

Capturing primordial non-Gaussian signatures in the late Universe by multiscale extrema of the cosmic log-density field

Yun Wang (王云)^{1,*} and Ping He (何平)^{1,2,†}

¹College of Physics, Jilin University, Changchun 130012, China

²Center for High Energy Physics, Peking University, Beijing 100871, China

 (Received 25 August 2024; accepted 21 December 2024; published 5 February 2025)

We construct two new summary statistics, the scale-dependent peak height function (scale-PKHF) and the scale-dependent valley depth function (scale-VLYDF) of matter density, and forecast their constraining power on primordial non-Gaussianity and cosmological parameters based on Quijote and Quijote-PNG simulations at $z = 0$. With the Fisher analysis, we demonstrate that these statistics outperform the power spectrum and bispectrum. Key findings include: (1) the constraint on the scalar spectral index n_s obtained from the scale-VLYDF/scale-PKHF is 1.59/1.10 times tighter than that from the joint analysis of power spectrum and bispectrum; (2) the combination of the two statistics yields a slight improvement in constraining $\{f_{\text{NL}}^{\text{local}}, f_{\text{NL}}^{\text{equil}}\}$ over the power spectrum-bispectrum combination, and provides a 1.39-fold improvement in the constraint on $f_{\text{NL}}^{\text{ortho}}$; and (3) after incorporating the power spectrum with our new statistics, parameter constraints surpass those from power spectrum-bispectrum combination by factors up to 2.93. This work offers an effective scheme for extracting primordial signals from the late Universe, paving the way for further breakthroughs in precision cosmology.

DOI: [10.1103/PhysRevD.111.L041302](https://doi.org/10.1103/PhysRevD.111.L041302)

Introduction. The study of the early Universe is an essential topic in modern cosmology, with profound implications for the origin of the cosmos, the formation of cosmic structures, and fundamental physics. A critical aspect of this study is detecting and constraining the non-Gaussianity of primordial density fluctuations, i.e., *primordial non-Gaussianity* (PNG), which is a powerful probe to discriminate inflationary models, and to investigate the high-energy physics of the early Universe [see, e.g., [1,2], for review].

The PNG is preserved throughout the evolution of cosmic matter distribution, leaving observable signatures in both the cosmic microwave background (CMB) and the large-scale structure (LSS) of the late Universe. To date, the most stringent constraints on the amplitudes of PNG, f_{NL}^{X} , come from measurements of the CMB anisotropies by the Planck satellite, which are $f_{\text{NL}}^{\text{local}} = -0.9 \pm 5.1$, $f_{\text{NL}}^{\text{equil}} = -26 \pm 47$, and $f_{\text{NL}}^{\text{ortho}} = -38 \pm 24$ at 68% C.L. [3], corresponding to local, equilateral, and orthogonal shapes of the primordial potential bispectrum, respectively [4]. However, the two-dimensional (2D) nature and Silk damping hamper the further improvement of CMB's constraining ability [5]. The ongoing and upcoming LSS surveys hold promise for offering enhanced sensitivity to PNG [1,6,7], since they can map a huge three-dimensional volume of our Universe with high-scale resolution. Yet, this approach faces substantial

complications due to the fact that feeble primordial information is obscured by the late-time non-Gaussianity induced by the nonlinear gravity and other astrophysical processes.

Confronted with the challenge, the scientific community has persistently strived to develop sophisticated methodologies that go beyond the vanilla power spectrum and bispectrum of 3D density field, including but not limited to marked power spectrum [8,9], power spectra in cosmic web environments [10,11], one-point probability distribution function [PDF; [12,13]], neural network [14,15], persistent homology [16–18], and field-level inference [19,20]. Which method optimizes the extraction of cosmological information from LSS remains an open question.

Motivated by the above, we explore another potential avenue in this article, focusing on the following crucial features of the late-time matter distribution: First, the density field's PDF is nearly log-normal [21,22]. Hence the logarithmic transform of the density field makes it more Gaussian-like and less nonlinear [23–25]. Second, the density field is manifested in a hierarchical weblike structure [26–28], which is most suitable to be analyzed with multiscale tools, such as continuous wavelet transform [CWT; [29–31]]. Third, the local extrema (halos/peaks and voids/valleys) of the density field have been shown to be particularly sensitive to the PNG [32,33]. Considering them all together, we first perform the wavelet transform of the log-density field, then identify local extrema on multiple scales and count them, thereby defining the *scale-dependent peak height function* (scale-PKHF) and

* Contact author: yunw@jlu.edu.cn

† Contact author: hep@jlu.edu.cn

scale-dependent valley depth function (scale-VLYDF). Here, we will demonstrate the outperformance of this pair of summary statistics in constraining PNG and standard cosmological parameters.

Multiscale extrema of the log density. With the aim of effectively mitigating the effects of late-time non-Gaussianity, we apply the logarithmic transform to the density field, which is given by

$$\rho_{\ln}(\mathbf{x}) = \ln[1 + \delta(\mathbf{x})], \quad (1)$$

where the density field is constructed by assigning the particle positions to a regular grid with $N_g = 512^3$ cells using the piecewise cubic spline window function [34]. Then convolved with a wavelet Ψ , the CWT of the log-density field ρ_{\ln} can be obtained as below:

$$\tilde{\rho}_{\ln}(w, \mathbf{x}) = \int \rho_{\ln}(\mathbf{x}') \Psi(w, \mathbf{x} - \mathbf{x}') d^3 \mathbf{x}', \quad (2)$$

in which $\Psi(w, \mathbf{x}) = w^{3/2} \Psi(w\mathbf{x})$ is the rescaled wavelet of scale w . In the frame of CWT, there are numerous wavelet options [31]. We use the isotropic Gaussian-derived wavelet [GDW; [35,36]], a Mexican-hat-shaped wavelet, as the mother wavelet, given its suitability for detecting peaks and valleys across multiple scales [37]. Its explicit form is presented below:

$$\Psi(\mathbf{x}) = C_N (6 - |\mathbf{x}|^2) e^{-|\mathbf{x}|^2/4}, \quad (3)$$

where $C_N = (15(2\pi)^{3/2})^{-1/2}$ is the normalization constant. For such isotropic wavelets, the log-density field can be reconstructed as follows [38]:

$$\rho_{\ln}(\mathbf{x}) = \langle \rho_{\ln} \rangle_V + \frac{1}{\mathcal{K}_\Psi} \int_0^{+\infty} w^{\frac{1}{2}} \tilde{\rho}_{\ln}(w, \mathbf{x}) dw, \quad (4)$$

where $\langle \rho_{\ln} \rangle_V$ is the mean log density over the whole space, and $\mathcal{K}_\Psi = \int_0^{+\infty} [\hat{\Psi}(k)/k] dk$ with $\hat{\Psi}(k)$ being the Fourier transform of $\Psi(\mathbf{x})$. It can be seen from Eqs. (2) and (4) that the CWT provides a complete multiscale picture of the matter distribution.

Next, we find the peaks/valleys of the CWT field $\tilde{\rho}_{\ln}(w, \mathbf{x})$ at a given scale w by locating cells with values above/below their neighbors. By counting those extrema, we can define the scale-PKHF $n_{\text{pk}}(w, \nu)$ /scale-VLYDF $n_{\text{vly}}(w, \nu)$ as the number density of CWT peaks/valleys with heights/depths falling in the bin $[\nu - d\nu/2, \nu + d\nu/2)$ per unit volume at scale w , which can be mathematically expressed as

$$n_{\text{pk}}(w, \nu) = \frac{d\mathcal{N}_{\text{pk}}(w)}{d\nu} \quad (5)$$

and

$$n_{\text{vly}}(w, \nu) = \frac{d\mathcal{N}_{\text{vly}}(w)}{d\nu}, \quad (6)$$

where $\mathcal{N}_{\text{pk}}(w)$ and $\mathcal{N}_{\text{vly}}(w)$ are the overall number densities of peaks and valleys at scale w , respectively. For comparing the scale-PKHF and scale-VLYDF with the power spectrum, we need to match the wavelet scale w with the wave number k by the correspondence $w = c_w k$ [see Appendix A of [40]], where $c_w = 2/\sqrt{7}$ for the isotropic GDW. Then our measurements will be restricted to (i) eight linearly spaced scales in the nonlinear regime of $0.1 \leq w/c_w \leq 0.5h \text{ Mpc}^{-1}$, (ii) ten linear peak-height bins corresponding to $0 < \nu \leq 4.5 \tilde{\rho}_{\ln, \text{rms}}$, and (iii) 12 linear valley-depth bins corresponding to $-5.4 \tilde{\rho}_{\ln, \text{rms}} \leq \nu < 0$, where $\tilde{\rho}_{\ln, \text{rms}} = \sqrt{\langle |\tilde{\rho}_{\ln}(w, \mathbf{x})|^2 \rangle_V}$ denotes the mean square root at the scale w . With this configuration, the space is ensured to have at least one peak/valley on the largest scale and in the highest peak/deepest valley bin. Henceforth for convenience, we will use k to replace w without any ambiguity. To gain physical intuition, we make a visualization of our basic idea and measurements in Fig. 1.

Simulations. The present work utilizes mock density fields at $z = 0$ derived from the Quijote [41] and its extension Quijote-PNG [42], which are publicly available large suites of N -body simulations with a wide range of cosmological parameter space. Each simulation tracks the gravitational evolution of 512^3 dark matter particles from $z = 127$ to $z = 0$ in a cubic box of side $L_{\text{box}} = 1h^{-1} \text{ Gpc}$ using the TreePM code GADGET-III [43]. The initial conditions of Quijote are Gaussian and generated by the 2LPTIC code [44], while those of Quijote-PNG are non-Gaussian and generated by the 2LPTPNG code [42,45]. The simulations are organized into different sets depending on their cosmological parameters. Among them, the fiducial set contains $N_{\text{fid}} = 15,000$ random realizations with parameters of $\{f_{\text{NL}}^{\text{local}} = 0, f_{\text{NL}}^{\text{equil}} = 0, f_{\text{NL}}^{\text{ortho}} = 0, \Omega_m = 0.3175, \Omega_b = 0.049, \sigma_8 = 0.834, n_s = 0.9624, h = 0.6711\}$, which can be used to compute the covariance matrix of the statistic. Corresponding to each parameter, there is a simulation set containing $N_{\text{deriv}} = 500$ pairs of realizations, in which this parameter is perturbed by a small step around its fiducial value leaving the others unchanged. Then in this way, one can compute the partial derivative of the statistic concerning the parameter. For the parameters we considered here, the step sizes are $\{df_{\text{NL}}^{\text{local}} = \pm 100, df_{\text{NL}}^{\text{equil}} = \pm 100, df_{\text{NL}}^{\text{ortho}} = \pm 100, d\Omega_m = \pm 0.010, d\Omega_b = \pm 0.002, d\sigma_8 = \pm 0.015, dn_s = \pm 0.020, dh = \pm 0.020\}$.

Fisher information analysis. The Fisher information matrix is a commonly used tool to assess the cosmological constraining power of a summary statistic [46–48], which is defined as

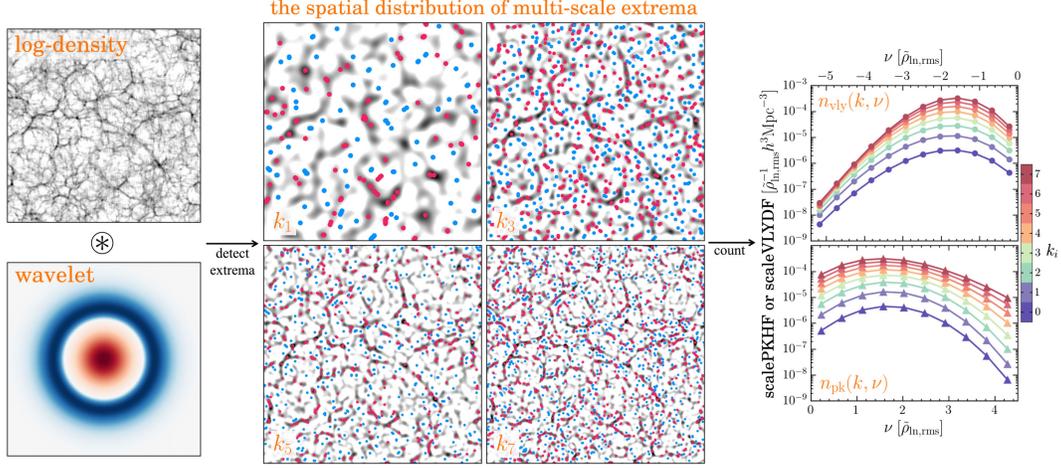


FIG. 1. The illustration of measurements of the scale-PKHF [Eq. (5)] and scale-VLYDF [Eq. (6)]. The left panels represent the CWT operation [Eq. (2)], in which the 2D log-density slice of $500 \times 500 \times 20 (h^{-1} \text{Mpc})^3$ is drawn from a fiducial simulation of Quijote at $z = 0$, and the 2D wavelet plot (with red for positive values and blue for negative values) is the cross section of the isotropic GDW [Eq. (3)] in X-Y plane. Here, the CWT operations are implemented on scales of $\{k_i | k_i \equiv w_i/c_w = (0.1 + i\Delta_k)h \text{Mpc}^{-1} \text{ with } 0 \leq i \leq 7 \text{ and } \Delta_k = 2/35\}$. Then we detect the local extrema of the CWT at each scale. For compactness, the middle panels show only the extrema of CWT fields on four selected scales, where gray regions indicate positive values of the CWT, blank regions indicate negative values, with red dots marking peaks and blue dots valleys. By counting the extrema, the measured scale-VLYDF $n_{\text{vly}}(k, \nu)$ and scale-PKHF $n_{\text{pk}}(k, \nu)$ are shown in the right panels.

$$\mathcal{F}_{ij} = \left(\frac{\partial \langle \mathbf{S} \rangle_{\text{deriv}}}{\partial \theta_i} \right) \mathcal{C}^{-1} \left(\frac{\partial \langle \mathbf{S} \rangle_{\text{deriv}}}{\partial \theta_j} \right)^{\text{T}}, \quad (7)$$

where $\langle \cdot \rangle_{\text{deriv}}$ denotes the ensemble average over N_{deriv} paired simulations for each parameter. The statistic vector \mathbf{S} is composed of the scale-PKHF, scale-VLYDF, and power spectrum, with elements ordered first by the scale k and then by ν at each scale. θ_i is the i th parameter of $\boldsymbol{\theta} = \{f_{\text{NL}}^{\text{local}}, f_{\text{NL}}^{\text{equil}}, f_{\text{NL}}^{\text{ortho}}, h, n_s, \Omega_m, \Omega_b, \sigma_8\}$, and \mathcal{C} is the covariance matrix of statistic defined as

$$\mathcal{C} = \frac{1}{N_{\text{fid}} - 1} \sum_{n=1}^{N_{\text{fid}}} (\mathbf{S}_n - \langle \mathbf{S} \rangle_{\text{fid}})^{\text{T}} (\mathbf{S}_n - \langle \mathbf{S} \rangle_{\text{fid}}), \quad (8)$$

in which the statistic \mathbf{S}_n is measured from the n th simulation, and $\langle \cdot \rangle_{\text{fid}}$ denotes the ensemble average over N_{fid} fiducial simulations. To get an unbiased estimate, we multiply the inverse of the covariance matrix by the Hartlap factor of $(N_{\text{fid}} - N_S - 2)/(N_{\text{fid}} - 1)$ [49], where N_S is the size of \mathbf{S} .

The inverse of the Fisher matrix \mathcal{F}^{-1} provides lower bounds on the parameter error covariance, with its diagonal elements estimating the $1-\sigma$ marginalized error on parameters

$$\sigma^2(\theta_i) \geq (\mathcal{F}^{-1})_{ii}. \quad (9)$$

Results. A theoretical understanding of the covariance matrix of summary statistics is particularly important for parameter forecasting from surveys. For this, we show in Fig. 2 the normalized covariance (i.e., correlation) matrix $r_{ij} \equiv C_{ij}/\sqrt{C_{ii}C_{jj}}$ of the scale-VLYDF, scale-PKHF, and power

spectrum [50], computed numerically using Eq. (8). As expected, the covariances of the scale-PKHF and scale-VLYDF are more diagonalized than those of the power spectrum, enabling more cosmological information to be retrieved. It can also be seen that there is a minimal correlation between those statistics, indicating that the information they provide is complementary.

Given the covariance matrix, we can also determine the cumulative signal-to-noise ratio, which is another useful proxy for the information content of the summary statistic. Its estimations are displayed in Fig. 3 for the studied statistics as a function of the maximum wave number k_{max} .

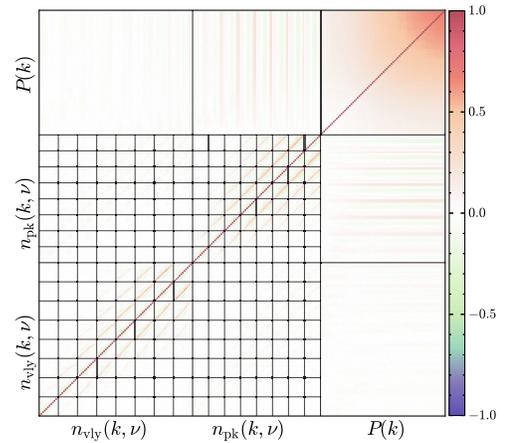


FIG. 2. The correlation matrix of the scale-VLYDF, scale-PKHF, and power spectrum. Note that for the two formers, gray lines partition the matrix into cells organized by scale.

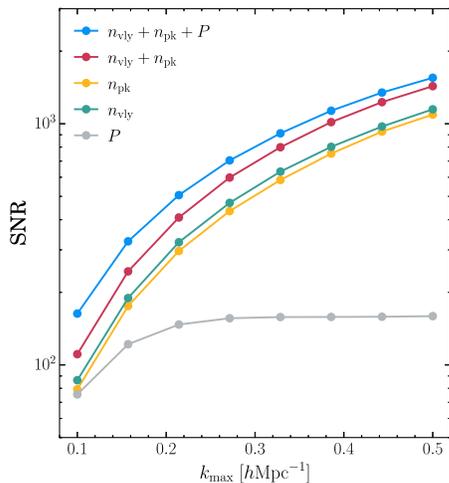


FIG. 3. The cumulative signal-to-noise ratio, $\text{SNR} = \sqrt{\langle \mathbf{S} \rangle_{\text{fid}} \mathbf{C}^{-1} \times \langle \mathbf{S} \rangle_{\text{fid}}^T}$, of the scale-VLYDF, scale-PKHF, power spectrum, and their combinations, as labeled. The eight maximum wave numbers are actually the scales at which we perform the CWT.

We see that the power spectrum SNR flattens out beyond $k \approx 0.3h \text{ Mpc}^{-1}$, which has also been reported in previous studies, and that the bispectrum follows the same feature with a lower SNR [10,42,51,52]. In contrast, both the scale-PKHF and scale-VLYDF do not experience such flattening and achieve a high SNR level. The combination of them both gives a much better SNR, up to 8.98 times higher than the power spectrum at $k_{\text{max}} = 0.5h \text{ Mpc}^{-1}$, and even 9.73 times when the power spectrum is included. We note that the combination of power spectra in cosmic web environments in [10] can achieve an eight times higher SNR than the ordinary power spectrum.

In Fig. 4, we present the Fisher forecast for $1-\sigma$ confidence contours of PNG and cosmological parameters at the maximum wave number $k_{\text{max}} = 0.5h \text{ Mpc}^{-1}$. Considering that the power spectrum carries negligible information on PNG, we also include the constraints from the bispectrum and its combination with the power spectrum for comparison [15,42]. For further clarity, we also list explicitly the improvement factors over the power spectrum in Table I for all statistics. We see that the use of either scale-VLYDF or scale-PKHF alone can boost constraints significantly on the three types of PNG and break key degeneracies between parameters, including those between different PNG types, those between PNG amplitudes and cosmological parameters, and those between different cosmological parameters, e.g., $n_s - \sigma_8$, $h - \sigma_8$, $\Omega_b - \sigma_8$, and $n_s - h$. It is noteworthy that the scalar spectral index n_s is constrained much more tightly by the scale-VLYDF and scale-PKHF, outperforming the power spectrum by factors of 12.37 and 8.61, the bispectrum by factors of 3.90 and 2.72, and the power spectrum-bispectrum combination by

1.59 and 1.10, respectively. However, they provide weaker constraints on Ω_m than that from the power spectrum.

After jointly combining the scale-PKHF and scale-VLYDF, all the parameter constraints become more stringent than those from the power spectrum. For the primordial parameter subset $\{f_{\text{NL}}^{\text{local}}, f_{\text{NL}}^{\text{equil}}, f_{\text{NL}}^{\text{ortho}}, n_s\}$ embedding the primordial information, constraints are improved approximately by factors of $\{2.11, 1.22, 2.40, 4.92\}$ over the bispectrum, and by $\{1.05, 1.03, 1.39, 2.00\}$ relative to the bispectrum and power spectrum combination. When the power spectrum is included in our combination, all constraints are further improved, reaching $\{1.72, 2.18, 1.50, 1.33, 2.93, 1.16, 1.05, 1.62\}$ times the joint constraints from the power spectrum-bispectrum combination for parameters $\{f_{\text{NL}}^{\text{local}}, f_{\text{NL}}^{\text{equil}}, f_{\text{NL}}^{\text{ortho}}, h, n_s, \Omega_m, \Omega_b, \sigma_8\}$. This demonstrates that tighter parameter constraints can be achieved by combining the power spectrum with the scale-VLYDF and scale-PKHF, bypassing the reliance on higher-order spectra (or correlation functions).

Conclusions. In this letter, we propose a pair of new summary statistics, the scale-PKHF and scale-VLYDF of 3D density field, which are well defined, easy to implement, and fully leverage the multiscale nature, log-normal property, and local extrema distribution of the matter distribution at late times. Based on massive datasets of Quijote and Quijote-PNG simulations, we apply the two statistics to forecast the primordial non-Gaussianity and cosmology with the Fisher matrix formalism.

We find that the covariance matrix of the scale-PKHF and scale-VLYDF shows less scale coupling than that of the ordinary power spectrum. Combining the scale-PKHF and scale-VLYDF with the power spectrum can achieve a high SNR of ~ 10 times the power spectrum at the maximum wave number of $0.5h \text{ Mpc}^{-1}$ without showing signs of flattening. These facts suggest that the two statistics can extract huge information content from the LSS. Further, we note that the sole use of scale-VLYDF or scale-PKHF can already put a tighter constraint on the scalar spectral index n_s than the power spectrum-bispectrum combination. Jointly considering the scale-PKHF and scale-VLYDF leads to much stronger constraints on the amplitudes of PNG $\{f_{\text{NL}}^{\text{local}}, f_{\text{NL}}^{\text{equil}}, f_{\text{NL}}^{\text{ortho}}\}$ than the bispectrum, while offering a modest improvement over the power spectrum-bispectrum combination, which highlights that the scale-PKHF and scale-VLYDF are very sensitive to the faint primordial signals in the LSS. By incorporating the power spectrum, scale-PKHF, and scale-VLYDF, all parameter constraints are tightened compared to those from the power spectrum-bispectrum combination. The greatest improvement is observed for n_s , followed by the PNG parameters and σ_8 , with the least improvement for the remaining parameters.

Overall, we conclude that our methodology shows great superiority in constraining PNG and cosmological

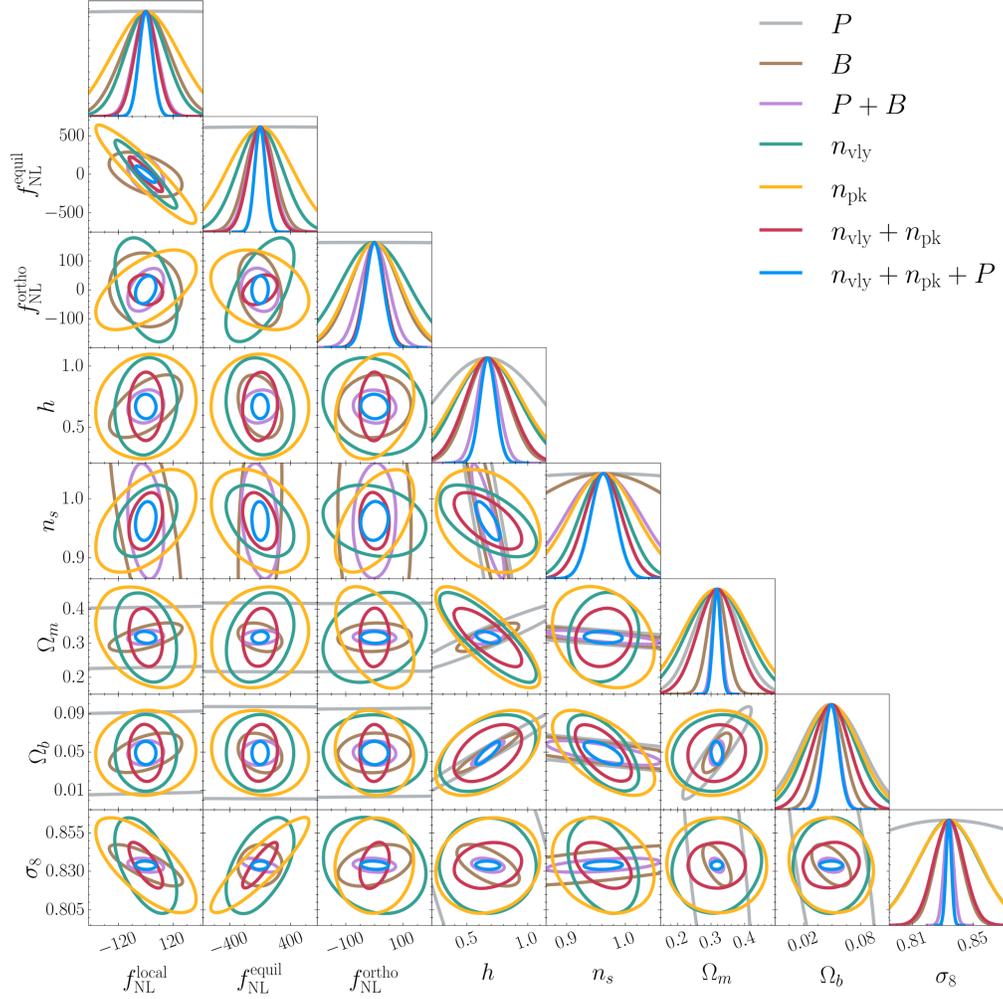


FIG. 4. The marginalized $1\text{-}\sigma$ confidence contours for PNG and standard cosmological parameters obtained from the scale-VLYDF, scale-PKHF, power spectrum, bispectrum, and some of their combinations, as labeled. In some subplots, the confidence ellipses from the power spectrum are too large to fit within the plotted range. The results for the bispectrum and its combination with power spectrum are sourced from [15] and consistent with those presented in [42], which are the measurements with the maximum wave number being cut off at $0.5h \text{ Mpc}^{-1}$ and without applying the reconstruction algorithm.

parameters. Notably, next-generation surveys like DESI [53] and Euclid [54], which aim to map the Universe at unprecedented resolution, can benefit significantly from these new statistics. By leveraging their ability to

extract complementary information from the LSS, these surveys can achieve tighter constraints on PNG and other cosmological parameters. Furthermore, the enhanced sensitivity to the scalar spectral index underscores the potential of our methods to refine our understanding of inflationary physics. Nonetheless, it is important to recognize that our analysis is confined to the dark matter density field, which cannot be observed directly and is instead traced by galaxies. Observational effects, such as galaxy bias, redshift-space distortions, survey geometry, and selection functions, can degrade the parameter constraints. To adapt our statistics for real data from LSS surveys and achieve optimal constraints, we plan to leverage the SimBIG [55,56] framework to produce galaxy mock catalogs, incorporating observational effects to robustly infer the posterior distribution of parameters.

We release our code for reproducing our results at [57].

TABLE I. The improvement factors of various statistics over the ordinary power spectrum for PNG amplitudes and cosmological parameters.

Parameters	σ_P/σ_B	σ_P/σ_{P+B}	$\sigma_P/\sigma_{n_{\text{vly}}}$	$\sigma_P/\sigma_{n_{\text{pk}}}$	$\sigma_P/\sigma_{n_{\text{vly}}+n_{\text{pk}}}$	$\sigma_P/\sigma_{n_{\text{vly}}+n_{\text{pk}}+P}$
$f_{\text{NL}}^{\text{local}}$	28.59	57.61	32.73	20.22	60.24	99.14
$f_{\text{NL}}^{\text{equil}}$	45.10	53.28	28.05	19.50	54.82	115.97
$f_{\text{NL}}^{\text{ortho}}$	43.52	74.96	29.82	39.32	104.36	112.42
h	2.59	4.94	1.66	1.53	2.36	6.59
n_s	3.17	7.80	12.37	8.61	15.61	22.82
Ω_m	2.47	5.11	0.78	0.69	1.20	5.92
Ω_b	2.37	3.83	1.20	1.09	1.62	4.01
σ_8	10.06	29.88	4.24	4.23	8.81	48.46

Acknowledgments. We would like to acknowledge the anonymous referee for valuable comments and suggestions. We are grateful to Thomas Flöss for providing parameter constraints estimated from the power spectrum, bispectrum, and their combination. We also thank Francisco Villaescusa-Navarro, William Coulton, and the whole Quijote team for

making their simulations publicly available. This work was supported by the National Science Foundation of China (No. 12147217 and No. 12347163), the China Postdoctoral Science Foundation (No. 2024M761110), and the Natural Science Foundation of Jilin Province, China (No. 20180101228JC).

-
- [1] P. D. Meerburg, D. Green, R. Flauger, B. Wallisch, M. C. D. Marsh, E. Pajer, G. Goon, C. Dvorkin, A. M. Dizgah, D. Baumann *et al.*, Primordial non-Gaussianity, *Bull. Am. Astron. Soc.* **51**, 107 (2019); [arXiv:1903.04409](#).
- [2] A. Achúcarro, M. Biagetti, M. Braglia, G. Cabass, R. Caldwell, E. Castorina, X. Chen, W. Coulton, R. Flauger, J. Fumagalli *et al.*, Inflation: Theory and observations, [arXiv:2203.08128](#).
- [3] Planck Collaboration, Planck 2018 results. IX. Constraints on primordial non-Gaussianity, *Astron. Astrophys.* **641**, A9 (2020).
- [4] D. Babich, P. Creminelli, and M. Zaldarriaga, The shape of non-Gaussianities, *J. Cosmol. Astropart. Phys.* **08** (2004) 009.
- [5] A. Kalaja, P. D. Meerburg, G. L. Pimentel, and W. R. Coulton, Fundamental limits on constraining primordial non-Gaussianity, *J. Cosmol. Astropart. Phys.* **04** (2021) 050.
- [6] M. Alvarez, T. Baldauf, J. R. Bond, N. Dalal, R. de Putter, O. Doré, D. Green, C. Hirata, Z. Huang, D. Huterer *et al.*, Testing inflation with large scale structure: Connecting hopes with reality, [arXiv:1412.4671](#).
- [7] D. Karagiannis, A. Lazu, M. Liguori, A. Raccanelli, N. Bartolo, and L. Verde, Constraining primordial non-Gaussianity with bispectrum and power spectrum from upcoming optical and radio surveys, *Mon. Not. R. Astron. Soc.* **478**, 1341 (2018).
- [8] E. Massara, F. Villaescusa-Navarro, S. Ho, N. Dalal, and D. N. Spergel, Using the marked power spectrum to detect the signature of neutrinos in large-scale structure, *Phys. Rev. Lett.* **126**, 011301 (2021).
- [9] E. Massara, F. Villaescusa-Navarro, C. Hahn, M. M. Abidi, M. Eickenberg, S. Ho, P. Lemos, A. M. Dizgah, and B. R.-S. Blancard, Cosmological information in the marked power spectrum of the galaxy field, *Astrophys. J.* **951**, 70 (2023).
- [10] T. Bonnaire, N. Aghanim, J. Kuruvilla, and A. Decelle, Cosmology with cosmic web environments. I. Real-space power spectra, *Astron. Astrophys.* **661**, A146 (2022).
- [11] T. Bonnaire, J. Kuruvilla, N. Aghanim, and A. Decelle, Cosmology with cosmic web environments. II. Redshift-space auto and cross-power spectra, *Astron. Astrophys.* **674**, A150 (2023).
- [12] C. Uhlemann, E. Pajer, C. Pichon, T. Nishimichi, S. Codis, and F. Bernardeau, Hunting high and low: Disentangling primordial and late-time non-Gaussianity with cosmic densities in spheres, *Mon. Not. R. Astron. Soc.* **474**, 2853 (2018).
- [13] O. Friedrich, C. Uhlemann, F. Villaescusa-Navarro, T. Baldauf, M. Manera, and T. Nishimichi, Primordial non-Gaussianity without tails—How to measure f_{NL} with the bulk of the density PDF, *Mon. Not. R. Astron. Soc.* **498**, 464 (2020).
- [14] U. Giri, M. Münchmeyer, and K. M. Smith, Robust neural network-enhanced estimation of local primordial non-Gaussianity, *Phys. Rev. D* **107**, L061301 (2023).
- [15] T. Flöss and P. D. Meerburg, Improving constraints on primordial non-Gaussianity using neural network based reconstruction, *J. Cosmol. Astropart. Phys.* **02** (2024) 031.
- [16] M. Biagetti, A. Cole, and G. Shiu, The persistence of large scale structures I: Primordial non-Gaussianity, *J. Cosmol. Astropart. Phys.* **04** (2021) 061.
- [17] M. Biagetti, J. Calles, L. Castiblanco, A. Cole, and J. Noreña, Fisher forecasts for primordial non-Gaussianity from persistent homology, *J. Cosmol. Astropart. Phys.* **10** (2022) 002.
- [18] J. H. T. Yip, M. Biagetti, A. Cole, K. Viswanathan, and G. Shiu, Cosmology with persistent homology: A Fisher forecast, *J. Cosmol. Astropart. Phys.* **09** (2024) 034.
- [19] D. Baumann and D. Green, The power of locality: Primordial non-Gaussianity at the map level, *J. Cosmol. Astropart. Phys.* **08** (2022) 061.
- [20] A. Andrews, J. Jasche, G. Lavaux, and F. Schmidt, Bayesian field-level inference of primordial non-Gaussianity using next-generation galaxy surveys, *Mon. Not. R. Astron. Soc.* **520**, 5746 (2023).
- [21] A. J. S. Hamilton, Galaxy clustering and the method of voids, *Astrophys. J. Lett.* **292**, L35 (1985).
- [22] P. Coles and B. Jones, A lognormal model for the cosmological mass distribution, *Mon. Not. R. Astron. Soc.* **248**, 1 (1991).
- [23] M. C. Neyrinck, I. Szapudi, and A. S. Szalay, Rejuvenating the matter power spectrum: Restoring information with a logarithmic density mapping, *Astrophys. J. Lett.* **698**, L90 (2009).
- [24] X. Wang, M. Neyrinck, I. Szapudi, A. Szalay, X. Chen, J. Lesgourgues, A. Riotto, and M. Sloth, Perturbation theory of the cosmological log-density field, *Astrophys. J.* **735**, 32 (2011).
- [25] H. Rubira and R. Voivodic, The effective field theory and perturbative analysis for log-density fields, *J. Cosmol. Astropart. Phys.* **03** (2021) 070.
- [26] R. K. Sheth and R. van de Weygaert, A hierarchy of voids: Much ado about nothing, *Mon. Not. R. Astron. Soc.* **350**, 517 (2004).
- [27] J. V. Sheth, Morphology of mock SDSS catalogues, *Mon. Not. R. Astron. Soc.* **354**, 332 (2004).

- [28] J. Shen, T. Abel, H. J. Mo, and R. K. Sheth, An excursion set model of the cosmic web: The abundance of sheets, filaments, and halos, *Astrophys. J.* **645**, 783 (2006).
- [29] I. Daubechies, *Ten Lectures on Wavelets* (SIAM, Philadelphia, 1992).
- [30] A. B. Romeo, O. Agertz, B. Moore, and J. Stadel, Discreteness effects in Λ CDM simulations: A wavelet-statistical view, *Astrophys. J.* **686**, 1 (2008).
- [31] P. S. Addison, *The Illustrated Wavelet Transform Handbook: Introductory Theory and Applications in Science, Engineering, Medicine and Finance* (CRC Press, Boca Raton, 2017).
- [32] N. Dalal, O. Doré, D. Huterer, and A. Shirokov, Imprints of primordial non-Gaussianities on large-scale structure: Scale-dependent bias and abundance of virialized objects, *Phys. Rev. D* **77**, 123514 (2008).
- [33] K. C. Chan, N. Hamaus, and M. Biagetti, Constraint of void bias on primordial non-Gaussianity, *Phys. Rev. D* **99**, 121304 (2019).
- [34] E. Sefusatti, M. Crocce, R. Scoccimarro, and H. M. P. Couchman, Accurate estimators of correlation functions in Fourier space, *Mon. Not. R. Astron. Soc.* **460**, 3624 (2016).
- [35] Y. Wang and P. He, The continuous wavelet derived by smoothing function and its application in cosmology, *Commun. Theor. Phys.* **73**, 095402 (2021).
- [36] Y. Wang, H.-Y. Yang, and P. He, Continuous wavelet analysis of matter clustering using the Gaussian-derived wavelet, *Astrophys. J.* **934**, 77 (2022).
- [37] E. Slezak, V. de Lapparent, and A. Bijaoui, Objective detection of voids and high-density structures in the first CfA redshift survey slice, *Astrophys. J.* **409**, 517 (1993).
- [38] The derivation of Eq. (4) can be referred to the one-dimensional case in Appendix A of [39].
- [39] Y. Wang and P. He, Comparisons between fast algorithms for the continuous wavelet transform and applications in cosmology: The 1D case, *RAS Tech. Instrum.* **2**, 307 (2023).
- [40] Y. Wang and P. He, Turbulence revealed by wavelet transform: Power spectrum and intermittency for the velocity field of the cosmic baryonic fluid, *Astrophys. J.* **974**, 107 (2024).
- [41] F. Villaescusa-Navarro, C. Hahn, E. Massara, A. Banerjee, A. M. Delgado, D. K. Ramanah, T. Charnock, E. Giusarma, Y. Li, E. Allys *et al.*, The Quijote simulations, *Astrophys. J. Suppl. Ser.* **250**, 2 (2020).
- [42] W. R. Coulton, F. Villaescusa-Navarro, D. Jamieson, M. Baldi, G. Jung, D. Karagiannis, M. Liguori, L. Verde, and B. D. Wandelt, Quijote-PNG: Simulations of primordial non-Gaussianity and the information content of the matter field power spectrum and bispectrum, *Astrophys. J.* **943**, 64 (2023).
- [43] V. Springel, The cosmological simulation code GADGET-2, *Mon. Not. R. Astron. Soc.* **364**, 1105 (2005).
- [44] M. Crocce, S. Pueblas, and R. Scoccimarro, Transients from initial conditions in cosmological simulations, *Mon. Not. R. Astron. Soc.* **373**, 369 (2006).
- [45] R. Scoccimarro, L. Hui, M. Manera, and K. C. Chan, Large-scale bias and efficient generation of initial conditions for nonlocal primordial non-gaussianity, *Phys. Rev. D* **85**, 083002 (2012).
- [46] R. A. Fisher, On the mathematical foundations of theoretical statistics, *Phil. Trans. R. Soc. A* **222**, 309 (1922).
- [47] M. Tegmark, A. N. Taylor, and A. F. Heavens, Karhunen-Loève eigenvalue problems in cosmology: How should we tackle large data sets?, *Astrophys. J.* **480**, 22 (1997).
- [48] J. Carron, On the assumption of Gaussianity for cosmological two-point statistics and parameter dependent covariance matrices, *Astron. Astrophys.* **551**, A88 (2013).
- [49] J. Hartlap, P. Simon, and P. Schneider, Why your model parameter confidences might be too optimistic. Unbiased estimation of the inverse covariance matrix, *Astron. Astrophys.* **464**, 399 (2007).
- [50] We compute the power spectrum by using the public Pylians 3rd library (see <https://pylians3.readthedocs.io/en/master/>) with wave number bins of width $k_F = 2\pi/L_{\text{box}}$, from k_F to $0.5h \text{ Mpc}^{-1}$.
- [51] K. C. Chan and L. Blot, Assessment of the information content of the power spectrum and bispectrum, *Phys. Rev. D* **96**, 023528 (2017).
- [52] W. R. Coulton, F. Villaescusa-Navarro, D. Jamieson, M. Baldi, G. Jung, D. Karagiannis, M. Liguori, L. Verde, and B. D. Wandelt, Quijote-PNG: The information content of the halo power spectrum and bispectrum, *Astrophys. J.* **943**, 178 (2023).
- [53] DESI Collaboration, The DESI experiment part I: Science, targeting, and survey design, [arXiv:1611.00036](https://arxiv.org/abs/1611.00036).
- [54] L. Amendola, S. Appleby, A. Avgoustidis *et al.*, Cosmology and fundamental physics with the Euclid satellite, *Living Rev. Relativity* **21**, 2 (2018).
- [55] C. Hahn, M. Eickenberg, S. Ho, J. Hou, P. Lemos, E. Massara, C. Modi, A. Moradinezhad Dizgah, B. Régaldo-Saint Blancard, and M. M. Abidi, simBIG: A forward modeling approach to analyzing galaxy clustering, *Proc. Natl. Acad. Sci. U.S.A.* **120**, e2218810120 (2023).
- [56] C. Hahn, P. Lemos, L. Parker, B. Régaldo-Saint Blancard, M. Eickenberg, S. Ho, J. Hou, E. Massara, C. Modi, A. Moradinezhad Dizgah, and D. Spergel, Cosmological constraints from non-Gaussian and nonlinear galaxy clustering using the simBIG inference framework, *Nat. Astron.* **8**, 1457 (2024).
- [57] <https://github.com/WangYun1995/CWTextrema>