Quijote-PNG: Optimizing the summary statistics to measure Primordial non-Gaussianity

Gabriel Jung, Andrea Ravenni, Michele Liguori, Marco Baldi, William R Coulton, Francisco Villaescusa-Navarro, Benjamin D Wandelt

The Astrophysical Journal, 976:109 (13pp), 2024 November 20 https://doi.org/10.3847/1538-4357/ad83bd

Quijote-PNG: Optimizing the Summary Statistics to Measure Primordial Non-Gaussianity

Gabriel Jung[1,14], Andrea Ravenni[2,3,4,14], Michele Liguori[2,3,5], Marco Baldi[6,7,8], William R. Coulton[9,10], Francisco Villaescusa-Navarro[11,12], and Benjamin D. Wandelt[11,13]
[1] Université Paris-Saclay, CNRS, Institut d'Astrophysique Spatiale, 91405, Orsay, France
[2] Dipartimento di Fisica e Astronomia "G. Galilei," Università degli Studi di Padova, via Marzolo 8, I-35131, Padova, Italy
[3] INFN, Sezione di Padova, via Marzolo 8, I-35131, Padova, Italy
[4] ICC, University of Barcelona, IEEC-UB, Martí i Franquès, 1, E-08028 Barcelona, Spain
[5] Dipartimento di Fisica, Università degli Studi di Trento, Via Sommarive 14, I-38123 Povo (TN), Italy
[6] Dipartimento di Fisica e Astronomia, Alma Mater Studiorum—University of Bologna, Via Piero Gobetti 93/2, 40129 Bologna BO, Italy
[7] INAF—Osservatorio Astronomico di Bologna, Via Piero Gobetti 93/3, 40129 Bologna BO, Italy
[8] INFN—Istituto Nazionale di Fisica Nucleare, Sezione di Bologna, Viale Berti Pichat 6/2, 40127 Bologna BO, Italy
[9] Kavli Institute for Cosmology Cambridge, Madingley Road, Cambridge CB3 0HA, UK
[10] DAMTP, Centre for Mathematical Sciences, University of Cambridge, Wilberforce Road, Cambridge CB3 OWA, UK
[11] Center for Computational Astrophysics, Flatiron Institute, 162 5th Avenue, New York, NY 10010, USA
[12] Department of Astrophysical Sciences, Princeton University, 4 Ivy Lane, Princeton, NJ 08544, USA
[13] Sorbonne Université, CNRS, UMR 7095, Institut d'Astrophysique de Paris, 98 bis bd Arago, 75014 Paris, France
Received 2024 March 20; revised 2024 September 25; accepted 2024 October 4; published 2024 November 15

Abstract

We apply a suite of different estimators to the QUIJOTE-PNG halo catalogs to find the best approach to constrain Primordial non-Gaussianity (PNG) at nonlinear cosmological scales, up to $k_{\rm max} = 0.5\,h\,{\rm Mpc}^{-1}$. The set of summary statistics considered in our analysis includes the power spectrum, bispectrum, halo mass function, marked power spectrum, and marked modal bispectrum. Marked statistics are used here for the first time in the context of the PNG study. We perform a Fisher analysis to estimate their cosmological information content, showing substantial improvements when marked observables are added to the analysis. Starting from these summaries, we train deep neural networks to perform likelihood-free inference of cosmological and PNG parameters. We assess the performance of different subsets of summary statistics; in the case of $f_{\rm NL}^{\rm equil}$, we find that a combination of the power spectrum and a suitable marked power spectrum outperforms the combination of power spectrum and bispectrum, the baseline statistics usually employed in PNG analysis. A minimal pipeline to analyze the statistics we identified can be implemented either with our ML algorithm or via more traditional estimators, if these are deemed more reliable.

Unified Astronomy Thesaurus concepts: N-body simulations (1083); Cosmological parameters from large-scale structure (340); Non-Gaussianity (1116)

1. Introduction

Primordial non-Gaussianity (PNG) provides a potentially powerful tool to discriminate between different early Universe scenarios, and therefore its investigation plays an important role in observational cosmology.

However, extracting PNG information is a task made significantly difficult by the smallness of the expected signal, which at low redshifts is several orders of magnitude below that generated by nonlinear gravitational evolution of cosmological perturbations. For this reason, most observational studies of primordial non-Gaussian parameters have so far focused on linear cosmological probes, such as the cosmic microwave background (CMB; Planck Collaboration 2020), or the galaxy power spectrum and bispectrum on large scales (G. Cabass et al. 2022a, 2022b; G. D'Amico et al. 2022; M. S. Cagliari et al. 2023; M. M. Ivanov et al. 2024).

While challenging, as we have just stressed, a large scale structure (LSS) analysis at nonlinear scales could potentially be very rewarding: most of the PNG constraining power comes in fact from the cosmological perturbation bispectrum and a simple mode counting argument suggests that large improvements could be achieved in this regime, provided we are able to at least partially clean the total non-Gaussianity signal from late-time nonlinear contributions. In this work—which is connected to a series of previous studies in the QUIJOTE-PNG series (W. R. Coulton et al. 2023a; G. Jung et al. 2023a, 2023b)—we investigate this possibility by performing a thorough analysis of the dark matter halo field in *N*-body simulations with PNG initial conditions, testing scales up to $k_{\rm max} = 0.5\,h\,{\rm Mpc}^{-1}$.

The study of cosmological information at nonlinear scales, in a more general context than just PNG analysis, is actually a research line that has recently received considerable attention. The reason for this growing interest is likely twofold. On the one hand, a large amount of data from upcoming galaxy surveys is going to make small scales observably accessible with high precision. On the other hand, the past few years have seen significant methodological developments in cosmological data analysis, like field-level inference, whether it is performed using perturbative (F. Schmidt 2021; D. Baumann & D. Green 2022; G. Cabass et al. 2024) or

[14] Equal contributions.

Bayesian (J. Jasche & B. D. Wandelt 2013; A. Andrews et al. 2023) forward models, or via machine learning using simulations (M. Ntampaka et al. 2019; D. Ribli et al. 2019; T. L. Makinen et al. 2022; P. Villanueva-Domingo & F. Villaescusa-Navarro 2022; N. S. M. de Santi et al. 2023; A. Roncoli et al. 2023; H. Shao et al. 2023; P. Lemos et al. 2024). This kind of simulation-based approach is also explored with data compressed into suitable sets of summary statistics (J. Alsing et al. 2018, 2019; N. Jeffrey et al. 2021; C. Hahn et al. 2023a, 2023b, 2023c, 2024; M. Gatti et al. 2024; J. Hou et al. 2024; M. M. Ivanov et al. 2024; B. Régaldo-Saint Blancard et al. 2024; B. Tucci & F. Schmidt 2024).

Here, we take the latter approach to the problem, first compressing the data into a set of predetermined summary statistics. While potentially leading to some loss of information with respect to a full field-level analysis, this procedure presents some advantages, as it is less computationally demanding, potentially easier to implement when analyzing actual data and, above all, it often leads to results that are more amenable to a clear physical interpretation. The feature of interpretability may also aid in separating effects due to systematics from those due to the signal under study. Beyond quantifying the amount of information on PNG contained in different summary statistics computed from the halo density field at late times, the main goal of our work is to determine a suitable subset, which reaches the optimal compromise between being informative and relatively easy to analyze.

In previous works, we started pursuing this program by considering different combinations of the halo power spectrum, bispectrum, and halo mass function. Here, we start by extending this analysis with the inclusion of additional summaries, the marked power spectrum and marked bispectrum. These marked statistics provide a flexible and easy way to measure weighted combinations of n-point correlation functions and previous studies showed that they are able to tightly constrain neutrino masses at nonlinear scales (O. H. E. Philcox et al. 2020; E. Massara et al. 2021, 2023). They are therefore a natural option to consider also in a PNG analysis and we will indeed show in this paper that they can provide a powerful tool to constrain PNG parameters.

Besides adding new summaries, the present work also contains some significant methodological extensions of our previous studies. In previous works (W. R. Coulton et al. 2023a; G. Jung et al. 2023a, 2023b), we used the Fisher matrix formalism to provide figures of merit for the various statistics. These were calculated, at fully nonlinear scales for a fiducial cosmology, by numerically evaluating derivatives and covariance matrices through Monte Carlo averaging of tens of thousands of N-body realizations of the halo field. The Fisher matrices—obtained following the algorithm by W. R. Coulton & B. D. Wandelt (2023) based on score compression into a minimal set of summaries—were also used to build and test quadratic estimators, which are nearly optimal for parameter values near the fiducial cosmology. One drawback of this approach is that it makes the exploration of a wide range of parameter values very computationally demanding, as it requires the production of new sets of tens of thousands of simulations, for many different choices of fiducial parameters. We try in this work to overcome this limitation by using deep neural networks (NN), trained on a suite of simulations whose parameters are arranged in a Latin-hypercube, to map our summaries directly into the final parameters, without having to explicitly evaluate any covariance or make any assumption about the likelihood function. We discuss a first application of our pipeline, mainly aimed at a detailed comparison of many NG statistics, in order to find their optimal combination in terms of PNG sensitivity and simplicity of implementation. This is a first step in the direction of future applications to PNG parameter inference on real data from galaxy surveys, which will require the use of realistic galaxy mocks.

The plan of the paper is as follows. In Section 2 we briefly describe the QUIJOTE-PNG simulation suite that was used to calculate the Fisher matrices, train the networks, and produce the final forecasts. In Section 3 we introduce the summary statistics considered in our analysis and discuss the methods used to extract them from the data. In Section 4 we discuss the implementation of the NNs and the metrics that we use to assess their performance. In Section 5 we show our results, which include Fisher forecasts and a comparison of different data preprocessing methodologies, followed by NN results for many different combinations of summaries. In Section 6 we draw our final conclusions.

2. Simulations

The analyses presented in this paper are based on the QUIJOTE (F. Villaescusa-Navarro et al. 2020) and QUIJOTE-PNG (W. R. Coulton et al. 2023b) sets of simulations.

These are dark matter only N-body simulations of volume $1\,h^{-3}\,\mathrm{Gpc}^3$, containing 512^3 particles each, and run using the TreePM code GADGET-III from initial conditions generated at $z = 127$ by the codes 2LPTIC (M. Crocce et al. 2006) and 2LPTPNG (R. Scoccimarro et al. 2012; W. R. Coulton et al. 2023b),[15] for the simulations without and with PNG, respectively. We focus on dark matter halos, which are identified in each simulation by the standard friends-of-friends algorithm (M. Davis et al. 1985) by setting the linking length parameter to $b = 0.2$ and considering halos with more than 20 dark matter particles.

We mainly use a set of 1000 simulations with varying amount of equilateral PNG, with $f_{\rm NL}^{\rm equil} \in [-600, 600]$ and varying cosmological parameters, where the parameters are distributed in a Latin-hypercube (LH). We also work with the 15,000 QUIJOTE simulations at a fiducial cosmology compatible with Planck CMB observations and without PNG and, to compute Fisher forecasts and compressed statistics, we use additional sets of 500 simulations, in which one parameter has been slightly displaced with respect to its fiducial value. We perform extra tests on the original QUIJOTE LH (2000 simulations with varying cosmological parameters and no PNG) and an LH with fixed cosmological parameters and varying local PNG, with $f_{\rm NL}^{\rm local} \in [-300, 300]$. The main characteristics of all these simulations are given in Table 1. In particular, the parameter uniform distributions within the specified boundaries of the LH act as our chosen prior in the moment network analysis. For a detailed description of the three PNG shapes considered here, see Section 2 of W. R. Coulton et al. (2023b).[16] We release the local and equilateral LH in complement to this work, which makes all simulations used here publicly available.[17]

[15] https://github.com/dsjamieson/2LPTPNG

[16] The orthogonal shape we use is the orthogonal-LSS one, which does not have a scale-dependent bias component.

[17] https://quijote-simulations.readthedocs.io/en/latest/png.html

Table 1
Parameters of the QUIJOTE and QUIJOTE-PNG Halo Catalogs Used in This Work

	$N_{\rm sims}$	σ_8	$\Omega_{\rm m}$	$\Omega_{\rm b}$	n_s	h	$f_{\rm NL}^{\rm local}$	$f_{\rm NL}^{\rm equil}$	$f_{\rm NL}^{\rm ortho}$
Fiducial	15000	0.834	0.3175	0.049	0.9624	0.6711	0	0	0
Displaced	500	±0.015	±0.01	±0.002	±0.02	±0.02	±100	±100	±100
LH $f_{\rm NL}^{\rm local}$	1000	0.834	0.3175	0.049	0.9624	0.6711	[−300, 300]	0	0
LH $f_{\rm NL}^{\rm equil}$	1000	[0.6, 1.0]	[0.1, 0.5]	0.049	[0.8, 1.2]	[0.5, 0.9]	0	[−600, 600]	0
LH QUIJOTE	2000	[0.6, 1.0]	[0.1, 0.5]	[0.03, 0.07]	[0.8, 1.2]	[0.5, 0.9]	0	0	0

3. Statistics

In this work, we test the performance of a variety of summary statistics calculated on the QUIJOTE-PNG halo catalogs in redshift space.

First, we consider the two- and three-point correlation functions of the halo density field in Fourier space $\delta(\boldsymbol{k})$, namely the power spectrum and bispectrum. We use the same estimation pipeline as in G. Jung et al. (2023a, 2023b). The halo power spectrum is estimated up to $k_{\rm max} = 0.5\,h\,{\rm Mpc}^{-1}$, using bins of size k_f (fundamental mode of the grid) and considering halos of mass above $M_{\rm min} = 3.2 \times 10^{13}\,M_\odot/h$. The halo bispectrum is computed up to the same nonlinear scales, using a modal estimator (J. R. Fergusson et al. 2010, 2012a, 2012b; M. M. Schmittfull et al. 2013). It simply consists on fitting well-chosen templates to the data, and it has been shown to be extremely efficient to compress the LSS bispectrum (J. Byun et al. 2021; J. Byun & E. Krause 2023; G. Jung et al. 2023a), with respect to a standard binned approach. Only a relatively small number of modes, typically less than 100, based on polynomial and tree-level matter bispectrum functions, are necessary to contain the full bispectral information up to $k_{\rm max} = 0.5\,h\,{\rm Mpc}^{-1}$.

Second, we study the marked (halo) power spectrum and bispectrum. These marked statistics are computed using the same pipeline as the standard power spectrum and bispectrum above, with an extra initial step of weighting the density field. For this weighting operation, we use the mark of M. White (2016),

$$m(\boldsymbol{x};\,R,\,p,\,\delta_s) = \left[\frac{1+\delta_s}{1+\delta_s+\delta_R(\boldsymbol{x})}\right]^p, \tag{1}$$

where $\delta_R(\boldsymbol{x})$ is the local density field, computed by smoothing the density field with a top-hat filter at scale R. The two other parameters, δ_s and p, can be chosen to modify the impact of $\delta_R(\boldsymbol{x})$ on the mark (sensitivity and enhancement of low/high density regions). As shown in E. Massara et al. (2021, 2023), measuring the power spectrum of such marked density field gives access to new cosmological information with respect to the power spectrum of the standard field, which comes in fact from higher order statistics. Here, we go one step further and also consider the bispectrum of the marked field. Due to the larger computational time required to estimate (marked) bispectra than their power spectrum counterparts, we restrict our analysis to the four different marks defined by the following choice of parameters, $R = [30, 25, 20, 30]\,h^{-1}$ Mpc, $p = [1, 1, 1, 1]$ and $\delta_s = [0.10, 0.25, 0.50, 0.50]$, and which were identified in E. Massara et al. (2023) as giving the most stringent constraints on cosmological parameters from galaxy catalogs constructed from the QUIJOTE N-body simulations. To construct the density fields of the QUIJOTE-PNG simulations, and to compute the corresponding marks, we use the PYLIANS3[18] library.

Note that the results presented in this paper use only the monopole of these different summary statistics, as we have verified that including the quadrupole of the standard and marked power spectra has only a negligible impact on constraints when considering jointly several of these observables.

Finally, we also include the halo mass function (HMF) in our analyses, as it was shown to contain significant information about PNG in G. Jung et al. (2023b). Here, the HMF is computed using 13 logarithmic mass bins between approximately 4.0×10^{13} and $4.6 \times 10^{15}\,M_\odot/h$.

Beyond these different summary statistics that will constitute our baseline analysis in Section 5, we also use compressed combinations of them calculated with

$$\tilde{s}_i = \left(\frac{\partial \bar{\boldsymbol{s}}}{\partial \theta_i}\right)_* \boldsymbol{C}_*^{-1}(\boldsymbol{s} - \bar{\boldsymbol{s}}_*), \tag{2}$$

where $\boldsymbol{s}$ is a chosen set of summary statistics of expected mean $\bar{\boldsymbol{s}}$ and covariance $\boldsymbol{C}$, and the subscript * denotes quantities evaluated at a chosen fiducial cosmology. This expression results in one compressed mode $\tilde{s}_i$ per parameter of interest θ_i and has been shown to be optimal (A. F. Heavens et al. 2000; J. Alsing & B. Wandelt 2018), i.e., preserving the full information about $\boldsymbol{\theta}$, if the summary statistics follow a Gaussian distribution where only the mean depends on parameters.

In what follows, covariances and derivatives are evaluated at the QUIJOTE fiducial cosmology (see Table 1). We use the set of 15,000 fiducial simulations, applying the Hartlap correction factor (J. Hartlap et al. 2007) to obtain unbiased estimates of the inverse covariances, and the couples of 500 displaced simulations to compute derivatives by finite difference.

4. Methods

4.1. Moment Network

As a way to quickly explore different summary statistics combinations while also covering a wide range of parameters, we train fully connected NNs to perform likelihood-free inference on different summary statistics, using the moment network methodology (N. Jeffrey & B. D. Wandelt 2020). These NNs will output two numbers, $\hat{\theta}$ and $\hat{\sigma}$ for each target parameter θ; the first being the mean and the second being the standard deviation of the marginalized posterior. To do so, we

[18] https://github.com/franciscovillaescusa/Pylians3

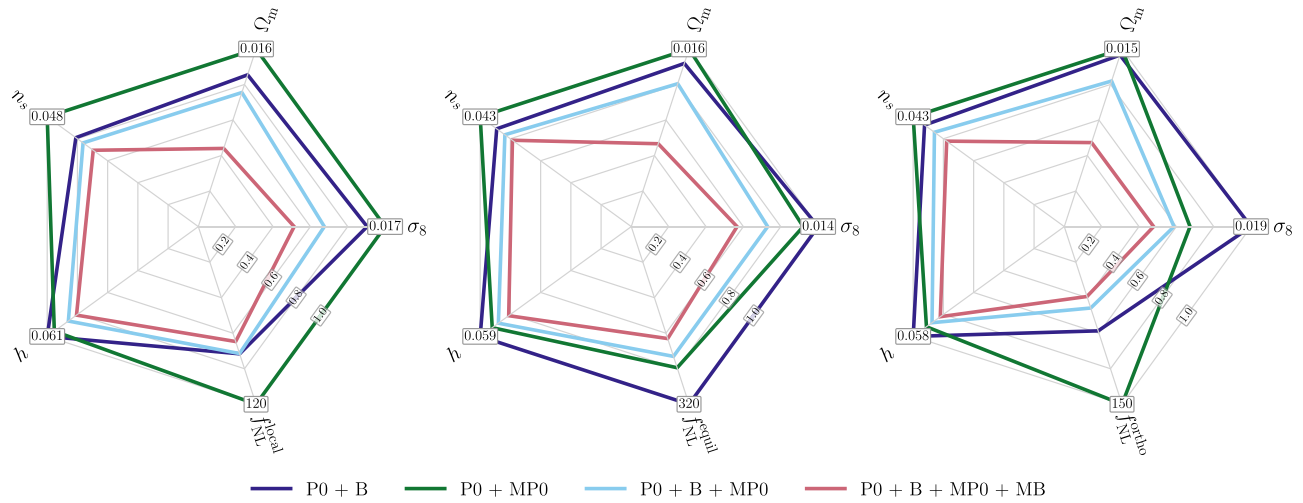

Figure 1. The 1σ Fisher error bars from different combinations of summary statistics measured in the QUIJOTE halo catalogs at $z = 0$, after marginalizing over $M_{\rm min}$. From left to right we consider the local, equilateral, or orthogonal PNG shapes jointly with cosmological parameters. The dark blue line corresponds to power spectrum + bispectrum (P0 + B) constraints, the green line to power spectrum + marked power spectrum (P0 + MP0), both are combined to obtain the light blue lines (P0 + B + MP0), and finally the red lines also include the marked bispectrum (P0 + B + MP0 + MB). The bold values are the largest 1σ Fisher error bars obtained for each parameter, used as normalization for the others. Marked statistics contain new information with respect to their standard counterparts on all parameters, which is significant for $f_{\rm NL}^{\rm equil}$, $f_{\rm NL}^{\rm ortho}$, σ_8, and $\Omega_{\rm m}$.

use the loss from, e.g., F. Villaescusa-Navarro et al. (2022)

$$\mathcal{L}_{\rm logMSE} = \sum_{i\in{\rm pars.}} \log\left[\sum_{j\in{\rm batch}} \Delta_{i,j}^2\right] + \sum_{i\in{\rm pars.}} \log\left[\sum_{j\in{\rm batch}} (\Delta_{i,j}^2 - \hat{\sigma}_{i,j}^2)^2\right], \quad (3)$$

where $\Delta_{i,j} \equiv \theta_{i,j} - \hat{\theta}_{i,j}$. The logarithms have been introduced to make both terms on the RHS of the same order of magnitude. Using this loss, it is guaranteed that the output of the network represents the first two moments of the posterior without making assumptions about its shape. Further details about the NNs architecture and training are provided in Appendix A.

4.2. Evaluating Performance

In our analysis, we consider several indicators to monitor the quality of the moment predictions.

First of all, for each parameter we calculate the coefficient of determination

$$R^2(\hat{\theta}) \equiv 1 - \frac{\sum_i (\theta_i - \hat{\theta}(x_i))^2}{\sum_i (\theta_i - \bar{\theta})^2}, \quad (4)$$

where i runs over simulations in the test set, θ_i is the input (true) parameter for the ith simulation, $\bar{\theta}$ is the average of the true parameter over the entire test set and $\hat{\theta}(x_i)$ is the posterior mean estimate, extracted from the ith simulation. We notice that $R^2 = 1$ if the true parameters are exactly recovered, whereas $R^2 = 0$ if the average value is always used as a prediction. Let us also stress that, despite the symbol used, R^2 can be negative if the estimator performs worse than just using the average value.

The coefficient of determination informs us of the quality of the posterior mean estimates, while a different metric needs to be used to monitor the second moment. Separately for each parameter again, we also calculate the coefficient

$$\chi^2(\hat{\theta}, \hat{\sigma}) \equiv \frac{1}{N}\sum_i \frac{(\theta_i - \hat{\theta}(x_i))^2}{\hat{\sigma}^2(x_i)}, \quad (5)$$

where $\hat{\sigma}(x_i)$ is the standard deviation prediction based on the ith simulation statistics, and N is the number of simulations in the considered set. This estimator is used to characterize the accuracy of the errors: the closer to 1, the more calibrated they are. For each trained network, we calculate χ^2 using the simulations in the validation set and discard all those where $|\chi^2 - 1| > 0.5$. Instead, all the instances of χ^2 shown below are calculated using the simulations in the test set.

5. Results

5.1. Fisher Forecasts

We start by evaluating the information content of the different statistics presented in Section 3 by considering the fiducial parameter values summarized in Table 1 and adopting the Fisher matrix formalism.

We use the combined Fisher estimator of W. R. Coulton & B. D. Wandelt (2023) to obtain unbiased results with the limited number of simulations at our disposal. For the details of the implementation, we refer the reader to G. Jung et al. (2023b), where we already applied this method to the same halo power spectra, bispectra, and HMF.

In Figure 1 we compare the 1σ Fisher error bars on cosmological parameters and PNG amplitudes for different combinations of summary statistics. An important result is that adding the marked power spectrum information to a standard power spectrum and bispectrum analysis improves the constraints on all parameters, except $f_{\rm NL}^{\rm local}$. This effect is the strongest for PNG of the equilateral and orthogonal types (more than 20% decrease), as well as σ_8 (close to 40%). A further gain is possible by including the marked bispectrum as well, where we even see a 10% improvement on $f_{\rm NL}^{\rm local}$, almost 50% for σ_8 and $\Omega_{\rm m}$, and close to 40% on $f_{\rm NL}^{\rm equil}$ and $f_{\rm NL}^{\rm ortho}$. Complementary analyses, including a study of the numerical

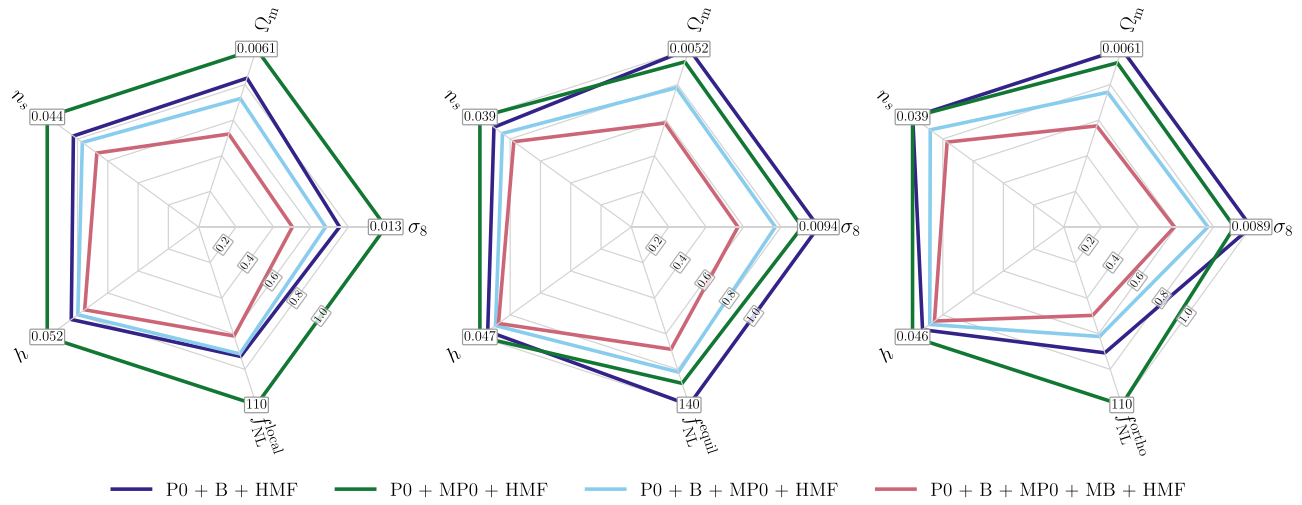

Figure 2. Same as Figure 1, adding the information of the HMF to every combination of summary statistics, confirming that improvements coming from the HMF or marked statistics have different origins.

stability of our results, 2D marginalized constraints, and Fisher forecasts focusing only on cosmological parameters are given in Appendix B.

In G. Jung et al. (2023b), it was shown that the HMF can also bring the same order of improvement on several parameters, especially $f_{\rm NL}^{\rm equil}$. In Figure 2, we verify that these improvements are in fact mostly independent from each other as the HMF and marked statistics bring complementary information.

Another interesting result shown in Figure 1 is that, in the case of equilateral PNG, a combination of the marked power spectrum and power spectrum performs slightly better than the standard power spectrum and bispectrum analysis. This is not the case for the other two shapes, noting, however, that the results are much better than they would be in a power-spectrum-only analysis (see, for example, the comparisons in G. Jung et al. 2023a), except for $f_{\rm NL}^{\rm local}$ itself where the constraint comes from the usual scale-dependent bias term of the power spectrum. This is a good indication that the marked power spectrum is a good alternative to the bispectrum in the search of PNG due to its simplicity of estimation with respect to the bispectrum. A comprehensive analysis should leverage both methods as they provide statistically independent validation of each other's result, and ultimately combine the two as, admittedly, the joint analysis can provide bounds a few $\times 10\%$ tighter than either one for most of the parameters.

In these analyses, we use jointly the four sets of marked statistics, defined by parameters $\{R, p, \delta_s\}$ given in Section 3. However, a large part of the improvement obtained by including marked statistics is already present when considering only one choice of mark, with a small dependence of the mark defining parameters. For example, the one with the smallest smoothing scale ($R = 20$, $p = 1$, $\delta_s = 0.5$) gives results as good as those of the three others combined, only slightly below (error bars a few percent larger at most) the case where the four marks are used jointly. This is why in many analyses below, where keeping the number of summary statistics as low as possible is important, we will focus on this specific choice of mark. Note also that in principle, it should be possible to improve the constraints even further by optimizing the parameters of the mark, for example, by exploring a wide range of them at the power spectrum level, before estimating the more computationally demanding bispectrum. By comparison to other analyses based on marked power spectrum (E. Massara et al. 2021, 2023), we do not expect any significant difference with the results reported here.

5.2. Neural Network Performance

As mentioned earlier, the main goal of this work is comparing and combining the various summary statistics described in Section 3, over a wide parameter range, by relying on a likelihood-free inference approach based on NNs. The next few sections are devoted to illustrate our analysis in detail and to discuss several tests aimed at its validation and interpretation.

Our main results are summarized in Figures 3 and 4. The former shows the accuracy of the mean parameter estimates as measured by the R^2 metric and the quality of the their uncertainty calibrations via the χ^2 metric, measured on the test set. The latter shows the comparison of the standard deviation estimated in the fiducial via Fisher forecast and using the moment network. As pointed out in Section 4.1, lower R^2 will correspond to larger standard deviations for a fixed χ^2, this being one if the uncertainties are estimated correctly.

Since some of the summary statistics are strongly correlated, and the training set we have at hand is limited in size, it is not surprising to sometimes see drops in accuracy in Figure 3 as more observables are added to the analysis. A larger number of input features usually requires bigger NN models, which in turn would require a larger training set. If the features one adds are highly informative, the accuracy of the model can still improve. If the new features are highly correlated with the ones already present, the amount of relevant information added may not counterbalance the worse training, and the net result is a drop in accuracy. The same applies to the standard deviations shown in Figure 4, which in some instances increase when more summary statistics are added. Moreover, to correctly interpret the results in Figures 3 and 4, we have to keep in consideration the prior we have imposed. Specifically, the combination of power spectrum and bispectrum is known to be informative about $f_{\rm NL}^{\rm equil}$, but the associated standard deviation, analyzing a volume of $1\,h^{-3}\,{\rm Gpc}^3$, is larger than the chosen prior.[19]

[19] This is in fact qualitatively consistent with the analysis of BOSS data (e.g., G. Cabass et al. 2022a), which has a larger volume than the one considered here.

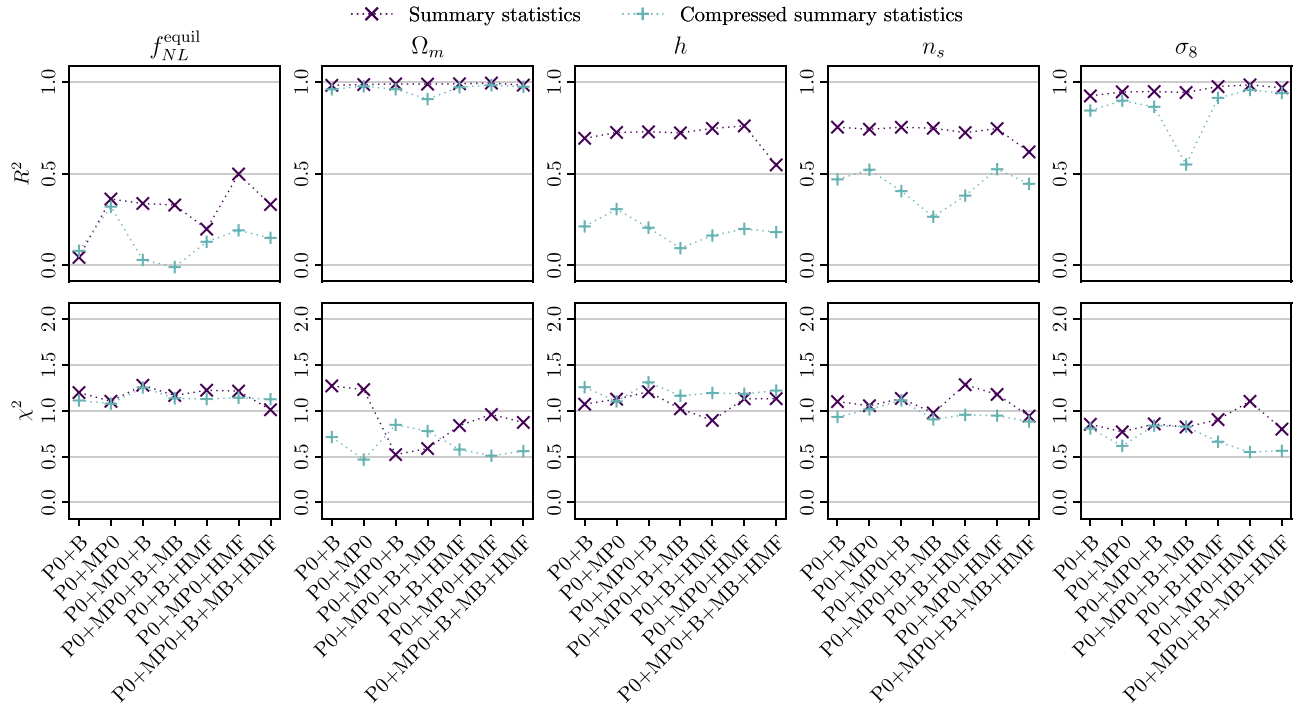

Figure 3. Accuracy of the five $f_{\rm NL}^{\rm equil}$ LH parameter predictions using power spectrum (P0), marked power spectrum (MP0), bispectrum (B), marked bispectrum (MB), and halo mass function (HMF), each column referring to a specific combination. The input of the NNs are the summary statistics either used as-is (purple cross markers) or compressed (light blue plus markers).

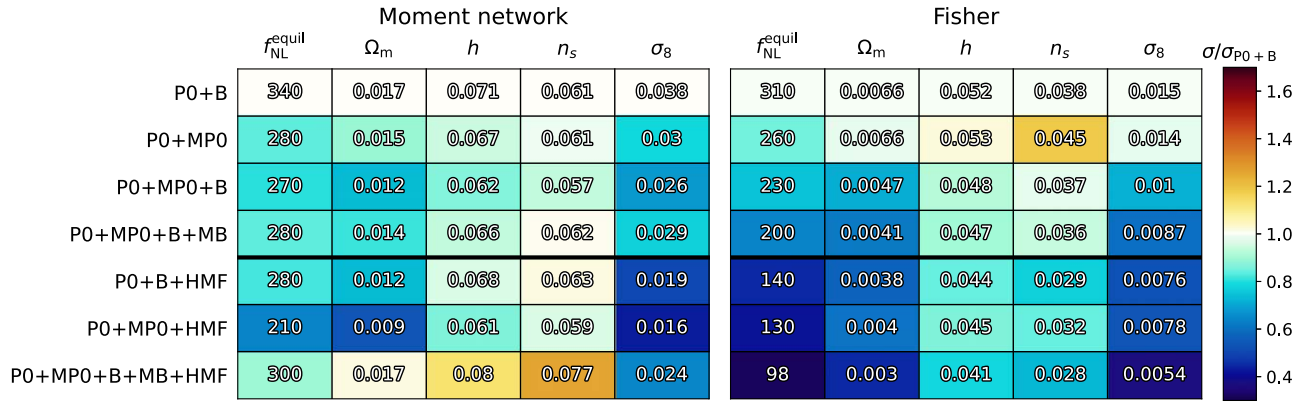

Figure 4. Standard deviation predictions for the models in Figure 3 applied to the fiducial simulations, and the corresponding 1σ Fisher constraints. The color scale gives the ratio of each error with respect to its equivalent computed by using only the power spectrum and bispectrum, highlighting the improvements for almost every parameter with the other combinations of summary statistics. Note that the parameter ranges defining the Latin-hypercube is included as prior to the Fisher calculations.

Therefore, the estimates are prior dominated, and always compatible with $f_{\rm NL}^{\rm equil} = 0$, as determined by the network.

In the next three sections, we discuss two different ways of preprocessing the data used as input of the NNs identified by different markers in Figure 3 (i.e., extracting the summary statistics that are used as-is or after the compression step in Equation (2)), we comment on the comparison between the Fisher bounds and the NN predictions for the fiducial cosmology shown in Figure 4, and we benchmark the performance of different combinations of summary statistics.

5.2.1. Data Preprocessing

We compare two different approaches to data preprocessing. Namely, in one case we directly feed the summary statistics to the networks, whereas in the other we adopt a precompression step. In principle, as discussed in J. Alsing et al. (2019), compressing the data decreases the amount of noise while preserving most of the information. Therefore, the NNs trained on the compressed statistic should require a comparatively smaller training set to reach the same accuracy as one naively trained on all the data. However, for the case of $f_{\rm NL}^{\rm equil}$—one of the main parameters in our analysis—this leads to underwhelming results.

This is shown in Figure 3, where we compare the network trained on the uncompressed summary statistics and the network trained on the compressed statistics. To explain why the compression hinders the training, we have to remember that it is meant to be performed with the maximum likelihood parameters to be optimal, or iteratively until the maximum

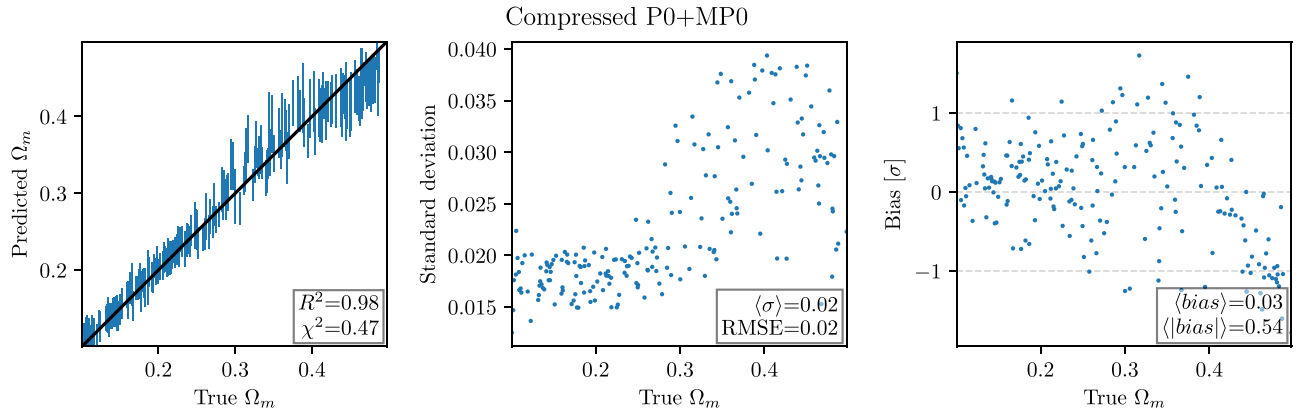

Figure 5. Predictions of $\Omega_{\rm m}$ of the best performing network trained on the compression of the P0 and MP0 statistics, and applied to the test set of the $f_{\rm NL}^{\rm equil}$ LH. Left panel: comparison of the true and predicted values of $\Omega_{\rm m}$. The lines span the predicted 1σ about the predicted value. Center panel: predicted standard deviations as a function of the true parameter value. The mean standard deviation is also shown, compared with the root mean squared errors (RMSE) of the mean predictions. Right panel: bias of the mean prediction in units of the predicted standard deviations.

likelihood is reached. Instead, we just calculate it once at the fiducial (near the central values of the ranges of parameters covered in the different LHs). In the case of the $f_{\rm NL}^{\rm equil}$ LH, where five different parameters are varied, none of the simulation has an input set close to the the fiducial, making the compression substantially suboptimal every time.

We will come back to the subject of data compression in Appendix C, where we will apply the same pipeline to different data sets, to glean some information about how many simulations are needed to properly train the network.

5.2.2. Comparison with Fisher Bounds

Besides looking at R^2 and χ^2, which are quantities calculated from the test set of the Latin-hypercubes (see Section 4.2), another useful test consists in comparing directly the standard deviations predicted by the NNs applied to the 15,000 fiducial simulations with the Fisher bounds. The corresponding results are shown in Figure 4. An apparent feature is that, even for the best performing NN, the standard deviations are larger than the Cramer–Rao bound. This means our estimates are somewhat conservative while at the same time the fact we are consistent with this bound provides a validation test of the NN training methodology. In order to correctly interpret this result, we need to remember that the NN is aimed at building reliable estimates on the whole Latin-hypercube and not at minimizing the errors for a specific set of parameters. Adding the fact that we have a relatively limited set of simulations at our disposal to train the network (see also the discussion in the next section), the observed suboptimality is to be expected. We discuss this aspect in Appendix C, and a detailed analysis of the convergence with more simulations will need to be carried out in the future. For example, in the slightly different setup of B. Tucci & F. Schmidt (2024), numerical convergence is reached with $\sim 10^4$ simulations.

The same also holds when NNs are trained on the compressed statistics. We already discussed how for parameters away from the fiducial, the compression is lossy. As in each simulation multiple parameters, if not all, are displaced from the fiducial, the NN learns to estimate the parameters and their standard deviations from the lossy statistic. This is combined with the fact that, due to the regularization applied in the training, the NN has to produce a smooth function of the target parameters. Thus, even when exposed to the compressed statistic calculated in the fiducial, the estimated error bars do not saturate the Cramer–Rao bound. In Figure 5, we see how the standard deviations depend on the input parameter, but do not drop in size when close to the fiducial. We use $\Omega_{\rm m}$ and the P0 and MP0, as this combination has a high R^2 and as such the errors are not driven by the prior.

5.3. Best Combination of Observables

In our analysis, we take for granted the use of the power spectrum (P0), which is relevant to constrain the standard cosmological parameters and to assess which other single summary statistic can be added to it to improve constraints on $f_{\rm NL}^{\rm equil}$. While we show the combination P0 + B in Figures 3 and 4 as reference, we find that the marked power spectrum outperforms the bispectrum, when both are complemented with P0. In particular, the mark $R = 20$, $p = 1$, $\delta_s = 0.5$ dominates the others (the figures show the analysis performed with that single marker). These findings are supported by both the Fisher forecasts and the NN analysis.

Adding the bispectrum, alone or with the marked bispectrum, to the analysis of the power spectrum and marked power spectrum shrink the Cramer–Rao bounds, but is quite inconsequential on the NN analysis. This is due to the strong correlation between the two sets of statistics.

Until now we refrained from discussing the inclusion of the HMF, as observationally speaking it may pose additional problems compared to the spectra. However, adding the HMF to any other combination of observables significantly improves both the accuracy and precision of our constraints. In fact, out of all possible combinations of observables, P0 + MP0 + HMF is the best performing one; while according to the Fisher forecast an extra $\sim$25% can be gained adding standard and marked bispectra.

In Figure 6, we show the $f_{\rm NL}^{\rm equil}$ prediction drawn by the NN trained on P0, MP0, and HMF applied to the test set of the $f_{\rm NL}^{\rm equil}$ LH. Notice that the results are still bounded by the prior. This explains the trend of the bias as a function of the true value of $f_{\rm NL}^{\rm equil}$. The scatter in the standard deviations is much more sizeable for $f_{\rm NL}^{\rm equil}$ than for $\Omega_{\rm m}$ (see Figure 5). However, this effect is due to the value of the other cosmological

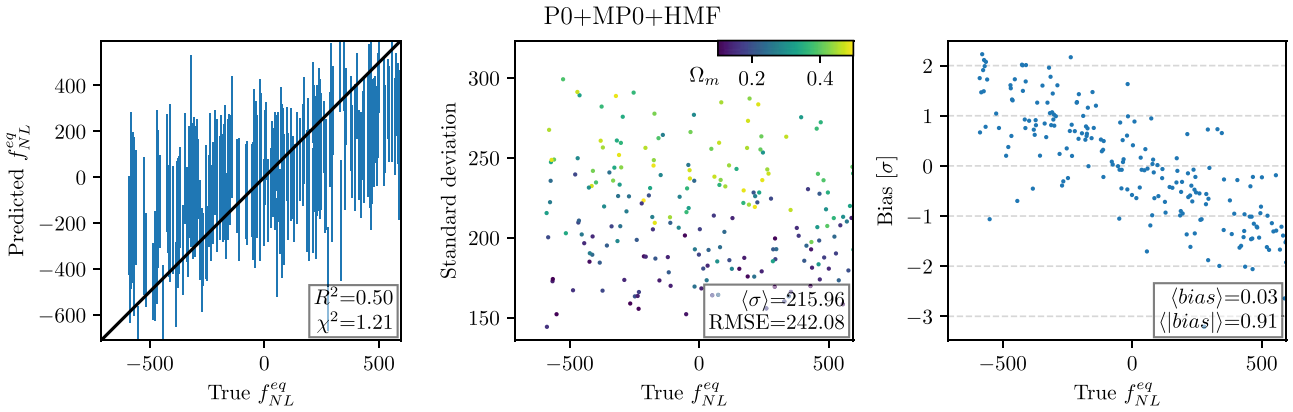

Figure 6. Predictions of $f_{\rm NL}^{\rm equil}$ of the best performing network trained on P0, MP0, and HMF and applied to the test set of the $f_{\rm NL}^{\rm equil}$ LH. The color scale in the middle panel corresponds to the true value of $\Omega_{\rm m}$, which is the cosmological parameter most correlated to the $f_{\rm NL}^{\rm equil}$ standard deviation. See Figure 5 for more details.

parameters in each simulation, the strongest correlation being with $\Omega_{\rm m}$ (correlation coefficient of 0.84).

6. Conclusion

In this work we systematically and quantitatively test and rank a variety of summary statistics, power spectrum, bispectrum, their marked counterparts, and the halo mass function, to plan how to optimally analyze LSS data to measure primordial non-Gaussianity.

We achieve this with two complementary approaches. Our Fisher forecast extends the analysis of G. Jung et al. (2023b), by adding marked observables for the first time in the PNG context. These forecasts are interesting as they show us the optimal errors we could achieve. While extra care has been taken to ensure numerical stability, using an additional data compression step, the results are bound by the choice of fiducial cosmology used to produce the simulations and rely on the assumption of Gaussian likelihood; a limitation that also applies to the quasi-maximum likelihood estimators that can be built with the same pipeline. These shortcomings are covered by using NNs trained on Latin-hypercubes to perform simulation-based inference. They allow us to construct likelihood-free estimators for each summary statistic combination from a relatively low number of simulations (here, an order of magnitude less than the number needed for fully converged Fisher forecast). Moreover, these estimators are reliable on a wide range of parameter values, thus freeing us from the choice of a specific fiducial cosmology.

We choose to use a standard moment network to estimate the marginalized posterior means and standard deviations, and we tested how using the raw summary statistics is a better choice than their score compression, due to computational limitation—the compression would require a new batch of simulations in many points, if not each, of the Latin-hypercube.

From the physical point of view, our main finding is that marked statistics show a great potential for the search of PNG in upcoming LSS data. Both the marked power spectrum and marked bispectrum helps to break degeneracies between PNG amplitudes and cosmological parameters, which are present at the standard power spectrum and bispectrum level, and thus decrease significantly Fisher error bars on all parameters. In addition, the marked power spectrum sets a tight constraint on $f_{\rm NL}^{\rm equil}$, which outperforms the bispectrum when either is used in conjunction with the power spectrum—we confirm this finding with the analysis of the Latin-hypercube. The same conclusions apply when the halo mass function information is added to the different combinations of summary statistics.

Using the moment network method, we predict $\sigma(f_{\rm NL}^{\rm equil})=280$ from the power spectrum and marked power spectrum measured up to $k_{\rm max}=0.5\,h\,{\rm Mpc}^{-1}$ in a volume of $1\,h^{-3}\,{\rm Gpc}^3$, and $\sigma(f_{\rm NL}^{\rm equil})=214$ adding the halo mass function. If we naively scale the moment network errors with the square root of the volume, we obtain $\sigma(f_{\rm NL}^{\rm equil})=36$ and 28 on a volume of $60\,h^{-3}\,{\rm Gpc}^3$.

This work allowed us to set the structure of an analysis pipeline. It will be useful to repeat it on simulations that include visible tracers of the dark matter halos to appropriately train the moment network, which will open up the possibility of analyzing available data.

Acknowledgments

G.J. acknowledges support from the ANR LOCALIZATION project, grant ANR-21-CE31-0019/490702358 of the French Agence Nationale de la Recherche. A.R. acknowledges support from PRIN-MIUR 2020 METE, under contract no. 2020KB33TP. M.L. acknowledges support by the MIUR Progetti di Ricerca di Rilevante Interesse Nazionale (PRIN) Bando 2022—grant 20228RMX4A.

Appendix A
Neural Network Architecture and Training

The architecture of our NNs consists of a set of fully connected layers. The input layer is followed by a normalization layer and a variable number of hidden layers, all with the same number of nodes. The output layer concatenates two sets of variables described in a moment. We use an ELU activation function (D.-A. Clevert et al. 2015) in all layers beside the output and apply a dropout (N. Srivastava et al. 2014) to each hidden layer of the network. To further regularize the network, we also use weight decay (I. Loshchilov & F. Hutter 2017) and stop the training if the validation loss does not decrease for 300 epochs, after which the best weights are restored. As the target parameters are the mean and standard deviation of each parameter posterior, and the latter is a strictly positive quantity, the output layer combines linear activation functions for the means and ELU+1 for the standard deviations. Weights are initialized according to the prescription

in K. He et al. (2015). The training is performed by the Adam optimizer (D. P. Kingma & J. Ba 2014), with a cyclical learning rate (L. N. Smith 2015). The value of various hyperparameters is set through Bayesian optimization (T. O'Malley et al. 2019) within some parameter range, which we verify a posteriori to be wide enough: the number of hidden layers (in [1, 8]), the number of their nodes (in [8, 2048]), the dropout rate (in [0.3, 0.7]), the weight decay rate (in [10^{-5}, 10^{-3}], with logarithmic sampling), and the base learning rate (in [10^{-6}, 10^{-2}], with logarithmic sampling).

Appendix B
Complementary Fisher Analyses

B.1. Fisher Contours

To illustrate further the gain of information obtained using marked statistics, we show in Figure 7 the joint 1σ and 2σ Fisher constraints on cosmological and PNG parameters. This confirms the complementarity of the marked power spectrum and bispectrum, which do not have the same parameter degeneracies.

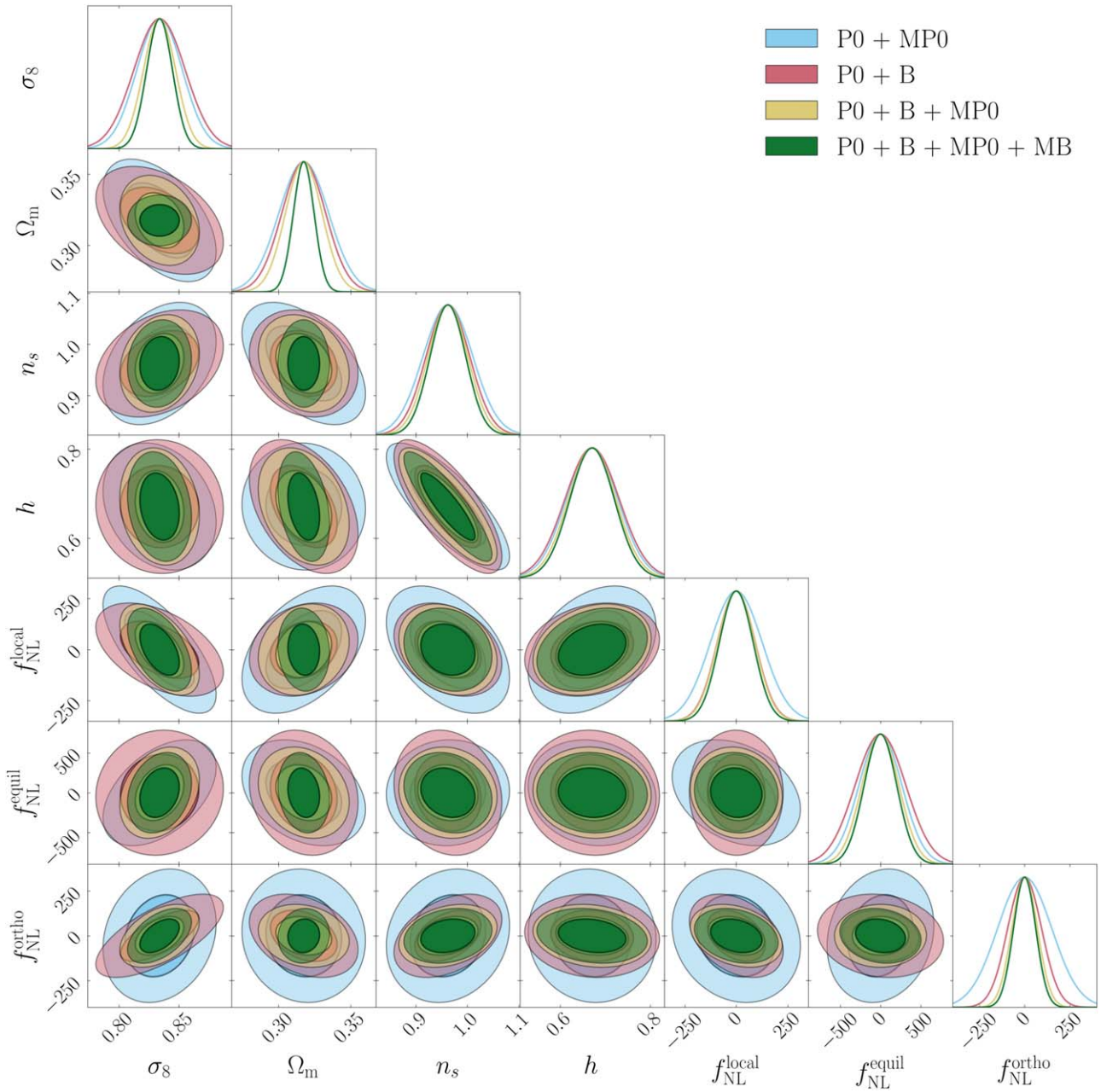

Figure 7. The constraining power of the halo power spectrum (P0), bispectrum (B), marked power spectrum (MP0), and marked bispectrum (MB) at $z = 0$, and determined from the QUIJOTE halo catalogs for $k_{max} = 0.5\,h$ Mpc^{-1}, on cosmological parameters and PNG amplitudes, after marginalizing over M_{min}.

B.2. Convergence with the Number of Simulations

We verify the numerical stability of the Fisher constraints we report in Section 5.1 by using subsets of the available simulations to compute both covariances and derivatives. In Figure 8, we show that our results are indeed stable, as the constraints would be the same (at the percent level) when using less than half of the full set.

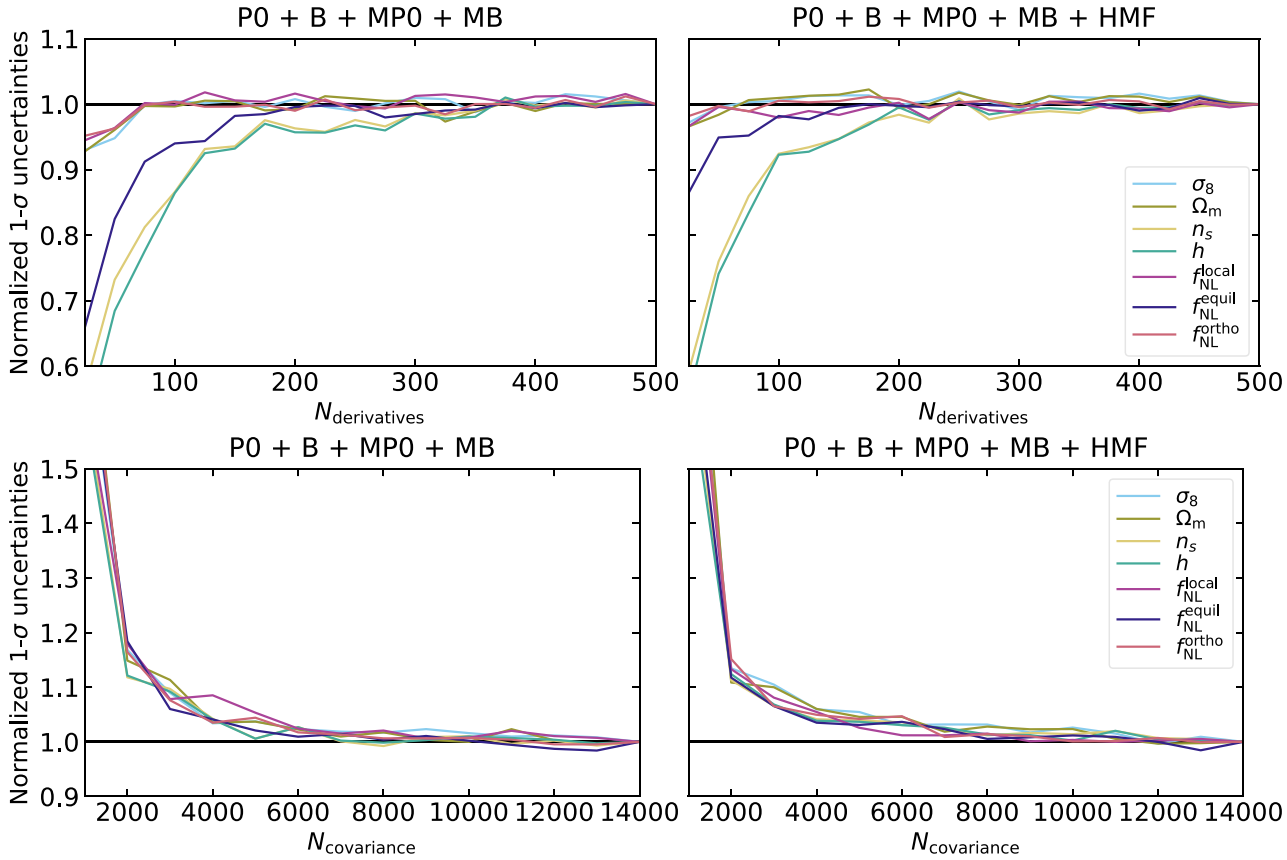

Figure 8. Stability of the Fisher 1σ error bars under variations of the number of simulations used to compute derivatives (top row) or covariances (bottom row). The analysis includes a combination of every summary statistic considered in this work (power spectrum, bispectrum, marked power spectrum, and marked bispectrum) in the left panels, with the addition of the HMF on the right. All error bars are normalized to their values computed with the full set of simulations.

B.3. Fisher Forecast without PNG

As shown in Section 5.1, the marked power spectrum and marked bispectrum are powerful observables to study PNG. Here, we verify that they also help in a standard cosmological parameter analysis, as illustrated in Figure 9. While using the power spectrum jointly with bispectrum or marked bispectrum yields similar results, the combination of these three statistics improves the constraints by around 20% for all parameters. Moreover, adding the marked bispectrum helps to disentangle σ_8 and Ω_m even further, making the Cramer–Rao bounds of these two parameters around 50% smaller than in the standard analysis.

This confirms that marked observables in general, and particularly the marked bispectrum, are interesting observables for the extraction of cosmological information on nonlinear scales.

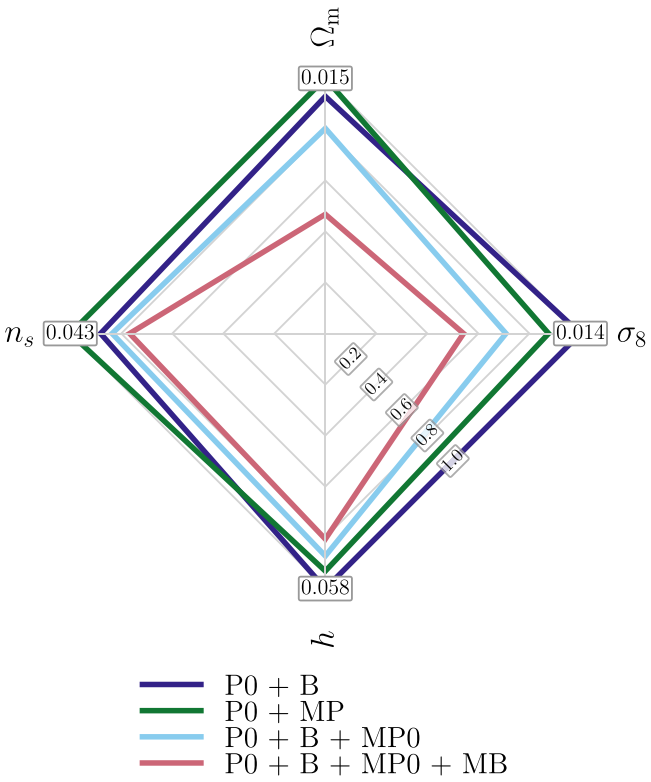

Figure 9. The 1σ Fisher error bars on several cosmological parameters from different combinations of summary statistics measured up to k_{max}= 0.5 h Mpc^{-1} in the QUIJOTE halo catalogs at $z = 0$, after marginalizing over M_{min}. More details can be found in Figure 1.

Appendix C
Small Training Sample Stability

In this appendix, we investigate the convergence of our results with the size of the training sets. We mentioned in Section 5.2 how the compression (if it can be performed correctly) may help to overcome the limits of a small training data set. However, this is not the case for our specific f_{NL}^{equil} analysis, in which we are limited by the currently available number of simulations and by the choice of parameters needed to produce the Latin-hypercube in our multidimensional parameter space; the same practical issues are present in the vanilla ΛCDM LH. Therefore, for the specific purpose of investigating convergence, we choose to study our set of simulations with varying f_{NL}^{local} LH. In this case, only one parameter is varied, while leaving all the others fixed at fiducial values. The size of the model (in terms of number of layers and number of nodes per layer) required to fit the data is thus much smaller than in the case of the f_{NL}^{equil} and of the vanilla ΛCDM LH sets. Thus, the complexity of the output is lower and the number of training simulations required for the compressed statistics to reach convergence is correspondingly smaller. Even the modest number contained in the PNG-LHs is now sufficient. In Figure 10 we show as an example the R^2 and χ^2 values of two sets of NNs trained on the power spectrum and bispectrum, for a variety of number of simulations in the training set N_{tr}. Only in this case, we do not remove the models that have a high χ^2, as it is the matter of the current discussion, but rather take the best model according to the validation loss. We can see how, without compression, R^2 increases throughout the tested range (from $N_{tr} = 100$ to $N_{tr} = 600$), whereas using the compressed statistic, a couple of hundreds of simulations are enough to give the best achievable estimate with these observables. For both the compressed and uncompressed statistic, more simulations are required to accurately estimate the error than to estimate the parameter, as shown in the χ^2 panel, and in Figure 11. However, even in this case, the compressed statistic outperforms the uncompressed statistic when few simulations are at hand. In either case, the training would benefit from a larger training set.

Despite the fact that the compression leads to unsatisfactory results in the analysis of the vanilla LH, since this case contains twice as many simulations as the f_{NL}^{equil} LH, it is still interesting to see how the training progresses and the results progressively improve when using a larger and larger training set, as shown in Figure 12. For completeness, in Figure 13 we also show how this applies to the f_{NL}^{equil} estimate.

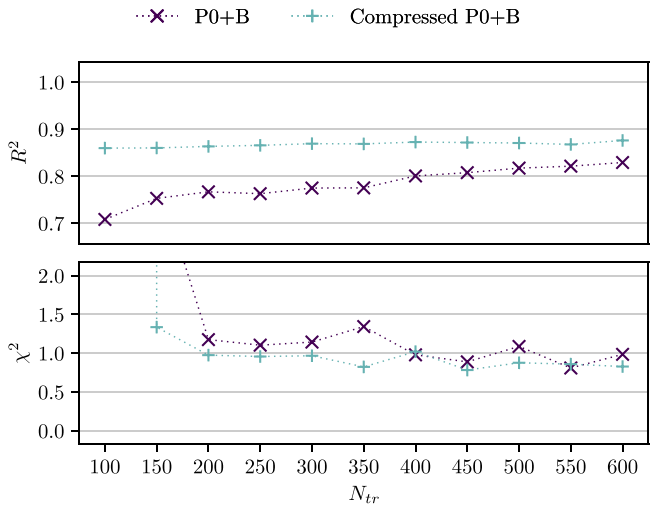

Figure 10. Accuracy of the $f_{\rm NL}^{\rm local}$ prediction as a function of the number of simulations used to train the model. Compressing the data before training the model simplifies the procedure, which needs comparatively less simulations.

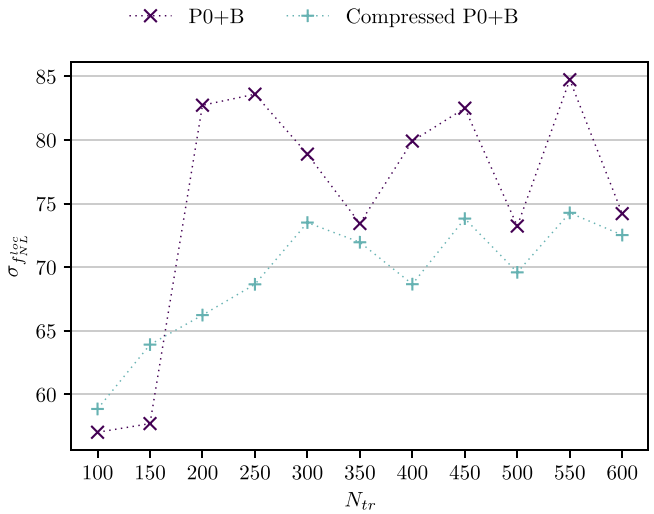

Figure 11. Error associated to the $f_{\rm NL}^{\rm local}$ prediction as a function of the number of simulations used to train the model. See caption of Figure 10.

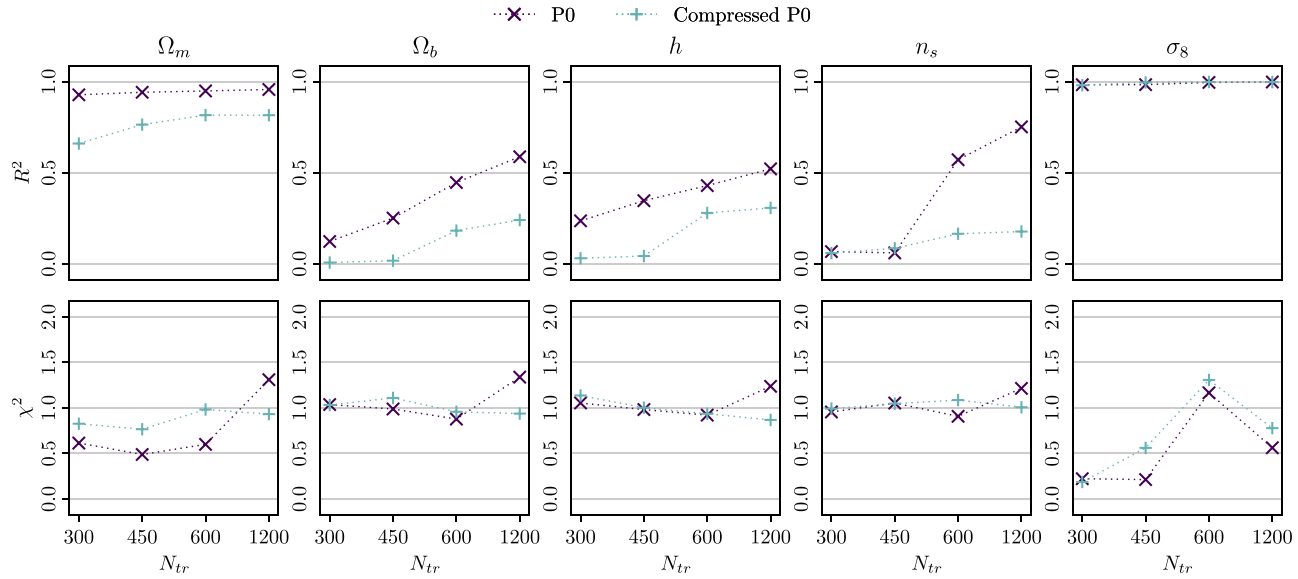

Figure 12. Accuracy of the five vanilla LH parameter predictions analyzing either the power spectrum or the compressed power spectrum, as a function of the training sample size. Notice that the x axis is not uniformly spaced.

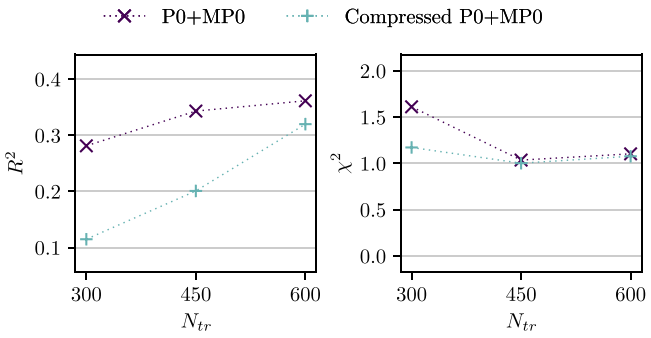

Figure 13. Same as Figure 12 for the $f_{\rm NL}^{\rm equil}$ parameter and combining the power spectrum with the marked power spectrum.

ORCID iDs

Gabriel Jung https://orcid.org/0000-0002-2133-6254
Andrea Ravenni https://orcid.org/0000-0002-1370-0546
Michele Liguori https://orcid.org/0000-0002-0383-2254
Marco Baldi https://orcid.org/0000-0003-4145-1943
William R. Coulton https://orcid.org/0000-0002-1297-3673
Francisco Villaescusa-Navarro https://orcid.org/0000-0002-4816-0455
Benjamin D. Wandelt https://orcid.org/0000-0002-5854-8269

References

Alsing, J., Charnock, T., Feeney, S., & Wandelt, B. 2019, MNRAS, 488, 4440
Alsing, J., & Wandelt, B. 2018, MNRAS, 476, L60
Alsing, J., Wandelt, B., & Feeney, S. 2018, MNRAS, 477, 2874
Andrews, A., Jasche, J., Lavaux, G., & Schmidt, F. 2023, MNRAS, 520, 5746
Baumann, D., & Green, D. 2022, JCAP, 08, 061
Byun, J., & Krause, E. 2023, MNRAS, 525, 4854
Byun, J., Oddo, A., Porciani, C., & Sefusatti, E. 2021, JCAP, 03, 105
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
Cabass, G., Simonović, M., & Zaldarriaga, M. 2024, PhRvD, 109, 043526
Cagliari, M. S., Castorina, E., Bonici, M., & Bianchi, D. 2024, JCAP, 2024, 036
Clevert, D.-A., Unterthiner, T., & Hochreiter, S. 2015, arXiv:1511.07289
Coulton, W. R., Villaescusa-Navarro, F., Jamieson, D., et al. 2023a, ApJ, 943, 178
Coulton, W. R., Villaescusa-Navarro, F., Jamieson, D., et al. 2023b, ApJ, 943, 64
Coulton, W. R., & Wandelt, B. D. 2023, arXiv:2305.08994
Crocce, M., Pueblas, S., & Scoccimarro, R. 2006, MNRAS, 373, 369
D'Amico, G., Lewandowski, M., Senatore, L., & Zhang, P. 2022, arXiv:2201.11518
Davis, M., Efstathiou, G., Frenk, C. S., & White, S. D. M. 1985, ApJ, 292, 371
de Santi, N. S. M., Shao, H., Villaescusa-Navarro, F., et al. 2023, ApJ, 952, 69
Fergusson, J. R., Liguori, M., & Shellard, E. P. S. 2010, PhRvD, 82, 023502
Fergusson, J. R., Liguori, M., & Shellard, E. P. S. 2012a, JCAP, 12, 032
Fergusson, J. R., Regan, D. M., & Shellard, E. P. S. 2012b, PhRvD, 86, 063511
Gatti, M., Jeffrey, N., Whiteway, L., et al. 2024, PhRvD, 109, 063534
Hahn, C., Eickenberg, M., Ho, S., et al. 2023a, JCAP, 04, 010
Hahn, C., Eickenberg, M., Ho, S., et al. 2023b, PNAS, 120, e2218810120
Hahn, C., Eickenberg, M., Ho, S., et al. 2024, PhRvD, 109, 083534
Hahn, C., Lemos, P., Parker, L., et al. 2023c, arXiv, arXiv:2310.15246
Hartlap, J., Simon, P., & Schneider, P. 2007, A&A, 464, 399
He, K., Zhang, X., Ren, S., & Sun, J. 2015, arXiv:1502.01852
Heavens, A. F., Jimenez, R., & Lahav, O. 2000, MNRAS, 317, 965
Hou, J., Moradinezhad Dizgah, A., Hahn, C., et al. 2024, PhRvD, 109, 103528
Ivanov, M. M., Cuesta-Lazaro, C., Mishra-Sharma, S., Obuljen, A., & Toomey, M. W. 2024, PhRvD, 110, 063538
Jasche, J., & Wandelt, B. D. 2013, MNRAS, 432, 894
Jeffrey, N., Alsing, J., & Lanusse, F. 2021, MNRAS, 501, 954
Jeffrey, N., & Wandelt, B. D. 2020, 34th Conf. on Neural Information Processing Systems, arXiv:2011.05991
Jung, G., Karagiannis, D., Liguori, M., et al. 2023a, ApJ, 948, 135
Jung, G., Ravenni, A., Liguori, M., et al. 2023b, ApJ, 957, 1
Kingma, D. P., & Ba, J. 2014, arXiv:1412.6980
Lemos, P., Parker, L., Hahn, C., et al. 2024, PhRvD, 109, 083536
Loshchilov, I., & Hutter, F. 2017, arXiv:1711.05101
Makinen, T. L., Charnock, T., Lemos, P., et al. 2022, OJAp, 5, 18
Massara, E., Villaescusa-Navarro, F., Hahn, C., et al. 2023, ApJ, 951, 70
Massara, E., Villaescusa-Navarro, F., Ho, S., Dalal, N., & Spergel, D. N. 2021, PhRvL, 126, 011301
Ntampaka, M., Eisenstein, D. J., Yuan, S., & Garrison, L. H. 2019, ApJ, 889, 151
O'Malley, T., Bursztein, E., Long, J., et al., 2019 Keras Tuner, https://github.com/keras-team/keras-tuner
Philcox, O. H. E., Massara, E., & Spergel, D. N. 2020, PhRvD, 102, 043516
Planck Collaboration 2020, A&A, 641, A9
Régaldo-Saint Blancard, B., Hahn, C., Ho, S., et al. 2024, PhRvD, 109, 083535
Ribli, D., Pataki, B. A., Zorrilla Matilla, J. M., et al. 2019, MNRAS, 490, 1843
Roncoli, A., Ćiprijanović, A., Voetberg, M., Villaescusa-Navarro, F., & Nord, B. 2023, arXiv:2311.01588
Schmidt, F. 2021, JCAP, 04, 033
Schmittfull, M. M., Regan, D. M., & Shellard, E. P. S. 2013, PhRvD, 88, 063512
Scoccimarro, R., Hui, L., Manera, M., & Chan, K. C. 2012, PhRvD, 85, 083002
Shao, H., Villaescusa-Navarro, F., Villanueva-Domingo, P., et al. 2023, ApJ, 944, 27
Smith, L. N. 2015, arXiv:1506.01186
Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. 2014, JMLR, 15, 1929
Tucci, B., & Schmidt, F. 2024, JCAP, 2024, 063
Villaescusa-Navarro, F., Genel, S., Angles-Alcazar, D., et al. 2022, ApJS, 259, 61
Villaescusa-Navarro, F., Hahn, C., Massara, E., et al. 2020, ApJS, 250, 2
Villanueva-Domingo, P., & Villaescusa-Navarro, F. 2022, ApJ, 937, 115
White, M. 2016, JCAP, 11, 057