bins of width $3k_F$ starting from half the fundamental to $k_{\text{max}} = 0.5 h \text{ Mpc}^{-1}$.

5. Power Spectrum and Bispectrum Measurements

We start by exploring the initial conditions used in our simulations. To do so, we generate 500 realizations of the Gaussian primordial potential field and 500 realizations for each of the primordial bispectra. In Figure 1(a), we show the impact of PNG on the primordial potential power spectrum. As is discussed in Wagner et al. (2010) and Scoccimarro et al. (2012), care needs to be taken that the kernels to generate the bispectrum do not generate unphysical corrections to the power spectrum. We see that there are no divergent corrections to the power spectrum introduced by our method. However, we do see that there are small corrections, $1 \times 10^{-5} f_{\text{NL}}$, to the power spectrum for the *orthogonal-CMB* non-Gaussianity. These corrections arise due to the loop correction to the power spectrum, some of which give k -dependent corrections while others are equivalent to an amplitude renormalization (see the Appendix of Wagner et al. (2010), for a detailed discussion). In principle, the scale-dependent contributions could be further minimized by different kernel choices; however, there is limited theoretical motivation to chose one parameterization over another, and hence we opted for the simplest implementation.

In Figure 2(a), we plot three slices showing how the bispectrum of the primordial initial conditions is affected by PNG. In addition, we plot the theoretically computed bispectrum. First, we note that the simulated initial conditions are in excellent agreement with the theoretical bispectrum, validating that our initial fields do contain the intended bispectrum shape. The deviations seen in the squeezed limit of the equilateral shape are statistical fluctuations that reduce when more simulations are considered. They are most visible in the squeezed limit of the equilateral shape, as the signal is smallest there. Second, it can clearly be seen that there are strong differences between the *orthogonal-LSS* and *orthogonal-CMB* templates. The differences are particularly large in the squeezed limit, as is expected, and will be very important for accurate modeling of effects such as scale-dependent bias (Dalal et al. 2008; Desjacques et al. 2009).

It is interesting to compare the primordial measurements to those at late times; in Figure 1(b), we show the matter power spectrum derivative with respect to PNG at $z=0.0$. We see that the impact of PNG on the power spectrum for the *local* and *equilateral* shapes is similar in structure. Our results for *local* and *equilateral* are consistent with the results found in Wagner et al. (2010), which is a nontrivial check because Wagner et al. (2010)

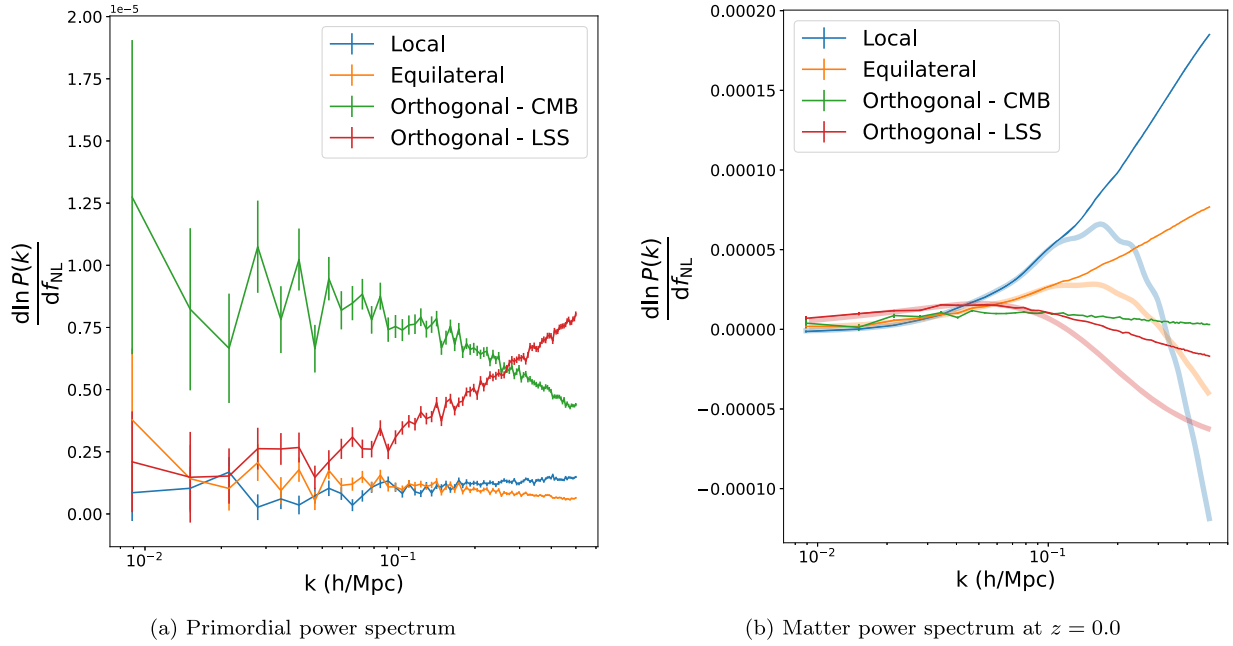


Figure 1. The derivative of the power spectrum with respect to the four different shapes of primordial non-Gaussianity. In Figure 1(a), we plot the derivatives of the initial conditions, whereas Figure 1(b) shows the derivatives obtained from the simulations at $z = 0.0$. The error bars denote the error on the mean. The thick lines show the one-loop EFT prediction.

use a different approach to generating the *equilateral* shape initial conditions. We also show the one-loop power spectrum prediction from EFT of LSS (Cabass et al. 2022a, 2022b; D’Amico et al. 2022); as expected, we find excellent agreement on the largest scales and significant differences on small, non-perturbative scales, where the one-loop term is inaccurate. It should be noted that a subtlety of the theory model is that the counter-term, required for the EFT of LSS to be a consistent theory, has an implicit dependence on f_{NL} that needs to be accounted for. The level of agreement is consistent with past work, e.g., that of Wagner et al. (2010), who compared simulated power spectra with PNG at $z=0$ to the time renormalization group prediction.

Finally, in Figure 2(b), we explore how the $z=0.0$ bispectrum is affected by PNG. We can see that, at $z=0.0$, the tree-level prediction for the bispectrum agrees on the largest scales, providing a simple validation of our simulations. However, as we move toward smaller scales, we find significant deviations from the tree-level prediction and the simulations—as is expected. Our results are in agreement with previous investigations of the matter bispectrum in Sefusatti et al. (2010), who measured the bispectrum in simulations with *local* non-Gaussianity up to $k_{\text{max}} \approx 0.3 h \text{ Mpc}^{-1}$. While our simulations are based off the same codes, this check is an important validation of the robustness to choices of the simulation parameters, e.g., resolution and force softening. Our results for the local case are also in agreement with the results of Enríquez et al. (2022), who investigated the similarities between *local* PNG and relativistic effects on the matter bispectrum. Note that the tree-level approximation is the leading method for most analyses and forecasts (Karagiannis et al. 2018; Slosar et al. 2019; Karagiannis et al. 2020; Cabass et al. 2022a, 2022b) to date. One-loop perturbative methods (e.g., Sefusatti 2009) have very recently been applied to galaxy survey measurements (D’Amico et al. 2022) and have been used to extend range of scales that can be accurately

modeled. These extract significantly more information about the primordial universe than the tree-level calculations, but it is impossible to use such methods on the maximum scales considered here, as they are non-perturbative. In this work, we investigate whether we can utilize the visible small-scale information.

6. Fisher Methodology

To estimate the information contained in the power spectrum and bispectrum, we use the Fisher formalism (Fisher 1935; Tegmark et al. 1997). The Fisher information, F_{IJ} , on parameters, θ , is defined as the variance of the score

$$F_{IJ} = \left\langle \frac{\partial \log \mathcal{L}(\mathbf{X}|\theta)}{\partial \theta_I} \frac{\partial \log \mathcal{L}(\mathbf{X}|\theta)}{\partial \theta_J} \right\rangle, \quad (14)$$

where $\mathcal{L}(\mathbf{X}|\theta)$ is the likelihood. The Fisher information is useful, as the minimum variance of an unbiased estimator, $\hat{\theta}$, for θ is given by Frechet & Frechet (1943), Aitken & Silverstone (1942), and Darrois (1945):

$$\text{Var}[\hat{\theta}] = F_{II}^{-1}, \quad (15)$$

with no summation over the repeated indices. Thus, by computing the inverse of the Fisher information, we can infer the maximum information that we can learn about θ from measurements of an observable, \mathbf{X} .

In this work, we assume that our observables, the matter power spectrum, and the bispectrum are well-approximated by a normal distribution. Thus,

$$\log \mathcal{L}(\mathbf{X}|\theta) = -\frac{1}{2} \sum_{ij} (O(k_i) - \bar{O}(k_i)) \Sigma_{ij}^{-1} (O(k_j) - \bar{O}(k_j)), \quad (16)$$

where $O(k_i)$ denotes either the power spectrum or bispectrum, \bar{O} is the observable mean, and Σ is the covariance matrix. We

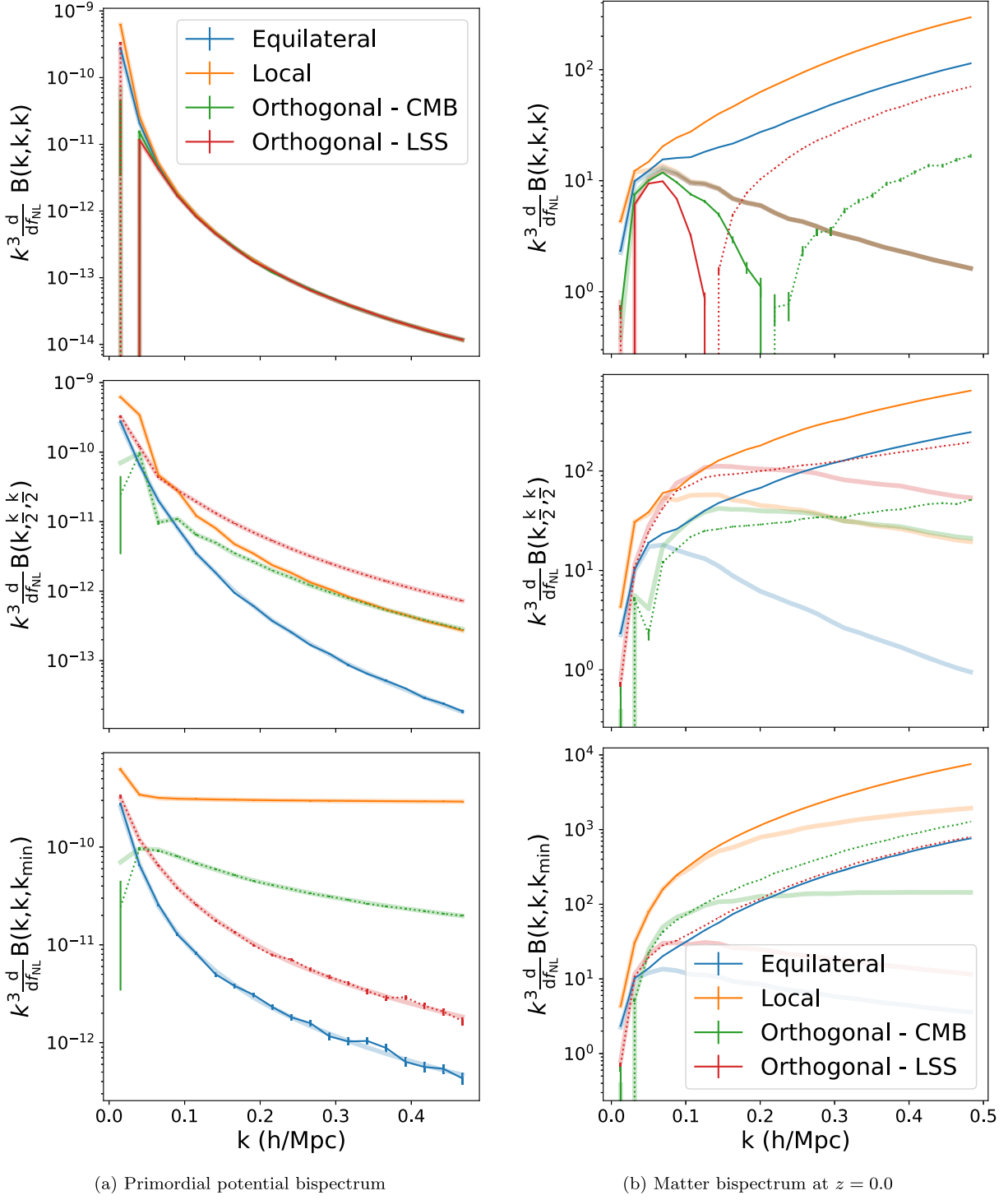


Figure 2. The equilateral (top), folded (middle), and squeezed bispectrum slices of the bispectrum derivative with respect to the four different shapes of primordial non-Gaussianity. In Figure 2(a), we plot the derivatives of the initial conditions, whereas Figure 2(b) shows the derivatives obtained from the simulations at $z = 0.0$. The dotted lines denote regions where the bispectrum is negative, the error bars are the error on the mean of our simulations, and the thick shaded line is the tree-level bispectrum prediction (the thickness of the theory curve was chosen solely to aid the visualization of overlapping lines).

also consider the information available from joint measurements of the power spectrum and bispectrum, in which case $O(k_i)$ is then the vector of both the power spectrum and bispectrum measurements. While the Gaussian likelihood assumption is not perfectly accurate (see, e.g., Scoccimarro 2000; Sellentin & Heavens 2018), it is sufficiently accurate to get an estimate of the information available.

For a Gaussian distribution, whose covariance is independent of the parameters of interest, the Fisher Information can be rewritten as

$$F_{IJ} = \frac{\partial \bar{O}(k_i)}{\partial \theta_I} \Sigma_{ij}^{-1} \frac{\partial \bar{O}(k_j)}{\partial \theta_J}. \quad (17)$$

In practice, the covariance matrix of the power spectrum and bispectrum will likely have some cosmological dependence; however, as was shown in Carron (2013), neglecting this dependence gives a better approximation for the true information of statistics like the power spectrum, whose true distribution is close to but not exactly Gaussian.

Thus, to assess the information content in the bispectrum and power spectrum, we require two ingredients: the derivative of our statistics of interest with respect to the parameters, and the covariance of the statistics.

The derivatives are computed numerically by central differencing: we compute mean power spectra and bispectra signal at $\theta + \delta\theta$ and $\theta - \delta\theta$ and estimate the derivative as

$$\frac{\partial \widehat{O}(k_i)}{\partial \theta_i} \approx \frac{\widehat{O}(k_i)|_{\theta=\theta+\delta\theta} - \widehat{O}(k_i)|_{\theta=\theta-\delta\theta}}{2\delta\theta}. \quad (18)$$

For the PNG derivatives, we use our new simulations, which were run with $\delta f_{\text{NL}} = \pm 100$, and for the other cosmological parameters considered (Ω_m , Ω_Λ , Ω_b , σ_8 , and h), we use the simulations summarized in Table 1 in Villaescusa-Navarro et al. (2020).

The covariance matrix is the second important part of our analysis. The contributions to the covariance matrix can be broken down into three terms:

$$\Sigma_{ij} = C_{ij}^{\text{Gaussian}} + C^{\text{connected}} + C^{\text{SSC}}. \quad (19)$$

The Gaussian contribution, C_{ij}^{Gaussian} , is straightforward to compute, whereas the connected, i.e., non-Gaussian, $C^{\text{connected}}$, and supersample covariance terms C^{SSC} are generally difficult to compute (see, e.g., Kayo et al. 2013; Chan & Blot 2017; Gualdi et al. 2018; Sugiyama et al. 2020). We compute the first two terms by measuring the covariance of the observables using 12,500 simulations in the fiducial cosmology (given the highly converged covariance matrix, we did not need to use all 15000 simulation available in the QUIJOTE suite). The supersample covariance term is computed as in Li et al. (2014a) and Chan et al. (2018); the details are summarized in Appendix B. Our base analysis does not include the supersample covariance term, but in Section 7.3 we assess the relative importance of it compared to the other terms. When inverting purely numerically estimated covariance matrices, we include a correction, the Hartlap/Anderson factor, to unbiased our estimate of the inverse covariance matrix or precision matrix (Hartlap et al. 2007).

We validate the accuracy of our results by verifying that the derivatives and covariance matrix are converged. In Figures 3 and 4, we examine how our Fisher forecasts change as we vary the number of simulations used to estimate the two components of our forecast. We see that the forecast is highly stable against variations in the number of simulations used in the calculation. We also see that the bispectrum and the combined power spectrum and bispectrum—hereafter joint—forecasts are also stable to changes in the number of simulations used to estimate the derivatives. However, we see that constraints from the power spectrum are not converged. This means that constraints from these statistics are likely to be overly optimistic, indicating that actual constraints derived from power spectrum measurements will be reduced compared to the results reported below.

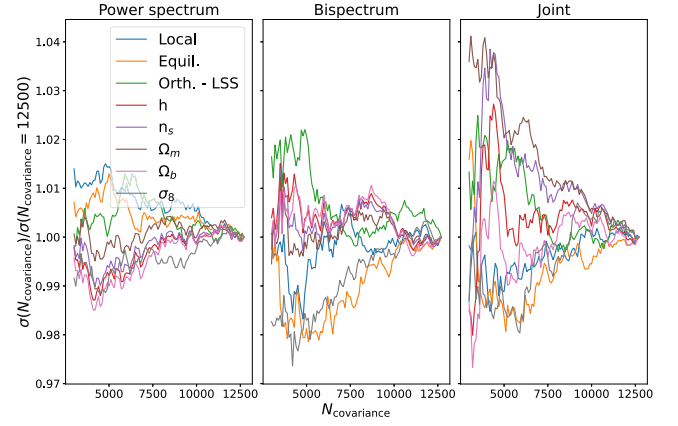


Figure 3. Variation in the marginalized parameter constraints as a function of number simulations used to compute the covariance matrix. The small fractional change suggests that the results are stable to the number of covariance simulations.

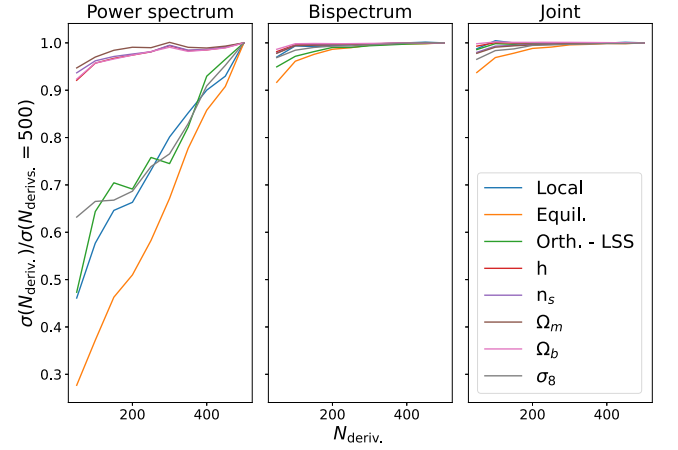


Figure 4. Variation in the marginalized parameter constraints as function of number simulations used to compute the Fisher derivatives. While the bispectrum and joint statistics are converged, the power spectrum shows large changes and is likely unconverged. For more details, see the discussion in Section 6.

Typically, one expects the power spectrum to converge more rapidly than higher-order statistics, so the convergence results in Figure 4 are at first surprising. In the case of primordial non-Gaussianity, however, the power spectrum is only sensitive to PNG through one-loop and higher-order terms, whereas the bispectrum is sensitive at the tree level. The power spectrum is thus only weakly sensitive to variations in the PNG parameters. Further, as will be seen in Section 7, the effect of PNG on the power spectrum is highly degenerate with other parameters. The strong degeneracies mean that the Fisher information needs to be estimated to a very high precision to be stable. While the inferences can be regularized, e.g., by the use of a prior as is explored in Appendix D, in the main text we prefer to present these unconverged forecasts. Alternatively, we could consider reparameterizations—for example, those of the form $\sigma_8 f_{\text{NL}}^\alpha$, which characterize the information probed by the power spectrum. However, we find that the bispectrum both measures different combinations and allows inferences on the physical parameters, e.g., f_{NL} . Thus, we retain the original parameterization. It is straightforward to see that the unconverged forecast is biased toward being overly constraining (see, e.g.,

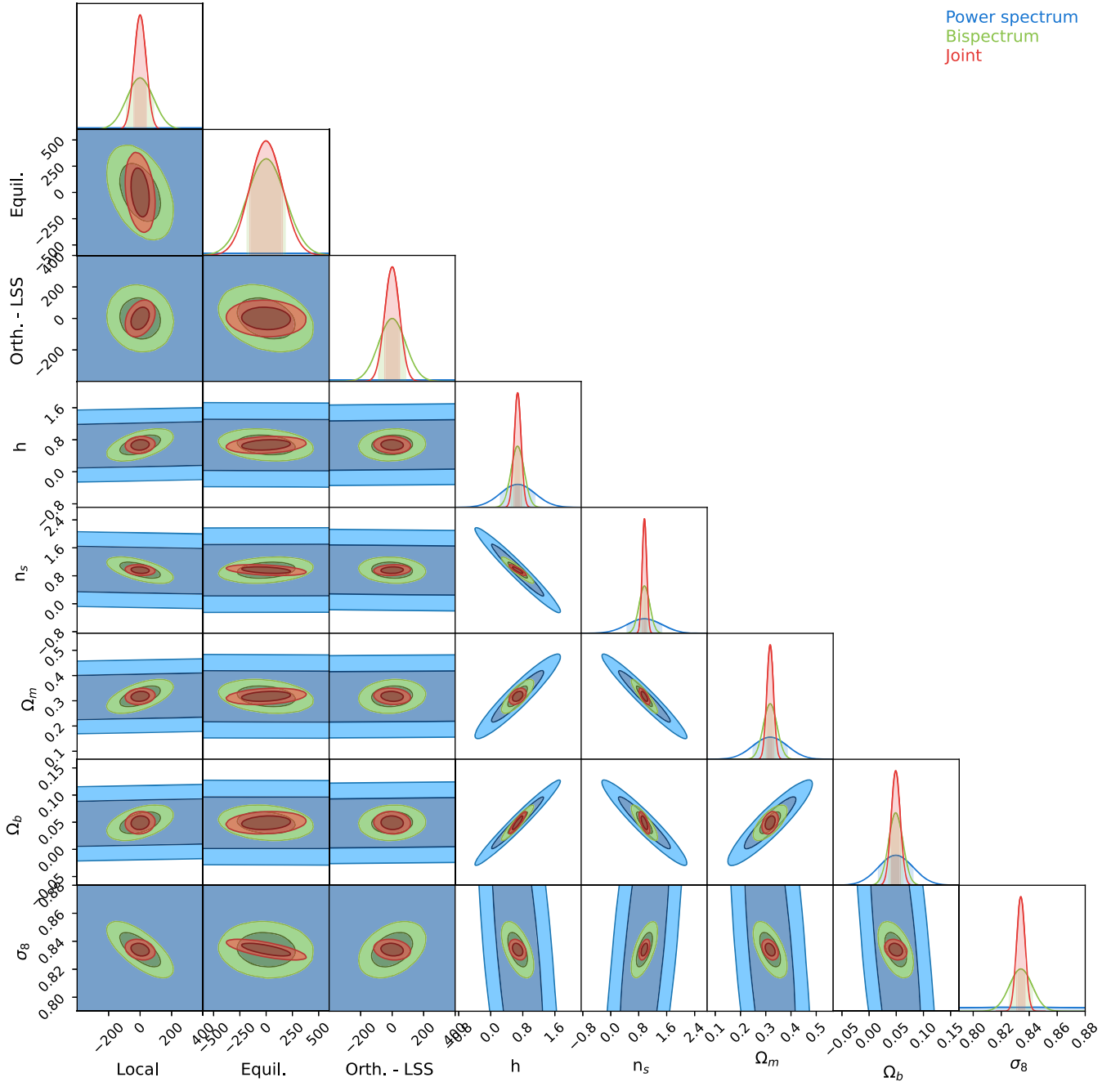


Figure 5. Joint constraints on the standard cosmological parameters and three shapes of primordial non-Gaussianity.

W. Coulton & Wandelt 2022, in preparation), and it is interesting to compare how broad even these optimistic power spectrum constraints are when compared to the bispectrum ones. It should be noted that, if we exclude the PNG parameters from the Fisher forecast, then the power spectrum forecasts are highly stable to variations in the number of derivative simulations used.

7. Cosmological Constraints

Next, we explore the constraining power from measurements at scales beyond the perturbation regime. Our constraints are obtained for a forecast experiment at $z=0.0$ and using modes from the fundamental mode, $k_F = 6.3 \times 10^{-3} h \text{ Mpc}^{-1}$, to $k_{\text{max}} = 0.5 h \text{ Mpc}^{-1}$.

7.1. Degeneracies

In Figure 5, we plot the constraints obtainable with measurements of the matter power spectrum and bispectrum including modes up to $k_{\text{max}} = 0.5 h \text{ Mpc}^{-1}$. Focusing first on the constraints from only power spectrum measurements, we find that they have negligible information on primordial non-Gaussianity when jointly measuring the different templates and marginalizing over cosmological parameters. This is expected, given the small impact of PNG on the power spectrum and the similarity of the induced changes to other cosmological parameters (e.g., n_s and s_8). Quantitatively, the impact of marginalization, as seen in Figure 6, can widen constraints by up to a factor of $\sim 100!$ Note that this marginalization also significantly degrades the cosmological parameter constraints

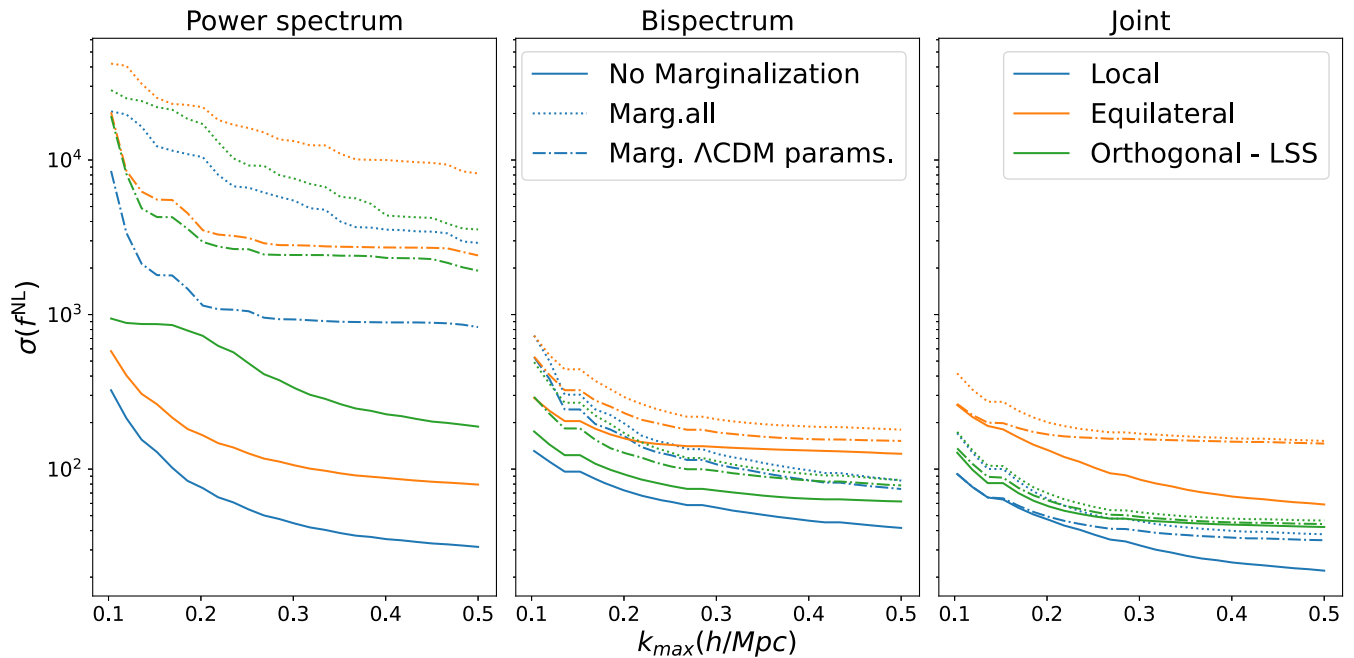


Figure 6. The constraining power of the power spectrum, bispectrum, and their combination as a function of scale for three shapes of non-Gaussianity. In solid lines, we show the unmarginalized constraints. In dotted–dashed lines, we show the constraints when marginalizing over the cosmological parameters. In dotted lines, we show the results from marginalizing over the cosmological parameters and the other shapes of primordial non-Gaussianity.

—especially σ_8 . The $\sigma_8 - f_{\text{NL}}$ degeneracy is very strong, as the PNG contributions to the matter power spectrum are very featureless—unlike, for example, h , which impacts the BAO and can be largely captured by a rescaling of the amplitude.

Focusing now on the bispectrum, we see that it offers vastly improved constraints. However, the bispectrum constraints also exhibit strong degeneracies with cosmological parameters. This indicates the challenge of separating the gravitationally sourced non-Gaussianity from PNG. When considering the joint power spectrum and bispectrum constraints, we see significant improvements that are largely driven by the improved constraining power on the standard, Λ CDM cosmological parameters. The joint analysis partially breaks the strong degeneracies (e.g., with Ω_m and σ_8) between the Λ CDM parameters, leading to better PNG constraints.

In Appendix C, we explore the degeneracies present when including two parameters beyond Λ CDM: w and $\sum m_\nu$.

7.2. The Value of Small Scales

Given the large information content available in the bispectrum, it is interesting to assess how much has been gained from the smaller scales. To answer this, we plot the constraining power of our probes as a function of the maximum included scale in Figure 6. We perform this investigation for three cases: (1) when constraining just one PNG shape, (2) when jointly constraining one PNG shape and the cosmological parameters, and (3) when we jointly fit the three PNG shapes and the cosmological parameters.

Focusing initially on case (1), we see that all probes yield significant gains by pushing the maximum scale up to $k_{\text{max}} \approx 0.3 h \text{ Mpc}^{-1}$. However, the rate of improvement slows significantly beyond this scale. This arises due to the increasingly strong non-Gaussian contributions to covariance matrix, which we explore further in Section 7.3. Importantly, the combined power spectrum and bispectrum analysis leads to

significant improvements over analyses of either probe alone. This is explored more in Jung et al. (2022).

By considering cases (2) and (3), when we include the effects of marginalization, we can see the strength of the degeneracies. For the power spectrum, the constraints degrade by up to two orders of magnitude. While the bispectrum constraints are less affected, we still find degradations of $\sim 100\%$, reflecting the correlations seen earlier in Figure 5. In this scenario, there is more benefit to pushing to smaller scales, as the extra information helps resolve the parameter degeneracies. The joint constraint panel shows that combining the probes mitigates the impact of marginalization, and again the constraints generally improve only modestly beyond $k_{\text{max}} = 0.3 h \text{ Mpc}^{-1}$. We also find minimal improvement for the *equilateral* shape, further underlying the degeneracy between the primordial shape and gravitational bispectra.

To contextualize the expected improvement with scale, we investigate the information available in the initial conditions. Given that the primordial universe is nearly Gaussian, we can compute the constraining power of an optimal estimator as

$$\sigma^{-2} = \frac{1}{6} \sum_{ijk} \frac{B^{\text{theory}}(k_i, k_j, k_k)^2}{P_\Phi(k_i)P_\Phi(k_j)P_\Phi(k_k)}, \quad (20)$$

where the sum counts the discrete Fourier modes in our simulation box. In Figure 7, we compare our constraints to the information available in the primordial bispectrum as a function of scale. As expected, the primordial constraints are tighter on all scales. Further, we see that pushing to increasingly small scales produces a far faster improvement on the constraining power (see Kalaja et al. 2021, for a more detailed discussion). As a cross-check, we have compared the primordial information that would be captured if we were to use our suboptimal, binned estimator. We find that the constraints degrade by 20%–30%. This could be reduced by optimizing the binning choice

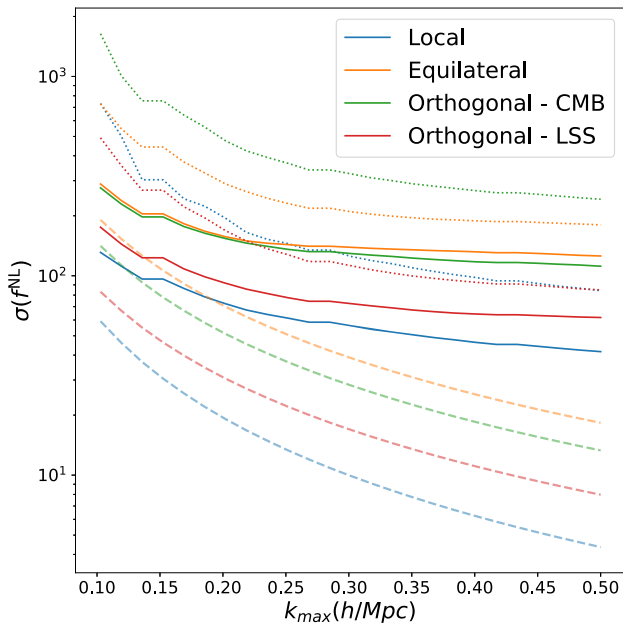


Figure 7. The constraining power of the bispectrum as a function of scale compared to the information available in the primordial field. The dashed lines denote the constraining power in the primordial potential (without any marginalization), the solid lines denote the constraining power at $z=0.0$ without marginalization, and the dotted lines are the constraints at $z=0.0$ with marginalization.

further (e.g., using a smaller bin width or having more bins for configurations where the bispectrum varies rapidly) or using an alternative estimator such as the modal estimator.

In this analysis, we also consider the information available to constrain the *orthogonal-CMB* PNG, but in our joint fits we only include either the *orthogonal-CMB* or *-LSS* shapes, not both. We do not include both simultaneously, as they are both approximations to the same EFTi bispectrum and thus are expected to be highly degenerate. We see similar behavior in the constraints for the *orthogonal-CMB bispectrum*.

7.3. The Importance of Covariance Matrix Modelling

The flattening of the information content arises as the bispectrum and power spectrum signal-to-noise (S/N) ratios flatten beyond $k \approx 0.3 h/\text{Mpc}$ as is seen in Figure 8. These results are consistent with those reported by Chan & Blot (2017), who found that the bispectrum S/N is degraded by a factor of ~ 10 at $k_{\text{max}} = 0.5 h \text{ Mpc}^{-1}$ at $z=0.0$ and that the S/N flattens above $k_{\text{max}} \approx 0.3 h \text{ Mpc}^{-1}$. This flattening arises as the nonlinear evolution of structure formation leads to strong correlations between the small-scale modes, reducing the available primordial information.

To further understand the flattening of our constraints, we explore how the different contributions to the covariance matrix propagate to parameter constraints. In Figure 9, we compared the bispectrum-only constraints obtained with the full covariance matrix to the constraints obtained when only a subset of the contributions are included. We find that including the non-Gaussian contributions, both the diagonal and off-diagonal elements, has a significant impact on the constraining power and leads to a degradation of the constraints by a factor of ~ 4 . This supports our assertion that the non-Gaussian terms cause the flattening of the S/N. These results agree with those seen in Barreira 2020, Gualdi 2020) and Biagetti et al. (2022).

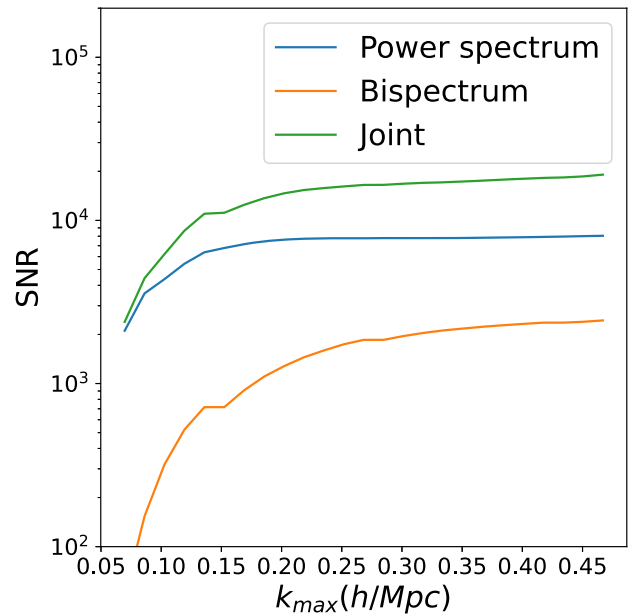


Figure 8. The cumulative signal-to-noise ratio for the power spectrum, bispectrum, and joint analysis.

These works show that the off-diagonal covariance matrix terms are highly important for constraints on *local* PNG and, when included, lead to significantly degraded constraints. Our work extends these results, highlighting their importance for other types of PNG.

Note that, at smaller k_{max} , the size of the non-Gaussian diagonal and off-diagonal terms is reduced but not negligible. At $k_{\text{max}} = 0.1 h \text{ Mpc}^{-1}$, there is still a degradation of some constraints by ~ 2 . It is interesting to note that the supersample covariance matrix terms have minimal impact on our parameter constraints despite being a nontrivial contribution to the covariance matrix, as was found in, e.g., Hamilton et al. (2006), Rimes & Hamilton (2006), Takahashi et al. (2009), and Chan et al. (2018). When we examine the unmarginalized constraints we find, for the power spectrum, that these supersample covariance terms are important. This is similar to the results found by Li et al. (2014b). Thus, for the power spectrum, the large increase in constraints due to marginalization swamps the impact of the supersample covariance terms, at least for the setup considered here. For the bispectrum, the limited degradation to the constraints arises because the strongly affected configurations are small-scale equilateral shapes (Chan et al. 2018; Barreira 2019) that contribute minimally to the constraining power, due to the large non-Gaussian covariance.

8. Conclusions

In this work, we have presented and validated a suite of N -body simulations that contain primordial non-Gaussianity: QUIJOTE-PNG. These augment the extensive QUIJOTE suite, matching the random seeds, configurations, and data formats of the previously run QUIJOTE simulations. This allows for consistent and seamless joint analyses of available information from the primordial universe and cosmological parameters.

The simulation data, including particle data at $z=0.0, 0.5, 1.0, 2.0,$ and 3.0 , along with the power spectrum and bispectrum measurements at $z=0$, are publicly available.

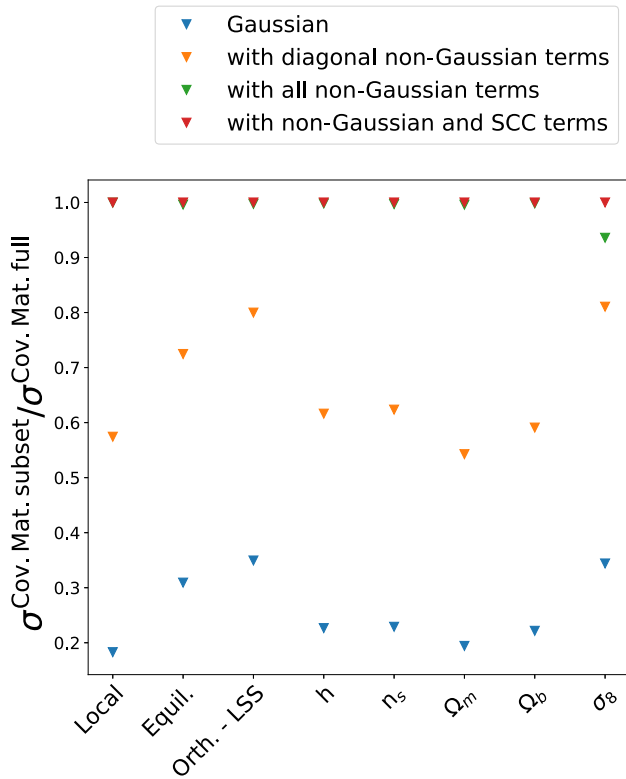


Figure 9. A comparison of the marginalized cosmological constraints from bispectrum-only measurements when using only a subset of the contributions to the covariance matrix. The result is normalized by the errors obtained using all the terms listed in Equation (19).

Please refer to the documentation at <https://quijote-simulations.readthedocs.io/en/latest/png.html>.

As a first use case, we investigated the information content accessible via measurements of the matter power spectrum and bispectrum. Our simulations allow us to extend the range of scales used in our analysis beyond the perturbative scales used in most advanced forecasts and analyses (e.g., Karagiannis et al. 2018; Cabass et al. 2022a; D’Amico et al. 2022). This analysis, along with our companion paper (Jung et al. 2022), are the first investigations of this kind. Our results show that significant gains can be made by including modes up to $\sim k_{\text{max}} \approx 0.3 h \text{ Mpc}^{-1}$; after this, the improvements are less dramatic.

Beyond these scales, the large non-Gaussian covariance of our statistical probes limits improvements, despite large signals from PNG on these scales, by saturating the S/N of our probes. This saturation is not captured when Gaussian approximations are used to model the covariance matrix, thus highlighting the importance of accurately including the non-Gaussian terms. Similar saturation has been observed in the power spectrum (Rimes & Hamilton 2005; Neyrinck & Szapudi 2007); however, these authors suggest that, when pushing to even smaller scales, $k > 1 h \text{ Mpc}^{-1}$, the information content again improves rapidly with scale. The resolution limitations of our simulations prevent us from investigating this, though we note that robustly exploiting such scales would be challenging to due the impact of baryonic processes (see, e.g., Chisari et al. 2018). Interestingly, we find that, on the scales considered here, the supersample contributions to the covariance matrix are unimportant. In our companion paper, Jung et al. (2022), we

examine the detailed features of the joint power spectrum–bispectrum covariance, which is an important component.

The challenge of modeling the large bispectrum covariance provides a large motivation to consider more efficient representations of the bispectrum information. Two such representation are explored in detail Jung et al. (2022). First, we measure the modal bispectrum, which—by means of a judicious choice of basis functions—allows the information in the bispectrum to be fully captured in $O(100)$ modal coefficients instead of the 1800 bispectrum bins used here. Second, we then implement a further compressed estimator that compresses the information down to just eight numbers.

We chose to work at $z = 0.0$, despite the limited volume of our universe that is observable at $z = 0.0$. This was done to allow a comparison to the information content in the halo field (Coulton et al. 2022b), where resolution limitations mean we need to work at $z = 0.0$ to obtain observationally relevant tracer densities. Measurements at higher redshifts will demonstrate similar features to those seen in our analysis, with the modification that the nonlinear scale will be moved to smaller scales—this is explicitly seen in our companion paper, Jung et al. (2022), where we investigate the information at $z = 1$. A second important choice is the size of box, and therefore k_{min} , which is small compared to the volumes probed by upcoming experiments (e.g., Doré et al. 2014; DESI Collaboration et al. 2016). PNG constraints are highly sensitive to the minimum scale used in the analysis; see Kalaja et al. (2021) for an extended discussion. Thus, the constraints on PNG obtainable from larger surveys will not just scale as $\sqrt{\text{Vol}}$ compared with those presented here. It is also expected that the degeneracies seen with cosmological parameters will be somewhat reduced, as the shapes of the gravitationally induced and the primordial bispectra will be more easily differentiated over a larger domain. Running larger simulations, while maintaining the small-scale resolution is computationally too expensive for this work; however, our results demonstrate the importance of small-scale measurements that will only be enhanced for larger-volume surveys.

A comparison between the information captured by our bispectrum measurements and the information available in the primordial field, which bounds how much we could possibly know about PNG, shows that the matter power spectrum and bispectrum estimators only capture $\sim 10\%$ of the total available information. This highlights the need to explore alternative approaches, such as topological measures (Biagetti et al. 2021), the matter pdf (Friedrich et al. 2020), machine-learning approaches (Giusarma et al. 2019; Villaescusa-Navarro et al. 2020), or field level approaches (Andrews et al. 2022), to fully capture the information available in the matter field. The simulations presented here are designed to facilitate future investigations along these lines.

Finally, it is important to emphasize that the 3D matter field considered in this work is not directly accessible. Observationally, we can access the 2D integrated matter field through gravitational lensing, and these observations will be impacted by observational systematics and baryonic processes, which are neglected here. Both of these effects will likely impact the observationally attainable information. We can also make biased estimates of the matter field from surveys of biased tracers, such as galaxy positions and velocities, and signatures of the neutral hydrogen distribution through 21 cm mapping. Thus, we stress the results of this work are not forecast

constraints for upcoming surveys, but they do provide a way to assess the value of smaller-scale bispectrum measurements and provide a suite of simulations for the community to use to develop and test new analysis tools for PNG. This work represents an important first step toward these more sophisticated probes. In a follow-up work, we will explore more direct observables, such as the halo and galaxy fields.

The authors are very grateful to Oliver Philcox, Daan Meerburg, Simone Ferraro, Shirley Ho, David Spergel, and Yin Li for their advice on numerous aspects of our analysis. G.J., M.L., and M.B. were supported by the project ‘‘Combining Cosmic Microwave Background and Large Scale Structure data: an Integrated Approach for Addressing Fundamental Questions in Cosmology,’’ funded by the MIUR Progetti di Ricerca di Rilevante Interesse Nazionale (PRIN) Bando 2017—grant 2017YJYZAH. D.K. is supported by the South African Radio Astronomy Observatory (SARAO) and the National Research Foundation (Grant No. 75415). B.D.W. acknowledges support by the ANR BIG4 project, grant ANR-16-CE23-0002 of the French Agence Nationale de la Recherche; and the Labex ILP (reference ANR-10-LABX-63) part of the Idex SUPER, and received financial state aid managed by the Agence Nationale de la Recherche, as part of the program Investissements d’avenir under the reference ANR-11-IDEX-0004-02. The Flatiron Institute is supported by the Simons Foundation. L.V. acknowledges ERC (BePreSySe, grant agreement 725327), PGC2018-098866- B-I00 MCIN/AEI/10.13039/501100011033 y FEDER ‘‘Una manera de hacer Europa,’’ and the ‘‘Center of Excellence Maria de Maeztu 2020-2023’’ award to the ICCUB (CEX2019-000918-M funded by MCIN/AEI/10.13039/501100011033).

Appendix A Initial Condition Generation

The kernels used to generate initial conditions with *local*, *equilateral*, and *orthogonal-CMB* non-Gaussianity are identical to those used in Scoccimarro et al. (2012). We use the notation

$$K_A^\alpha = \int \frac{d^3k_1}{(2\pi)^3} \frac{d^3k_2}{(2\pi)^3} K_A^\alpha(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}) \Phi(\mathbf{k}_1) \Phi(\mathbf{k}_2) \times (2\pi)^3 \delta^{(3)}(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}) \quad (\text{A1})$$

and the set of kernels defined in Table 3. As in Scoccimarro et al. (2012), we generate *local* non-Gaussianity using

$$\Phi^{\text{local}} = \Phi + f_{\text{NL}}^{\text{Local}} K_A^I \quad (\text{A2})$$

equilateral with

$$\Phi^{\text{Equil.}} = \Phi + f_{\text{NL}}^{\text{Equil.}} [-3K_A^I - 6K_C^I + 4K_B^I + 2K_B^{II}] \quad (\text{A3})$$

and *orthogonal-CMB* with

$$\Phi^{\text{Ortho-CMB}} = \Phi + f_{\text{NL}}^{\text{Ortho-CMB}} [-9K_A^I - 24K_C^I + 10K_B^I + 8K_B^{II}] \quad (\text{A4})$$

A.1. Orthogonal-LSS

In this section, we derive the kernels used in Equation (8) to generate initial conditions with the *orthogonal-LSS* non-Gaussianity. This derivation mirrors the derivation and choices made for the other shapes in Scoccimarro et al. (2012). The

bispectrum in Equation (5) is separable, and the first three terms are of the same form as the *equilateral* and *f* terms used in Scoccimarro et al. (2012). The next step is to enumerate the possible kernels associated with each term, and these are listed in Table 3.

Using this, we can show that the *orthogonal-LSS* bispectrum can be generated by the following operations:

$$\begin{aligned} & \frac{\Phi^{NG}(\mathbf{k})}{3} \\ &= - \left[1 + \frac{9p}{27} \right] \left[(1-u)K_A^I + \frac{u}{2}K_A^{II} \right] - \left[2 + \frac{60p}{27} \right] \frac{1}{3}K_C^I \\ &+ \left[1 - \frac{15p}{27} \right] [(1-t-s)K_B^I + tK_B^{II} + sK_C^{III}] \\ &+ \frac{p}{27} \left[(1-a)K_D^I + \frac{a}{2}K_D^{II} \right] \\ &- \frac{20p}{27} \left[(1-b)K_E^I + \frac{b}{2}K_E^{II} \right] \\ &+ \frac{6p}{27} [(1-c-d)K_F^I + cK_F^{II} + dK_F^{III}] \\ &+ \frac{15p}{27} [(1-e-f)K_G^I + eK_G^{II} + fK_G^{III}]. \end{aligned} \quad (\text{A5})$$

There is a large family of fields with the correct bispectrum with differing $N \neq 3$ -point functions. Our approach to choosing the terms is to impose the constraint that the largest corrections to the power spectrum should scale as k^{-2} , and for simplicity, to set as many of the remaining coefficients as possible to zero. For two kernels, $K^A(k_1, k_2, k_3) = k_1^a k_2^b k_3^c$ and $K^B(k_1, k_2, k_3) = k_1^\alpha k_2^\beta k_3^\gamma$, the leading correction to the power spectrum is given as

$$\begin{aligned} \delta P(k_3) &= k_3^{c+\gamma} \int \frac{d^3k_1}{(2\pi)^3} [k_1^{a+\alpha-3} (|\mathbf{k}_1 + \mathbf{k}_3|)^{b+\beta-3} \\ &+ k_1^{b+\alpha-3} (|\mathbf{k}_1 + \mathbf{k}_3|)^{a+\beta-3}]. \end{aligned} \quad (\text{A6})$$

Thus, to satisfy our constraints we require

$$\begin{aligned} & \frac{15p}{27}f - \frac{6p}{27}d - \frac{20p}{27} \frac{(1-b)}{2} + \frac{p}{27} \frac{a}{2} = 0 \\ & - \left(1 + \frac{9p}{27} \right) \frac{u}{2} + \left(1 + \frac{15p}{27} \right) s - \frac{6p}{27}c = 0 \\ & \left(1 + \frac{15p}{27} \right) t + \frac{15p}{27}e - \frac{1}{6} \left(2 + \frac{60p}{27} \right) = 0. \end{aligned} \quad (\text{A7})$$

We thus choose

$$\begin{aligned} b &= 1 \\ t &= \frac{\left(2 + \frac{60p}{27} \right)}{6 \left(1 + \frac{15p}{27} \right)} \end{aligned} \quad (\text{A8})$$

$$a = c = d = e = f = u = s = 0. \quad (\text{A9})$$

Table 3
The Kernels Associated with Each Term in the Bispectrum (as Shown in the First Column)

| Bispectrum Term | Label(α) | K_α^I | K_α^{II} | K_α^{III} |
|-------------------------------|-------------------|--|--|---|
| $\frac{1}{k_1^3 k_2^3}$ | A | 1 | $\frac{1}{2} \frac{1}{k_3^3} [k_1^3 + k_2^3]$ | ... |
| $\frac{1}{k_1 k_2^2 k_3^3}$ | B | $\frac{1}{k_3} [k_1 + k_2]$ | $\frac{1}{k_3^2} [k_1^2 + k_2^2]$ | $\frac{1}{k_3^3} [k_1^2 k_2 + k_1 k_2^2]$ |
| $\frac{1}{k_1^2 k_2^2 k_3^3}$ | C | $\frac{1}{3} \frac{1}{k_3^2} k_1 k_2$ | ... | ... |
| $\frac{k_1^2}{k_2^2 k_3^4}$ | D | $\frac{k_3^2}{k_1 k_2}$ | $\frac{1}{2} \frac{1}{k_3^4} \left[\frac{k_1^5}{k_2} + \frac{k_2^5}{k_1} \right]$ | ... |
| $\frac{1}{k_1 k_2 k_3^4}$ | E | $\frac{k_1^2 k_2^2}{k_3^4}$ | $\frac{1}{2} \frac{1}{k_3} \left[\frac{k_1^2}{k_2} + \frac{k_2^2}{k_1} \right]$ | ... |
| $\frac{k_1}{k_2^2 k_3^4}$ | F | $k_3 \left[\frac{1}{k_1} + \frac{1}{k_2} \right]$ | $\frac{1}{k_3^3} \left[\frac{k_1^4}{k_2} + \frac{k_2^4}{k_1} \right]$ | $\frac{1}{k_3^4} [k_1^4 + k_2^4]$ |
| $\frac{1}{k_2^2 k_3^4}$ | G | $\left[\frac{k_2}{k_1} + \frac{k_1}{k_2} \right]$ | $\frac{1}{k_3^2} \left[\frac{k_1^3}{k_2} + \frac{k_2^3}{k_1} \right]$ | $\frac{1}{k_3^4} [k_1^3 k_2 + k_1 k_2^3]$ |

Notes. Each bispectrum term can typically be generated by multiple kernels. While the result is equivalent at the bispectrum level, for other statistics they have different properties.

Bringing these terms together, we have

$$\begin{aligned} \frac{\Phi^{NG}(\mathbf{k})}{3} = & - \left[1 + \frac{9p}{27} \right] K_A^I - \left[2 + \frac{60p}{27} \right] K_C^I \\ & + \left[1 + \frac{15p}{27} \right] [(1-t)K_B^I + tK_B^{II}] \\ & + \frac{p}{27} K_D^I - \frac{20p}{27} K_E^{II} + \frac{6p}{27} K_F^I + \frac{15p}{27} K_G^I, \quad (\text{A10}) \end{aligned}$$

or equivalently

$$\begin{aligned} \frac{\Phi^{NG}(\mathbf{x})}{3} = & - \left[1 + \frac{9p}{27} \right] \Phi^2 - \left[2 + \frac{60p}{27} \right] \frac{1}{3} \partial^{-2} [\partial \Phi \partial \Phi] \\ & + 2 \left[1 + \frac{15p}{27} \right] [(1-t) \partial^{-1} [\Phi \partial \Phi] + t \partial^{-2} [\Phi \partial^2 \Phi]] \\ & + \frac{p}{27} \partial^2 [\partial^{-1} \Phi \partial^{-1} \Phi] - \frac{20p}{27} \partial^{-1} [\partial^2 \Phi \partial^{-1} \Phi] \\ & - \frac{12p}{27} \partial [\Phi \partial^{-1} \Phi] + \frac{30p}{27} \partial \Phi \partial^{-1} \Phi. \quad (\text{A11}) \end{aligned}$$

These operators have been derived assuming a scale-invariant spectrum ($n_s = 1$) for simplicity. It is simple to adapt to a non-scale-invariant spectrum by replacing k with $P(k)^{-\frac{1}{3}}$. While this replacement alters the primordial bispectrum, the distortion is small, as $n_s \sim 1$, and it allows us to simply use non-scale-invariant simulations in our analysis (i.e., $n_s \neq 1$), as is required to be consistent with observations (Planck Collaboration VI 2020).

Appendix B Covariance Matrix Computation

The covariance matrix given in Equation (19) is composed of several terms: the Gaussian, connected, and supersample covariance terms.

B.1. Gaussian Covariance Matrix

The computation of the Gaussian covariance matrix for the power spectrum estimator, Equation (10), is given by

Feldman et al. (1994):

$$C_{ij}^{\text{Gaussian}} = \frac{2}{N_{\text{modes}}^2} \sum_{k_\alpha \in k_i} P(k_\alpha)^2 \delta_{k_i, k_j}, \quad (\text{B1})$$

where N_{modes} is the number of modes in the power spectrum binned. For the binned bispectrum estimator, the equivalent expression is

$$\begin{aligned} C^{\text{Gaussian}}(\hat{B}_{ijk}, \hat{B}_{ljk}) \\ = \frac{g_{ijk}^{llk}}{N_{\text{triplets}}^2} \sum P(k_\alpha) P(k_\beta) P(k_\gamma), \quad (\text{B2}) \end{aligned}$$

where N_{triplets} is the number of bispectrum configurations in the bin, the sum is over all bispectrum triplets, and g_{ijk}^{llk} selects only the diagonal elements of the covariance matrix and is 1, 2, or 6 depending on whether there are 3, 2, or 1 unique values in $\{ijk\}$.

B.2. Supersample Covariance

Generally, the supersample covariance contribution to the covariance matrix is given by

$$C_{X,Y}^{\text{SSV}}(k, k') = \sigma_W^2 \frac{\partial X(K)}{\partial \delta_b} \Big|_{\delta_b=0} \frac{\partial Y(K')}{\partial \delta_b} \Big|_{\delta_b=0}, \quad (\text{B3})$$

where

$$\sigma_b^2 = \int \frac{d^3k}{(2\pi)^3} \frac{W(\mathbf{k})}{V_W} \frac{W(\mathbf{k})}{V_W} P_L(k). \quad (\text{B4})$$

We use follow Li et al. (2014a) and use separate universe simulations to compute the responses to a long-wavelength mode (the derivative terms in Equation (B3)). For the power spectrum of global mean observables, this is given by

$$\frac{\partial P(K)}{\partial \delta_b} \Big|_{\delta_b=0} = P(k) + \frac{\partial P_{\text{SU}}(K)}{\partial \delta_b} - \frac{1}{3} \frac{\partial P(k)}{\partial \ln k}, \quad (\text{B5})$$

and for local power spectra observables we have an additional $-2P(k)$. For the bispectrum of global mean observables, this is

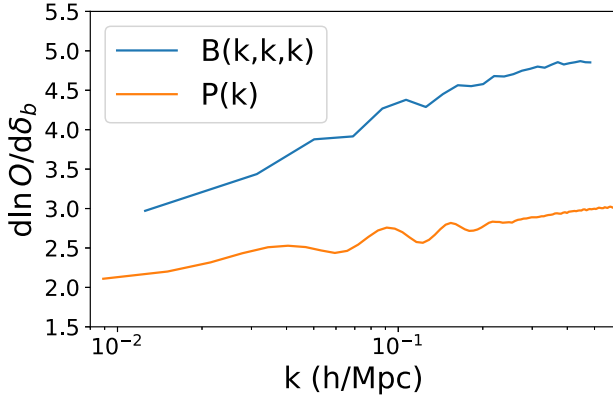


Figure 10. The response functions of the power spectrum and bispectrum to a long-wavelength mode. These have been computed using separate universe simulations as described in Li et al. (2014a).

(Chan et al. 2018)

$$\left. \frac{\partial B(k_1, k_2, k_3)}{\partial \delta_b} \right|_{\delta_b=0} = B(k_1, k_2, k_3) + \frac{\partial B_{\text{SU}}(k_1, k_2, k_3)}{\partial \delta_b} - \sum_i \frac{1}{3} \frac{\partial B(k_1, k_2, k_3)}{\partial \ln k_i}, \quad (\text{B6})$$

and for local mean observables there is an additional $-3B(k_1, k_2, k_3)$. In both cases, the derivative $\partial P_{\text{SU}}(k)/\partial \delta_b$ and $\partial B_{\text{SU}}(k_1, k_2, k_3)/\partial \delta_b$ are computed from the separate universe simulations using finite differences.

In Figure 10, we plot the response functions, and these show great agreement with the equivalent plots in Li et al. (2014a)

and Chan et al. (2018). Further, we find that the relative importance of the supersample covariance terms is very similar to that found in Chan et al. (2018), providing validation of our implementation.

Appendix C

Joint Constraints with Parameters beyond Λ CDM

In this appendix, we extend our analysis to include two common extensions: the dark energy equation of state parameter, w , and the sum of the masses of the neutrinos $\sum m_\nu$. We use the simulations described in Villaescusa-Navarro et al. (2020) to compute the derivatives with respect to w and the simulations in Hahn et al. (2020) for the neutrino mass constraints. For the neutrino mass, we use two modifications to our method: first, we compute the properties with respect to the total matter field (rather than just the dark matter field). This is trivial to do by including the neutrinos when gridding the simulation outputs. Second, we use the third-order difference method, Equation (4.5) in Hahn et al. (2020), to compute the derivatives. This is used because the neutrino mass is positive definite and thus we cannot use the central difference method used for the other parameters. Additionally, this higher-order method provides a more accurate estimate of the derivatives.

In Figure 11, we plot the joint constraints of all the cosmological and PNG parameters. The inclusion of w primarily degrades the constraints on PNG by decreasing the constraining power on the Λ CDM parameters. On the other hand, we find strong degeneracies between $\sum m_\nu$ and the f_{NL} parameters. This is not unexpected, given that the PNG parameters exhibit degeneracies with σ_8 and the impact of $\sum m_\nu$ is similar to σ_8 on these scales.

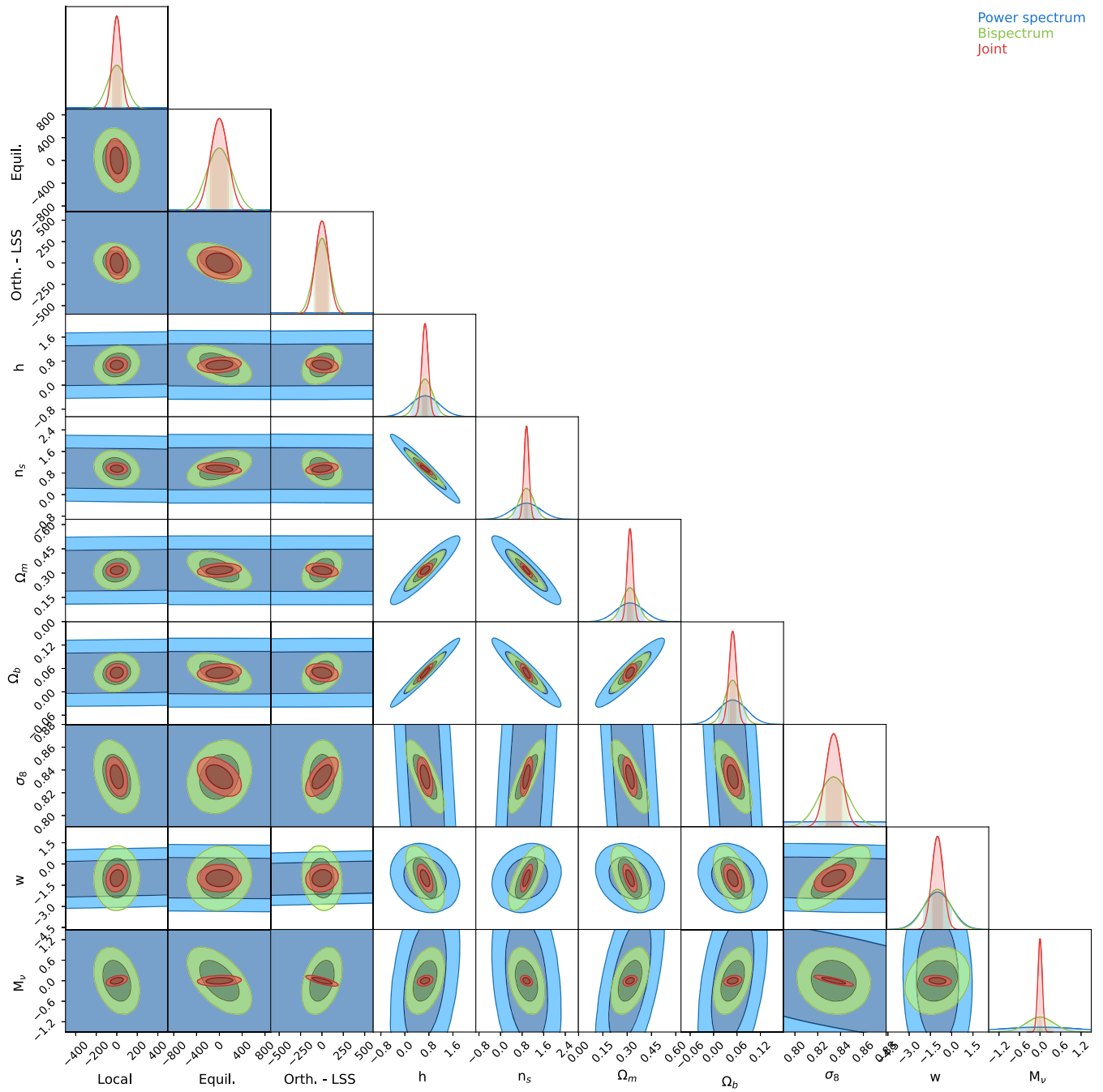


Figure 11. Joint constraints on the standard cosmological parameters, the equation of state of dark energy, w , the sum of the masses of the neutrinos, Σm_ν , and three shapes of primordial non-Gaussianity.

Appendix D

Impact of a Prior on the Power Spectrum Convergence

In Section 6, we found that the power spectrum constraints were not converged. As can be seen in Figure 6, the power spectrum constraints are increased by a factor of ~ 10 – 100 when marginalizing, and this large degeneracy is the reason for the lack of convergence. To test this idea, we perform our convergence test when imposing a Gaussian prior of width

$f_{\text{NL}} = 1000$ for all the PNG shapes. In Figure 12, we see the convergence—and in this case, we find that all derivatives are sufficiently converged. We show the parameter constraints obtained when including this prior in Figure 13. The only effect of the prior is for the power spectrum constraints, where it improves the constraints on almost all of the cosmological parameters by $\sim 20\%$, with the exception of s_8 , which is significantly improved.

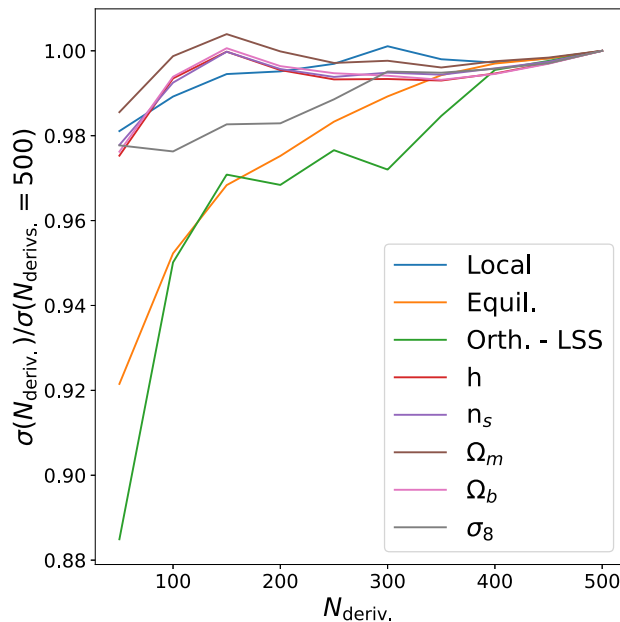


Figure 12. The convergence of the power spectrum measurements when using a Gaussian prior of width $f_{\text{NL}} = 1000$.

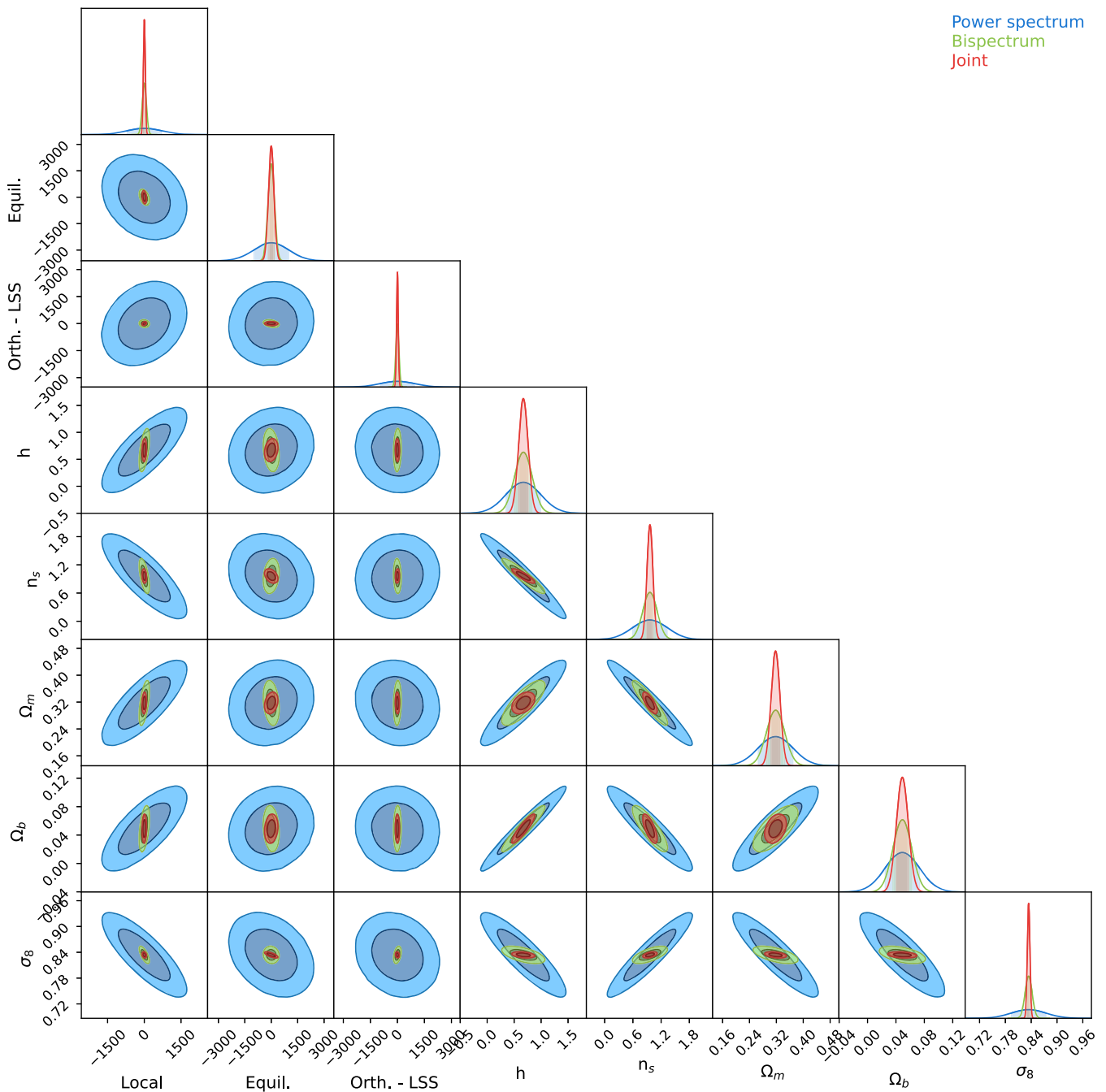


Figure 13. The cosmological constraints obtained when a Gaussian prior of width $\sigma(f_{\text{NL}}) = 1000$ is applied to the power spectrum.

ORCID iDs

William R Coulton <https://orcid.org/0000-0002-1297-3673>
 Drew Jamieson <https://orcid.org/0000-0001-5044-7204>
 Marco Baldi <https://orcid.org/0000-0003-4145-1943>
 Gabriel Jung <https://orcid.org/0000-0002-2133-6254>
 Benjamin D. Wandelt <https://orcid.org/0000-0002-5854-8269>

References

Abazajian, K. N., Adshead, P., Ahmed, Z., et al. 2016, arXiv:1610.02743
 Achúcarro, A., Biagetti, M., Braglia, M., et al. 2022, arXiv:2203.08128
 Ade, P., Aguirre, J., Ahmed, Z., et al. 2019, *JCAP*, 2019, 056
 Aitken, A. C., & Silverstone, H. 1942, *Proc. Math. Phys. Sci.*, 61, 186
 Alam, S., Ata, M., Bailey, S., et al. 2017, *MNRAS*, 470, 2617

Alishahiha, M., Silverstein, E., & Tong, D. 2004, *PhRvD*, 70, 123505
 Amendola, L., Appleby, S., Avgoustidis, A., et al. 2018, *LRR*, 21, 2
 Andrews, A., Jasche, J., Lavaux, G., & Schmidt, F. 2022, arXiv:2203.08838
 Arkani-Hamed, N., Creminelli, P., Mukohyama, S., & Zaldarriaga, M. 2004, *JCAP*, 2004, 001
 Baldauf, T., Seljak, U., & Senatore, L. 2011, *JCAP*, 2011, 006
 Barreira, A. 2019, *JCAP*, 2019, 008
 Barreira, A. 2020, *JCAP*, 2020, 031
 Baumann, D., & Green, D. 2022, *JCAP*, 2022, 51
 Bennett, C. L., Larson, D., Weiland, J. L., et al. 2013, *ApJS*, 208, 20
 Biagetti, M., Calles, J., Castiblanco, L., Cole, A., & Noreña, J. 2022, *JCAP*, 2022, 2
 Biagetti, M., Castiblanco, L., Noreña, J., & Sefusatti, E. 2022, *JCAP*, 2022, 009
 Biagetti, M., Cole, A., & Shiu, G. 2021, *JCAP*, 2021, 061
 Cabass, G., Ivanov, M. M., Philcox, O. H. E., Simonović, M., & Zaldarriaga, M. 2022a, *PhRvL*, 129, 021301

- Cabass, G., Ivanov, M. M., Philcox, O. H. E., Simonović, M., & Zaldarriaga, M. 2022b, *PhRvD*, **106**, 043506
- Carron, J. 2013, *A&A*, **551**, A88
- Castorina, E., Hand, N., Seljak, U., et al. 2019, *JCAP*, **2019**, 010
- Chan, K. C., & Blot, L. 2017, *PhRvD*, **96**, 023528
- Chan, K. C., Moradinezhad Dizgah, A., & Noreña, J. 2018, *PhRvD*, **97**, 043532
- Chen, X. 2010, *AdAst*, **2010**, 638979
- Chen, X., Huang, M.-x., Kachru, S., & Shiu, G. 2007, *JCAP*, **2007**, 002
- Cheung, C., Fitzpatrick, A. L., Kaplan, J., & Senatore, L. 2008a, *JCAP*, **2008**, 021
- Cheung, C., Fitzpatrick, A. L., Kaplan, J., Senatore, L., & Creminelli, P. 2008b, *JHEP*, **2008**, 014
- Chisari, N. E., Richardson, M. L. A., Devriendt, J., et al. 2018, *MNRAS*, **480**, 3962
- Coulton, W., Miranthis, A., & Challinor, A. 2022a, arXiv:2208.12270
- Coulton, W. R., Meerburg, P. D., Baker, D. G., et al. 2020, *PhRvD*, **101**, 123504
- Coulton, W. R., Villaescusa-Navarro, F., Jamieson, D., et al. 2022b, arXiv:2206.15450
- Creminelli, P., D'Amico, G., Musso, M., Noreña, J., & Trincherini, E. 2011, *JCAP*, **2011**, 006
- Creminelli, P., Nicolis, A., Senatore, L., Tegmark, M., & Zaldarriaga, M. 2006, *JCAP*, **2006**, 004
- Creminelli, P., & Zaldarriaga, M. 2004, *JCAP*, **2004**, 006
- Dalal, N., Doré, O., Huterer, D., & Shirokov, A. 2008, *PhRvD*, **77**, 123514
- d'Amico, G., Gleyzes, J., Kokron, N., et al. 2020, *JCAP*, **2020**, 005
- D'Amico, G., Lewandowski, M., Senatore, L., & Zhang, P. 2022, arXiv:2201.11518
- Darmois, G. 1945, *Revue de l'Institut International de Statistique/Review of the International Statistical Institute*, **13**, 9
- Davis, M., Efstathiou, G., Frenk, C. S., & White, S. D. M. 1985, *ApJ*, **292**, 371
- DESI Collaboration, A., Aghamousa, J., Aguilar, S., et al. 2016, arXiv:1611.00036
- Desjacques, V., Seljak, U., & Iliev, I. T. 2009, *MNRAS*, **396**, 85
- Doré, O., Bock, J., Ashby, M., et al. 2014, arXiv:1412.4872
- Dvali, G., Gruzinov, A., & Zaldarriaga, M. 2004, *PhRvD*, **69**, 023505
- Enríquez, M., Hidalgo, J. C., & Valenzuela, O. 2022, *JCAP*, **2022**, 048
- Fixsen, D. J., Weiland, J. L., Brodd, S., et al. 1997, *ApJ*, **490**, 482
- Feldman, H. A., Kaiser, N., & Peacock, J. A. 1994, *ApJ*, **426**, 23
- Fergusson, J. R., Regan, D. M., & Shellard, E. P. S. 2012, *PhRvD*, **86**, 063511
- Ferreira, P. G., Magueijo, J., & Górski, K. M. G. 1998, *ApJL*, **503**, L1
- Fisher, R. A. 1935, *J. R. Stat. Soc.*, **98**, 39
- Foreman, S., Coulton, W., Villaescusa-Navarro, F., & Barreira, A. 2020, *MNRAS*, **498**, 2887
- Frechet, M., & Frechet, M. 1943, *Revue de l'Institut International de Statistique/Review of the International Statistical Institute*, **11**, 182
- Friedrich, O., Uhlemann, C., Villaescusa-Navarro, F., et al. 2020, *MNRAS*, **498**, 464
- Giannantonio, T., Ross, A. J., Percival, W. J., et al. 2014, *PhRvD*, **89**, 023511
- Giusarma, E., Reyes Hurtado, M., Villaescusa-Navarro, F., et al. 2019, arXiv:1910.04255
- Green, D., & Porto, R. A. 2020, *PhRvL*, **124**, 251302
- Gualdi, D., Manera, M., Joachimi, B., & Lahav, O. 2018, *MNRAS*, **476**, 4045
- Gualdi, D., & Verde, L. 2020, *JCAP*, **2020**, 041
- Hahn, C., Villaescusa-Navarro, F., Castorina, E., & Scoccimarro, R. 2020, *JCAP*, **2020**, 040
- Hamilton, A. J. S., Rimes, C. D., & Scoccimarro, R. 2006, *MNRAS*, **371**, 1188
- Hartlap, J., Simon, P., & Schneider, P. 2007, *A&A*, **464**, 399
- Hill, J. C. 2018, *PhRvD*, **98**, 083542
- Ho, S., Agarwal, N., Myers, A. D., et al. 2015, *JCAP*, **2015**, 040
- Hockney, R. W., & Eastwood, J. W. 1981, *Computer Simulation Using Particles* (Boca Raton, FL: CRC Press)
- Ivanov, M. M., Simonović, M., & Zaldarriaga, M. 2020, *JCAP*, **2020**, 042
- Jing, Y. P. 2005, *ApJ*, **620**, 559
- Jung, G., Karagiannis, D., Liguori, M., et al. 2022, *ApJ*, **940**, 71
- Jung, G., Racine, B., & Tent, B. v. 2018, *JCAP*, **2018**, 047
- Kalaja, A., Meerburg, P. D., Pimentel, G. L., & Coulton, W. R. 2021, *JCAP*, **2021**, 050
- Karagiannis, D., Lazanu, A., Liguori, M., et al. 2018, *MNRAS*, **478**, 1341
- Karagiannis, D., Slosar, A., & Liguori, M. 2020, *JCAP*, **2020**, 052
- Kayo, I., Takada, M., & Jain, B. 2013, *MNRAS*, **429**, 344
- Kobayashi, Y., Nishimichi, T., Takada, M., & Miyatake, H. 2022, *PhRvD*, **105**, 083517
- Komatsu, E., Kogut, A., Nolta, M. R., et al. 2003, *ApJS*, **148**, 119
- Komatsu, E., Wandelt, B. D., Spergel, D. N., Banday, A. J., & Górski, K. M. G. 2002, *ApJ*, **566**, 19
- Lehners, J.-L. 2010, *AdAst*, **2010**, 903907
- Leistedt, B., Peiris, H. V., & Roth, N. 2014, *PhRvL*, **113**, 221301
- Lewis, A., Challinor, A., & Lasenby, A. 2000, *ApJ*, **538**, 473
- Li, Y., Hu, W., & Takada, M. 2014a, *PhRvD*, **89**, 083519
- Li, Y., Hu, W., & Takada, M. 2014b, *PhRvD*, **90**, 103530
- LoVerde, M., & Smith, K. M. 2011, *JCAP*, **2011**, 003
- Lucchin, F., & Matarrese, S. 1988, *ApJ*, **330**, 535
- Lyth, D., & Wands, D. 2002, *PhLB*, **524**, 5
- Lyth, D. H. D. H., & Riotto, A. A. 1999, *PhR*, **314**, 1
- Maldacena, J. 2003, *JHEP*, **2003**, 013
- Martin, J. 2016, *ApSSP*, **45**, 41
- Matarrese, S., Pilo, L., & Rollo, R. 2021, *JCAP*, **2021**, 062
- Meerburg, P. D., Green, D., Flauger, R., et al. 2019, *BAAS*, **51**, 107
- Meerburg, P. D., Münchmeyer, M., Muñoz, J. B., & Chen, X. 2017, *JCAP*, **2017**, 050
- Meerburg, P. D., van der Schaar, J. P., & Jackson, M. G. 2010, *JCAP*, **2010**, 001
- Mueller, E.-M., Rezaie, M., Percival, W. J., et al. 2021, arXiv:2106.13725
- Neyrinck, M. C., & Szapudi, I. 2007, *MNRAS*, **375**, L51
- Pajer, E., Schmidt, F., & Zaldarriaga, M. 2013, *PhRvD*, **88**, 083502
- Planck Collaboration I 2020, *A&A*, **641**, A1
- Planck Collaboration IX 2020, *A&A*, **641**, A9
- Planck Collaboration VI 2020, *A&A*, **641**, A6
- Planck Collaboration XXIV 2014, *A&A*, **571**, A24
- Pullen, A. R., & Hirata, C. M. 2013, *PASP*, **125**, 705
- Rezaie, M., Ross, A. J., Seo, H.-J., et al. 2021, *MNRAS*, **506**, 3439
- Rimes, C. D., & Hamilton, A. J. S. 2005, *MNRAS*, **360**, L82
- Rimes, C. D., & Hamilton, A. J. S. 2006, *MNRAS*, **371**, 1205
- Ross, A. J., Percival, W. J., Carnero, A., et al. 2013, *MNRAS*, **428**, 1116
- Schmidt, F., & Kamionkowski, M. 2010, *PhRvD*, **82**, 103002
- Schmittfull, M. M., Regan, D. M., & Shellard, E. P. S. 2013, *PhRvD*, **88**, 063512
- Scoccimarro, R. 2000, *ApJ*, **544**, 597
- Scoccimarro, R., Hui, L., Manera, M., & Chan, K. C. 2012, *PhRvD*, **85**, 083002
- Sefusatti, E. 2009, *PhRvD*, **80**, 123002
- Sefusatti, E., Crocce, M., & Desjacques, V. 2010, *MNRAS*, **406**, 1014
- Sefusatti, E., Crocce, M., Scoccimarro, R., & Couchman, H. M. P. 2016, *MNRAS*, **460**, 3624
- Sellentini, E., & Heavens, A. F. 2018, *MNRAS*, **473**, 2355
- Senatore, L., Smith, K. M., & Zaldarriaga, M. 2010, *JCAP*, **2010**, 028
- Slosar, A., Ahmed, Z., Alonso, D., et al. 2019, *BAAS*, **51**, 53
- Slosar, A., Hirata, C., Seljak, U., Ho, S., & Padmanabhan, N. 2008, *JCAP*, **2008**, 031
- Springel, V. 2005, *MNRAS*, **364**, 1105
- Sugiyama, N. S., Saito, S., Beutler, F., & Seo, H.-J. 2020, *MNRAS*, **497**, 1684
- Takahashi, R., Yoshida, N., Takada, M., et al. 2009, *ApJ*, **700**, 479
- Tegmark, M., Taylor, A. N., & Heavens, A. F. 1997, *ApJ*, **480**, 22
- Tomlinson, J., Jeong, D., & Kim, J. 2019, *AJ*, **158**, 116
- Uhlemann, C., Pajer, E., Pichon, C., et al. 2018, *MNRAS*, **474**, 2853
- Valageas, P. 2002, *A&A*, **382**, 431
- Villaescusa-Navarro, F., Hahn, C., Massara, E., et al. 2020, *ApJS*, **250**, 2
- Wagner, C., Verde, L., & Boubekeur, L. 2010, *JCAP*, **2010**, 022