

Equivalence of the field-level inference and conventional analyses on large scales

Francesco Spezzati^{1,2}, Marco Marinucci^{1,2} and Marko Simonović^{3,4}

¹*Dipartimento di Fisica Galileo Galilei, Università di Padova, I-35131 Padova, Italy*

²*INFN Sezione di Padova, I-35131 Padova, Italy*

³*Dipartimento di Fisica e Astronomia, Università di Firenze, I-50019 Sesto Fiorentino, Italy*

⁴*INFN, Sezione di Firenze, I-50019 Sesto Fiorentino, Italy*



(Received 1 August 2025; accepted 25 February 2026; published 18 March 2026)

We study a simple setup with dark matter halos in real space, with the amplitude of the linear density field A as the only free cosmological parameter. We show that Eulerian perturbation theory is adequate for describing this system on large scales, compute the leading n -point functions and perform a joint power spectrum, bispectrum, and trispectrum analysis. Beyond the bispectrum, which is crucial for breaking the degeneracy between A and the linear bias, we find that addition of the trispectrum reduces the error on A by only 20%–30%. Our results for the joint analysis are in good agreement with recent field-level analyses in the same setup. This implies that the field-level inference on large scales does not get significant information from large displacements beyond those in Eulerian kernels or higher-order n -point functions beyond the trispectrum. We provide further evidence for this showing that the dependence of the error bars on the maximum wave numbers used in the analysis is the same in the two approaches. Our results are in disagreement with some of the recent joint power spectrum and bispectrum analyses using likelihood-free inference based on perturbative forward modeling. We discuss a possible origin of this discrepancy and highlight the importance of resolving it in order to have the optimal results in cosmological analyses based on perturbation theory.

DOI: [10.1103/yhg1-c63b](https://doi.org/10.1103/yhg1-c63b)

I. INTRODUCTION AND MAIN RESULTS

Field-level Bayesian inference (FBI) of cosmological parameters has recently emerged as an alternative to conventional n -point function analyses. This approach offers several advantages, such as being optimal and allowing for a simpler inclusion of data systematics. However, in the context of galaxy clustering, field-level inference which relies on detailed cosmological simulations of the observed galaxy density field remains infeasible, due to very high computational cost. A practical alternative is FBI based on perturbative forward modeling. Within the framework of the effective field theory (EFT) of large-scale structure (LSS) [1–9], perturbative forward modeling [10–19] provides a universal and reliable description of the galaxy density field on large scales with a handful of free parameters to be fitted from the data. Considerable effort has been made to test FBI with perturbative forward

modeling and apply it in practice [20–29] (see also [30] for a recent application of the LSS bootstrap [31–34] to FBI).

A notable recent cosmological analysis using FBI was presented in [35], focusing on comparison of the field-level and conventional n -point function-based approaches. By using the exact same perturbative forward model in both FBI and the simulation-based¹ inference (SBI) for the power spectrum (P) and bispectrum (B), the authors found that FBI achieves a 3–5 times more precise measurement of the amplitude of the initial density field A than the standard $P + B$ combination. The leading hypotheses for explaining this dramatic difference is that the FBI uses more information from higher-order n -point functions and exploits at the field level some nonperturbative features of the nonlinear evolution that come with no new free parameters (e.g., large displacements). However, due to complexity of both FBI and SBI even in the perturbative framework, this has not been rigorously proven and a clear understanding of where the additional information in FBI would exactly come from is still lacking.

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

¹Here “simulation” refers not to the real cosmological simulations, but generation of nonlinear realizations based on the same perturbative forward model as in FBI. For more details, see Ref. [35].

Similar analysis was presented in the data challenge for beyond the two-point galaxy clustering statistics [36]. For example, Fig. 22 in [36] shows the constraints on the amplitude of the linear field for the FBI and the standard EFT-based joint power spectrum and bispectrum analysis for the same volume as in [35], although with less dense sample. In this case, the difference between the two methods is less dramatic (see Sec. 6.3 of [36] for more details about the comparison), but it is hard to straightforwardly compare them. This is mainly due to the fact that the k_{max} used in the P + B analysis is larger from the scale cuts in the FBI. In addition, for this larger k_{max} in the “restricted” EFT P + B analysis only the errors are presented (the posterior is centered at the best-fit values for the analysis with lower k_{max}). This leaves several open questions. Is the P + B analysis biased for smaller k_{max} ? Are the error bars different for smaller k_{max} ? And how do they compare to the FBI results and the SBI errors from [35]? Providing clear answers is very important, both to get a better understanding of theoretical aspects of FBI and to get firmer basis for its applications in practice and comparison to the conventional analyses.

The goal of this paper is to shed new light on these questions, focusing on the usual perturbation theory P + B analysis in the extremely simplified but realistic settings. We use a simple, tractable and intuitive Eulerian perturbative model (without infrared resummation), with the simple Gaussian likelihood and covariance. Validity of such model and likelihood on large scales can be easily verified in practice and we will provide several arguments for it. In a nutshell, such simplified approach is possible given that the analysis presented in [35] follows a particularly simple setup. It uses a simulation box with a relatively small volume of $V = (2 \text{ Gpc}/h)^3$, which implies relatively large statistical error bars. The studied sample consists of dark matter halos in real space with a very high number density. Additionally, it focuses only on relatively high redshifts $z \geq 0.5$ and large scales, with the maximum wave number used in the analysis being $k_{\text{max}} \leq 0.12 \text{ h/Mpc}$. Such setup with regular periodic box, low shot noise, simple tracers without redshift-space distortions or other observational effects, and small fluctuations on very large scales is ideal for understanding FBI using simple analytical methods. Importantly, some of our results then can be easily derived in a very clear and transparent way, without the need to rely on Markov chain Monte Carlo (MCMC) analyses only. Our approach and results crucially rely on several key assumptions that simplify the description of the system significantly. We would like to clearly state them at the very beginning, given that we will use them throughout the paper.

First, we assume that the field of dark matter halos in real space for given data cuts and redshifts is well described by Eulerian perturbation theory. Although the forward model used in [35] is based on Lagrangian perturbation theory,

one can show that the two schemes are practically indistinguishable on large scales, given the statistical error bars of the specified simulation box with $V = (2 \text{ Gpc}/h)^3$. We provide justification and more details in Sec. II A. An important consequence of this fact is that the large displacements beyond those contained in Eulerian kernels cannot play a major role in explaining the results of [35].

Assuming that the Eulerian description is sufficient, we further assume that all nonlinearities are controlled by only one parameter²—the variance of the density field Δ^2 at the scale k —and that this parameter is small. It is given by

$$\Delta^2(k) = \frac{1}{2\pi^2} \int_0^k q^2 dq P_{\text{lin}}(q), \quad (1)$$

where $P_{\text{lin}}(k)$ is the linear matter power spectrum. In a Λ cold-dark-matter- (Λ CDM) like cosmology, for $k_{\text{max}} = 0.12 \text{ h/Mpc}$ and $z = 0.5$, the variance of the density field is approximately $\Delta^2 \approx 0.13$, and it gets smaller at higher redshifts and for smaller values of k_{max} . Importantly, we do not only assume that the relative size of each individual term in perturbative calculation is given by $\Delta^2(k_{\text{max}})$, but that the *total* loop contribution to the correlation functions has the same size. This is a nontrivial assumption, since the total one-loop contributions to the higher-order n -point function can have a large number of terms multiplied by a large number of different nuisance parameters. In other words, we assume that, while all bias parameters and counterterms can be numbers of order $\mathcal{O}(1)$, for a realistic tracer they cannot conspire to give large loop corrections. We will explicitly check that this is the case for a typical sample using large-volume numerical simulations and show that deviations of measured correlation functions from the leading-order perturbative predictions are roughly given by $\Delta^2(k)$ on large scales. Even though our assumption puts some constraints on the allowed values of nuisance parameters, it is essential for the perturbation theory to make sense and it is sometimes referred to as the perturbativity prior [37].

With the Eulerian forward model where all nonlinearities are controlled by a single parameter Δ^2 it is easy to compute a few leading-order correlation functions. Since we are interested in comparison to the FBI analysis of [35], we use the cubic model. Within this model, one can consistently compute the one-loop power spectrum, the tree-level bispectrum, and the tree-level trispectrum (T). Higher-order n -point functions and additional loop corrections require quartic and higher-order nonlinearities. Furthermore, since we are focused on large scales only, their impact on cosmological constraints should be much less significant. To estimate this, we can approximate the

²Other parameters, such as the Lagrangian radius of dark matter halos, are irrelevant, given the large number density of the sample, which implies low halo mass.

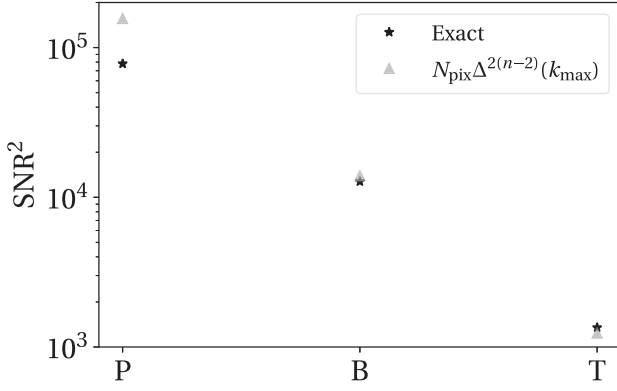


FIG. 1. SNR for the power spectrum, the bispectrum, and the trispectrum. The stars correspond to the exact evaluation using the standard Gaussian likelihood described in Sec. IV (using fiducial bias parameters from Sec. VI), while the triangles correspond to the estimate in Eq. (2). The agreement between the two for the higher-order n -point functions is quite good. A factor of 2 difference for the power spectrum is expected, since the exact result for the SNR is given by $(\text{SNR})_2^2 = N_{\text{pix}}/2$.

signal-to-noise ratio (SNR) in a connected n -point function of the galaxy density field δ_g as follows:

$$(\text{SNR})_n^2 \approx \int^{k_{\text{max}}} \frac{d^3 \mathbf{k}_1}{(2\pi)^3} \cdots \int^{k_{\text{max}}} \frac{d^3 \mathbf{k}_n}{(2\pi)^3} \frac{\langle \delta_g(\mathbf{k}_1) \cdots \delta_g(\mathbf{k}_n) \rangle_c^2}{P_g(k_1) \cdots P_g(k_n)} \sim N_{\text{pix}} \Delta^{2(n-2)}(k_{\text{max}}) [1 + \mathcal{O}(\Delta^2(k_{\text{max}}))], \quad (2)$$

where $P_g(k)$ is the galaxy power spectrum, $N_{\text{pix}} \equiv V \int^{k_{\text{max}}} d^3 \mathbf{k} / (2\pi)^3$ is the number of pixels or Fourier modes in a survey, and $\Delta^2(k_{\text{max}}) \lesssim 0.1$ for the range of redshifts and scales of interest. Note that this estimate is correct up to $\mathcal{O}(1)$ overall multiplicative factor which depends on fiducial values of bias parameters, which we neglect for simplicity. For example, for the power spectrum ($n = 2$), the correct estimate is $(\text{SNR})_2^2 = N_{\text{pix}}/2$, up to small corrections³ of order $\Delta^2(k_{\text{max}})$. One can compute the exact SNR using the standard Gaussian likelihood that we will describe below and check that for higher-order n -point functions the agreement with the estimate above is rather good for small values of nonlinear bias parameters. This is shown in Fig. 1 where the exact calculations for the SNR of P, B, and T are compared to Eq. (2). Note that the signal-to-noise ratio in each higher-order n -point function is roughly 1 order of magnitude smaller than in the previous one, given the value of $\Delta^2(k_{\text{max}})$. This implies that, in the

³Note that the same parameter Δ^2 controls both the SNR for the n -point functions and the corrections to the SNR. This is expected in perturbation theory where all nonlinearities are controlled by Δ^2 . Small corrections to Eq. (2) are captured by the non-Gaussian covariance matrix. The smallness of Δ^2 justifies why using the simple Gaussian covariance matrix is sufficient on large scales.

absence of almost perfect degeneracies, the information content of the nonlinear field on large scales is almost saturated by the leading summary statistics. One well-known strong degeneracy for biased tracers in real space is between the amplitude of the linear fluctuations A and the linear bias b_1 . This degeneracy is broken by combining the power spectrum and bispectrum, and hence the total constraining power of real space data is suppressed by Δ^2 compared to the total signal in the power spectrum.

What do our assumptions imply for the FBI analysis? Given the Eulerian description, perturbativity prior, and the negligible shot noise, one can compute the posterior for the FBI perturbatively. It is then possible to argue that FBI is equivalent, order by order in perturbation theory, to the analysis based on all n -point correlation functions that can be reliably computed up to the given order in the perturbative expansion used for the nonlinear field in the FBI. On the level of equations this equivalence has been explicitly demonstrated in [38] at leading order in Δ^2 and more recently in [39] at subleading order in Δ^2 and for the full cubic bias model. We refer the reader to these references for further details. Showing this connection using simulation data is more challenging and beyond the scope of this paper. Here we will ask two concrete simple questions that can still shed some light on the connection between n -point functions and FBI. First, we can ask how well the amplitude A can be measured with P + B combination using a simple model and the usual Gaussian likelihood. Second, using the Fisher matrix, we can ask if addition of the trispectrum to the P + B analysis significantly improves the errors.

In order to answer these questions, in Sec. III we consistently compute the one-loop power spectrum, the tree-level bispectrum, and the tree-level trispectrum for dark matter halos in real space. We explicitly check that our one-loop power spectrum and tree-level bispectrum model is unbiased, performing the real analysis on the perturbation theory (PT) challenge simulation suite [40],⁴ as shown in Sec. V. We then perform analyses and forecasts using mock data and measure the amplitude of the linear density field for a simulation box with volume $V = (2 \text{ Gpc}/h)^3$ and number density of tracers as in [35,36]. Our two main findings are as follows.

- (i) First, we find that the errors for the amplitude A in the P + B analysis are (a) in agreement with the simple mode counting estimate of Eq. (2) and (b) similar to the errors obtained in [35] using FBI (and therefore significantly smaller than in the SBI P + B analysis of [35]). Our findings are in line with the previous literature on the joint P + B analysis and we demonstrate that our constraints are

⁴The power spectrum data for the PT challenge simulations are publicly available at <https://www2.yukawa.kyoto-u.ac.jp/takahiro.nishimichi/data/PTchallenge/>, see Ref. [40] for more details.

robust against inclusion of the theoretical error that mimic addition of higher-order nonlinearities in the analysis.

- (ii) Second, we find that, in a generic setup, adding the tree-level trispectrum to the P + B analysis only slightly improves the error on A . The improvement we find is of the order $\mathcal{O}(20\%)$ and in agreement with the estimate from Eq. (2). We use a number of simplifying assumptions about the theoretical model for T and its cross-covariance with P and B , which all maximize the impact of T in the joint analysis.

Our results, presented in Sec. VI, indicate that in the simple setup, where the only cosmological parameter is the amplitude of the linear field A , the information content in the nonlinear field on large scales is almost saturated with the power spectrum and bispectrum and that P + B + T analysis leads to very similar errors as the FBI of [35,36]. These results are in line with general expectations in perturbation theory and explicit derivations presented in [38,39]. One remaining puzzle is the origin of the difference between the conventional P + B analysis presented in this paper and the SBI method for P + B used in [35] (see also [28]). We will comment on differences in these two approaches and possible explanations for different results in Sec. VII.

In order to perform our analyses, we relied mostly on the well-known results for n -point functions of biased tracers. However, we also derive several new results such as the covariance matrices for trispectrum analysis and the detailed model for the noise contributions to the trispectrum. Somewhat orthogonal to the main goals of this paper, we also estimate the information content in the baryon acoustic oscillations (BAO) wiggles beyond $k_{\max} = 0.12 \, h/\text{Mpc}$ and discuss how much information on A can come from large displacements at nonlinear scales.

II. PERTURBATIVE DESCRIPTION OF THE HALO DENSITY FIELD

Galaxy density field δ_g in perturbation theory can be written as a sum of two contributions

$$\delta_g = \delta_g^{\text{model}} + \epsilon_g. \quad (3)$$

The first term is the model δ_g^{model} which depends on the initial conditions and cosmological parameters. The second term is the “noise” ϵ_g , which captures the stochastic nature of halo and galaxy formation. This contribution can be understood to arise from integrating over small scales, to which we do not have access in observations (scales smaller than the typical intergalactic separation). In the perturbative expansion valid on large scales the noise is uncorrelated with the matter density field (see Sec. II C for details), but it can be modulated by the long-wavelength modes which leads to several nontrivial contributions to the galaxy

bispectrum and trispectrum. In this section, we will remind the reader how perturbative descriptions for both δ_g^{model} and ϵ_g look and why the simple Eulerian theory for δ_g^{model} is sufficient to capture all relevant nonlinearities in real space on large enough scales.

A. Why is Eulerian description sufficient on large scales?

It is well known that Lagrangian perturbation theory (LPT) reproduces very well the nonlinear fields of dark matter or biased tracers, in real and redshift space. Even when using the Zel’dovich approximation only, one can generate realizations with familiar filamentary structures, which correlate well with the simulations. For this reason, LPT is often employed when making perturbative maps or doing detailed comparisons of simulations and perturbation theory. Furthermore, LPT automatically accounts for effects such as broadening of the BAO peak (or other features) and can be easily implemented in practice on a grid. Given all this, it is natural that the field-level analyses use forward modeling based on LPT to evolve a given initial linear density field (provided cosmological, noise, and bias parameters). LPT up to cubic order with all relevant bias operators in the initial field was used in [35], as well as in the previous works [18,19,26–29].

However, most of the field-level inference has been done so far on very large scales. For example, the largest $k_{\max}$ cut used in [35] is $k_{\max} = 0.12 \, h/\text{Mpc}$. In this regime, there is a significant simplification in the nonlinear dynamics and Eulerian perturbation theory (EPT) is sufficient to take into account all relevant nonlinearities in the field. While all theoretical calculations can be done in LPT as well, the advantage of EPT is that it is easier to work with, it expresses the nonlinear field in terms of the initial conditions in a simple way, and allows one to more easily disentangle different physical effects that can affect the analysis (such as displacements and broadening of the BAO peak versus the growth of structure). Since one of the goals of this paper is to better analytically understand the error bars for the field-level and conventional analyses, we will use EPT.

Why is EPT sufficiently good on large enough scales? There are two main reasons:

- (i) First, the nonlinear fields generated using cubic EPT and LPT up to $k_{\max} = 0.12 \, h/\text{Mpc}$ and at redshifts $z = 0.5$ and $z = 1$ are almost indistinguishable, i.e., their cross-correlation coefficient is close to 1. This has been verified explicitly at the field level [41] and can be easily understood analytically. We will review the main arguments below.
- (ii) Second, even though the FBI is based on comparing the fields predicted in theory and data, it eventually boils down to measuring some summary statistics due to the averaging over all possible initial conditions, which is equivalent to computing the

expectation values [38,39]. It is well known that in EPT and LPT the expectation values are in much closer agreement than what the cross-correlation coefficient indicates, making EPT even more adequate in approximating LPT.

Let us review these two arguments in detail. We begin by estimating the cross-correlation coefficient between EPT and LPT. Let us imagine that in the initial (Lagrangian) coordinates $\mathbf{q}$ we have a linear field $\delta_1(\mathbf{q})$. We can imagine that we evolve this field in two different ways. In EPT, if we use just the linear evolution, the final field is just proportional to $\delta_1(\mathbf{q})$. On the other hand, in LPT we have to displace particles from their initial coordinates $\mathbf{q}$ to their final Eulerian positions $\mathbf{x}$, such that $\mathbf{x} = \mathbf{q} + \boldsymbol{\psi}(\mathbf{q})$. The leading contribution to the full displacement field $\boldsymbol{\psi}(\mathbf{q})$ can be easily computed in the linear theory. Integrating in time the continuity equation with the linear velocity $\mathbf{v}_1$ we can write

$$\delta_1(\tau, \mathbf{q}) = -\nabla \int^\tau d\tau' \mathbf{v}_1(\tau', \mathbf{q}) \equiv -\nabla \boldsymbol{\psi}_1(\tau, \mathbf{q}), \quad (4)$$

which defines the displacement field in the Zel'dovich approximation. Taking the inverse Laplacian (and suppressing the time dependence for simplicity) we can evaluate the linear displacement field starting from the linear density fluctuations as follows:

$$\boldsymbol{\psi}_1(\mathbf{q}) = -\nabla / \nabla^2 \delta_1(\mathbf{q}), \quad (5)$$

where this operation is usually done in Fourier space, such that for each Fourier mode one writes $\boldsymbol{\psi}_1(\mathbf{k}) = i\mathbf{k} / k^2 \delta_1(\mathbf{k})$. Using the linear displacement, the leading-order (LO) LPT field is given by

$$\delta_{\text{LPT}}^{\text{LO}}(\mathbf{k}) = \int d^3\mathbf{q} e^{-i\mathbf{k} \cdot \boldsymbol{\psi}_1(\mathbf{q})} e^{-i\mathbf{k} \cdot \mathbf{q}}. \quad (6)$$

Note that here we assume that all quantities are evaluated in linear theory at the final redshift and we suppress explicit time dependence.

Starting from this equation, one can compute the cross-correlation coefficient between the leading-order LPT and EPT fields. It is given by

$$r^{\text{LO}}(k) \equiv \frac{\langle \delta_{\text{EPT}}^{\text{LO}}(\mathbf{k}') \delta_{\text{LPT}}^{\text{LO}}(\mathbf{k}) \rangle'}{\sqrt{P_{\text{EPT}}(k) P_{\text{LPT}}(k)}} \approx e^{-\frac{1}{2} k^2 \sigma_v^2}, \quad (7)$$

where prime on the correlation function indicates that $(2\pi)^3 \delta^D(\mathbf{k} + \mathbf{k}')$ has been removed, $P_{\text{EPT}}(k)$ and $P_{\text{LPT}}(k)$ are leading-order power spectra, and the velocity dispersion is defined as

$$\langle \boldsymbol{\psi}_i(\mathbf{q}) \boldsymbol{\psi}_j(\mathbf{q}) \rangle' \equiv \sigma_v^2 \delta_{ij}^K, \quad (8)$$

such that it is given by

$$\sigma_v^2(z) = \frac{1}{6\pi^2} \int_0^\infty dq P_{\text{lin}}(q, z). \quad (9)$$

Note that what is commonly referred to as the velocity dispersion is actually related to the variance of the displacement field and therefore has units of length. Derivation of Eq. (7) is given in Appendix A. Let us here provide an intuitive justification for it. The major difference between EPT and LPT is the displacement field. Therefore, on scales larger than the typical variance of the displacement field, given by σ_v^2 , the EPT and LPT fields are the same. On smaller scales they rapidly decorrelate and $r^{\text{LO}}(k)$ goes to zero. The value of σ_v^2 in a Λ CDM-like cosmology is approximately

$$\sigma_v^2(z) = 36 \left(\frac{D_+(z)}{D_+(0)} \right)^2 (\text{Mpc}/h)^2, \quad (10)$$

where the time dependence is inherited from the linear power spectrum in Eq. (9). This implies a significant decorrelation between the leading-order EPT and LPT fields, even at $k_{\text{max}} = 0.1 \text{ h}/\text{Mpc}$.

However, in our analysis we will go beyond the leading-order approximation. If we use cubic models for both EPT and LPT, the agreement between the two improves. In this case, a part of the nonlinearities induced by the large displacements is taken into account in EPT. Since the two theories match at one-loop order on large scales, the cross-correlation coefficient at next-to-leading order (NLO) will be better and must scale as

$$r^{\text{NLO}}(k) = 1 + \mathcal{O}(k^4 \sigma_v^4), \quad (11)$$

in the low- k limit, where $\mathcal{O}(k^4 \sigma_v^4)$ is the leading contribution from the unaccounted two-loop terms. While deriving the exact expression is more involved in this case, we can provide a simple estimate of the cross-correlation coefficient at NLO. In a way identical to what happens in the so-called infrared resummation, we can write

$$r^{\text{NLO}}(k) \approx \left(1 + \frac{1}{2} k^2 \sigma_v^2 \right) e^{-\frac{1}{2} k^2 \sigma_v^2}, \quad (12)$$

which is guaranteed to produce the correct low- k limit for the cubic EPT and LPT models, while keeping the same exponential suppression on small scales. In Fig. 2 we plot the cross-correlation coefficient for linear and cubic models, corresponding to Eqs. (7) and (12), respectively. Our estimates agree very well with measurements of the cross-correlation coefficient from explicit EPT and LPT maps built at different order in perturbation theory (for example, see Fig. 10 in [41]). Going to cubic order in EPT and having

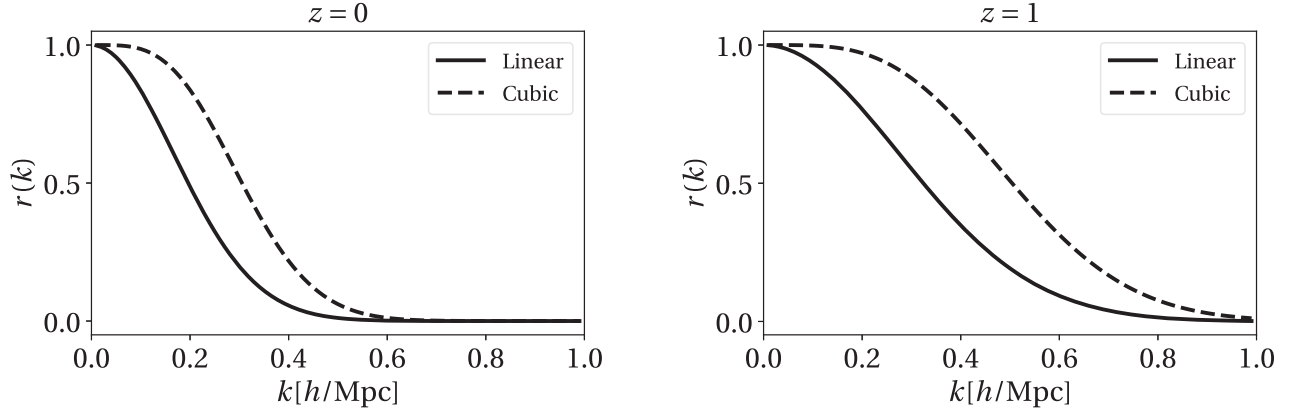


FIG. 2. Estimate of the cross-correlation coefficient between the EPT and LPT fields when both theories are computed up to same order in perturbation theory (linear or cubic model), at $z = 0$ (left) and $z = 1$ (right). These curves have the usual s shape and are in excellent agreement with the findings from the explicit grid computations of EPT and LPT (for example, see Fig. 10 in [41]).

the correct one-loop power spectrum makes the fields correlate much better.

The cross-correlation coefficient is related to the error that one makes approximating LPT with EPT. If we define the “error” field as the difference $\delta_{\text{err}} \equiv \delta_{\text{LPT}} - \delta_{\text{EPT}}$, then the power spectrum of this difference can be approximated as follows:

$$P_{\text{err}}(k) = \langle |\delta_{\text{LPT}}(\mathbf{k}) - \delta_{\text{EPT}}(\mathbf{k})|^2 \rangle' = P_{\text{LPT}}(k) + P_{\text{EPT}}(k) - 2\sqrt{P_{\text{LPT}}(k)P_{\text{EPT}}(k)}r(k) \approx P_{\text{PT}}(k)(1 - r^2(k)), \quad (13)$$

where in order to make the last approximation we used that $P_{\text{LPT}} \approx P_{\text{EPT}} = P_{\text{PT}}$ and that $r \approx 1$ on large scales, such that

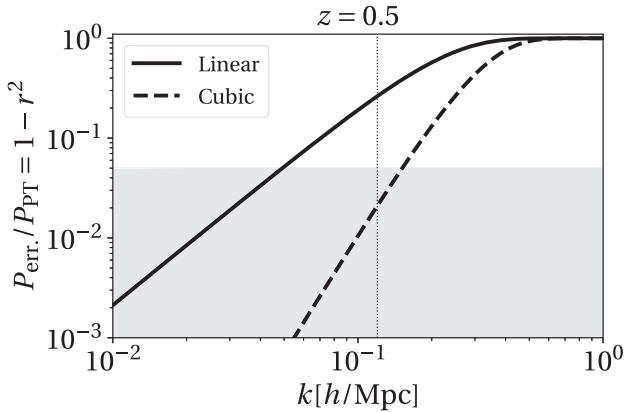


FIG. 3. Power spectrum of the difference between LPT and EPT fields at $z = 0.5$ in linear and cubic models, normalized by P_{PT} , see Eq. (13). The shaded region corresponds to a relative error smaller than 5%. Note that $P_{\text{err}}/P_{\text{PT}} \approx 2\Delta A/A$. Thin vertical line marks the maximum wave number used in the analysis, $k_{\text{max}} = 0.12 \text{ h/Mpc}$. At higher redshifts the agreement between EPT and LPT is even better.

$2(1 - r) \approx (1 + r)(1 - r) = 1 - r^2$. In other words, the quantity $1 - r^2$ measures the relative size of the error power spectrum compared to the P_{PT} . In Fig. 3, we plot $1 - r^2$ for redshift $z = 0.5$. We can see that the relative error of approximating LPT with EPT is at the percent level for the scale cuts used in the analyses, $k_{\text{max}} = 0.1$ and $k_{\text{max}} = 0.12 \text{ h/Mpc}$. This is much smaller than the $\mathcal{O}(5\%)$ relative error on the amplitude of the linear field that can be obtained from a $(2 \text{ Gpc/h})^3$ volume box of biased tracers in real space [35]. This implies that, on large scales and within given statistical errors, Eulerian and Lagrangian perturbative forward models are practically indistinguishable.

One may object that the arguments we presented so far apply only to the dark matter field. However, the situation for the biased tracers is in some sense even better, thanks to the additional source of the error due to the discrete sampling of the halo density field. Assuming constant shot noise power spectrum on large scales with the amplitude P_{noise} , we can write the ratio of the total error⁵ power spectrum to the noise as follows:

$$\frac{P_{\text{err}}}{P_{\text{noise}}} \approx 1 + (1 - r^2) \frac{b_1^2 P_{\text{lin}}(k)}{P_{\text{noise}}}. \quad (14)$$

As expected, for biased tracers with large enough shot noise, all modeling errors are negligible compared to the shot noise P_{noise} on large scales, when r goes to 1. In the left panel of Fig. 4, we plot the ratio from the previous formula for a very dense tracer with $\bar{n} \sim 4.3 \times 10^{-2} \text{ (h/Mpc)}^3$. This agrees very well with measurements from simulations [for example, see the dark orange line on the left panel of Fig. (20) in [12]]. In the right panel of Fig. 4 we show the

⁵For galaxies, we assume $\delta_{g,\text{LPT}} - \delta_{g,\text{EPT}} = \delta_{\text{err}} + \epsilon_g$, where $\langle \epsilon_g \epsilon_g \rangle' = P_{\text{noise}}$.

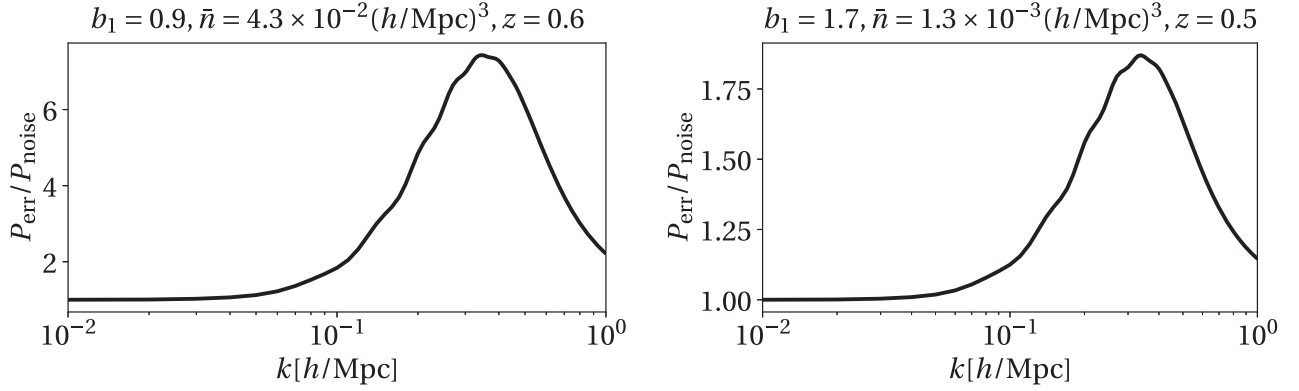


FIG. 4. $P_{\text{err}}/P_{\text{noise}}$ as function of scale for two biased tracers similar the ones considered in [12] (left) and in [35] (right).

same ratio but for the tracer considered in [35]. While the modeling error is roughly 10% larger than the sampling noise at $k_{\text{max}} = 0.12 \text{ h/Mpc}$, the sampling noise itself is just a small fraction of the total power spectrum on these scales. Therefore, the cross-correlation coefficient between EPT and LPT can be very easily understood both for dark matter field and biased tracers. In both cases, the two perturbative approaches agree at the percent level on relevant scales.

Let us also briefly discuss the second reason why EPT is a good approximation for LPT on large scales. So far we estimated the cross-correlation coefficient and the error power spectrum and showed that on scales relevant for our analysis the EPT and LPT fields match at the percent level. However, in the FBI one performs the marginalization over all initial conditions, resulting in effective measurement of some summary statistics from the maps. In the perturbative regime, these statistics are just the leading n -point functions [38]. Regardless of the details, we want to argue that the estimate of the mismatch based on the cross-correlation coefficient is the most pessimistic one. It is dominated by the velocity dispersion induced by the large bulk flows. However, it is well known that due to the equivalence principle the effects of these bulk flows exactly cancel in all summary statistics (with the exception of features such as BAO, which is irrelevant on large scales) [42–48]. Therefore, the agreement between EPT and LPT predictions can be only better in practice than the $1 - r^2$ estimate indicates. For example, even though the Zel’dovich and linear Eulerian fields strongly decorrelate for $k \gtrsim 1/\sigma_v$, the two power spectra do not differ dramatically. The same is true for any higher-order correlation function or other statistical quantities.

In conclusion, for the range of scales and redshifts used in [35] and our analysis in this paper, the standard Eulerian perturbation theory (without the IR resummation) and Lagrangian perturbation theory describe the halo density field equally well. Since there is no distinction between the two nonlinear fields on the level of summary statistics, all conclusions based on EPT must be valid for LPT as well.

B. Halo density field in Eulerian perturbation theory

Given the arguments that justify using the EPT, we now turn to the model for the galaxy density field. In EPT the nonlinear density field of biased tracers in real space on large scales has a simple form in terms of local⁶ operators [52],

$$\begin{aligned} \delta_g^{\text{model}}(\mathbf{x}) = & b_1 \delta(\mathbf{x}) + \frac{b_2}{2} (\delta^2(\mathbf{x}) - \langle \delta^2 \rangle) + b_{\mathcal{G}_2} \mathcal{G}_2(\mathbf{x}) \\ & + \frac{b_3}{6} \delta^3(\mathbf{x}) + b_{\mathcal{G}_3} \mathcal{G}_3(\mathbf{x}) + b_{\mathcal{G}_2\delta} \mathcal{G}_2(\mathbf{x}) \delta(\mathbf{x}) \\ & + b_{\Gamma_3} \Gamma_3(\mathbf{x}) - c_s^2 \frac{\nabla^2}{k_{\text{NL}}^2} \delta(\mathbf{x}), \end{aligned} \quad (15)$$

where δ is the nonlinear matter field (to be computed in Eulerian perturbation theory up to cubic order) and bias operators $\mathcal{G}_2$, $\mathcal{G}_3$, and Γ_3 are built from tidal fields of gravitational or velocity potential. These terms are sufficient to describe the one-loop power spectrum, tree-level bispectrum, and tree-level trispectrum and they match the cubic Lagrangian perturbation theory. Note that the last contribution in Eq. (15) is a higher derivative term, which has a different expansion parameter, given by k/k_{NL} and not by the variance of δ . For a Λ CDM-like power spectrum this term is comparable to other cubic contributions and has to be kept. It is important to remember that there are two physically distinct effects that are described by the higher derivative operator. One is the response of dark matter halos to $\nabla^2 \delta$ and the other one is the dark matter one-loop counterterm. Since at this order in perturbation theory these

⁶Since dark matter halo formation happens over a long time with the characteristic scale $t_{\text{gal form}} \sim H^{-1}$, the overdensity of tracers at a given time depends on the entire history of the long-wavelength dark matter density field evolution, making the most general bias expansion nonlocal in time [4,5,49–51]. However, it has been shown that up to cubic order these nonlocal contributions can be rewritten such that the expansion (15) looks local [4,5].

two effects are completely degenerate, we will keep only one free parameter in our analysis.⁷

In summary, for the purposes of this work, we will use eight unknown parameters that depend on the halo or galaxy properties and have to be measured from the data. These are

$$\theta_{\text{bias}} = \{b_1, b_2, b_{\mathcal{G}_2}, b_3, b_{\mathcal{G}_3}, b_{\mathcal{G}_2\delta}, b_{\Gamma_3}, c_s^2\}. \quad (16)$$

Biases and counterterms are (strongly) degenerate with some cosmological parameters and measuring multiple correlation functions is crucial to break these degeneracies.

In our analysis, we will focus only on one cosmological parameter, which measures the amplitude of the linear density fluctuations A . We define the following parameter:

$$\alpha \equiv \frac{A}{A^{\text{fid}}} = \frac{\sigma_8}{\sigma_8^{\text{fid}}}, \quad (17)$$

where the second equality holds if all other cosmological parameters are fixed as in [35].

We will do all calculations and the analysis in Fourier space. Given the real space model and parameters, we can write

$$\delta_g^{\text{model}}(\mathbf{k}, z) = \sum_n \alpha^n D^n(z) \int_{q_1} \cdots \int_{q_n} (2\pi)^3 \delta^D(\mathbf{k} - \mathbf{q}_{1\dots n}) K_n(\mathbf{q}_1, \dots, \mathbf{q}_n; \theta) \delta_1(\mathbf{q}_1) \cdots \delta_1(\mathbf{q}_n) - \alpha c_s^2 \frac{k^2}{k_{\text{NL}}^2} D(z) \delta_1(\mathbf{k}), \quad (18)$$

where $\int_{\mathbf{q}} \equiv \int d^3\mathbf{q}/(2\pi)^3$, $\mathbf{q}_{1\dots n} \equiv \mathbf{q}_1 + \cdots + \mathbf{q}_n$, $D(z) \equiv D_+(z)/D_+(0)$ is the normalized growth factor at redshift z such that $D(0) = 1$, $\delta_1(\mathbf{k})$ is the linear density field, and K_n are perturbation theory kernels that contain all information about the nonlinear evolution. These kernels are well known in the literature (see, for example, [3,4,31,53]) and here we just write them down for completeness. They are given as

$$K_1(\mathbf{q}; \theta) = b_1, \quad (19)$$

$$K_2(\mathbf{q}_1, \mathbf{q}_2; \theta) = b_1 F_2(\mathbf{q}_1, \mathbf{q}_2) + \frac{b_2}{2} + b_{\mathcal{G}_2} \left(\frac{(\mathbf{q}_1 \cdot \mathbf{q}_2)^2}{q_1^2 q_2^2} - 1 \right), \quad (20)$$

$$\begin{aligned} K_3(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3; \theta) = & b_1 F_3(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) + \frac{b_2}{3} (F_2(\mathbf{q}_1, \mathbf{q}_2) + F_2(\mathbf{q}_1, \mathbf{q}_3) + F_2(\mathbf{q}_2, \mathbf{q}_3)) + 2b_{\mathcal{G}_2} F_{\mathcal{G}_2}(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) \\ & + \frac{b_3}{6} + b_{\mathcal{G}_3} F_{\mathcal{G}_3}(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) + b_{\Gamma_3} F_{\Gamma_3}(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) + \frac{b_{\mathcal{G}_2\delta}}{3} \left(\frac{(\mathbf{q}_1 \cdot \mathbf{q}_2)^2}{q_1^2 q_2^2} + \frac{(\mathbf{q}_1 \cdot \mathbf{q}_3)^2}{q_1^2 q_3^2} + \frac{(\mathbf{q}_2 \cdot \mathbf{q}_3)^2}{q_2^2 q_3^2} - 3 \right). \end{aligned} \quad (21)$$

In these expressions, F_2 and F_3 are the standard perturbation theory kernels, see Ref. [54], and additional terms are defined by

$$F_{\mathcal{G}_2}(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) \equiv \frac{1}{3} F_2(\mathbf{q}_1, \mathbf{q}_2) \left(\frac{((\mathbf{q}_1 + \mathbf{q}_2) \cdot \mathbf{q}_3)^2}{|\mathbf{q}_1 + \mathbf{q}_2|^2 q_3^2} - 1 \right) + 2 \text{ perm.}, \quad (22)$$

$$F_{\mathcal{G}_3}(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) \equiv -\frac{1}{2} \left(1 - \frac{(\mathbf{q}_1 \cdot \mathbf{q}_2)^2}{q_1^2 q_2^2} - \frac{(\mathbf{q}_1 \cdot \mathbf{q}_3)^2}{q_1^2 q_3^2} - \frac{(\mathbf{q}_2 \cdot \mathbf{q}_3)^2}{q_2^2 q_3^2} + 2 \frac{(\mathbf{q}_1 \cdot \mathbf{q}_2)(\mathbf{q}_1 \cdot \mathbf{q}_3)(\mathbf{q}_2 \cdot \mathbf{q}_3)}{q_1^2 q_2^2 q_3^2} \right), \quad (23)$$

$$F_{\Gamma_3}(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) \equiv -\frac{4}{21} \left(\frac{(\mathbf{q}_1 \cdot \mathbf{q}_2)^2}{q_1^2 q_2^2} - 1 \right) \left(\frac{((\mathbf{q}_1 + \mathbf{q}_2) \cdot \mathbf{q}_3)^2}{|\mathbf{q}_1 + \mathbf{q}_2|^2 q_3^2} - 1 \right) + 2 \text{ perm.} \quad (24)$$

⁷In principle, there are two different scales associated with the two different effects. The dark matter counterterm has the nonlinear scale k_{NL} , while the dark matter halos have another scale k_* related to the Lagrangian size of a typical halo in the sample. For massive halos, these two can be very different, but we will have in mind dense samples as in [35] where the two scales are comparable. Anyway, since we assume that the higher derivative bias and the counterterm are both unknown, this simplification does not impact our analysis.

The kernels K_n can be used to compute the one-loop power spectrum, tree-level bispectrum, and tree-level trispectrum. We will do that in the next section. Before that, we discuss the second important ingredient, which is the stochastic noise.

C. Stochastic noise

As we already explained, the galaxy density field in perturbation theory is not fully deterministic, but it also

contains stochastic noise which comes from our ignorance of the small-scale modes that determine galaxy formation and that are not a part of observations (or perturbation theory). This noise does not correlate with long-wavelength density fluctuations, but it can be modulated by the long modes. Therefore, the noise field can be written in a similar manner as the bias expansion

$$\epsilon(\mathbf{x}) = \epsilon_0(\mathbf{x}) + \epsilon_1(\mathbf{x})\delta(\mathbf{x}) + \frac{1}{2}\epsilon_2(\mathbf{x})(\delta^2(\mathbf{x}) - \langle\delta^2\rangle) + \epsilon_{\mathcal{G}_2}(\mathbf{x})\mathcal{G}_2(\mathbf{x}) + \dots \quad (25)$$

Note that $\epsilon_i(\mathbf{x})$ are non-Gaussian fields that can have significant three-point functions. It is clear from this expression that $\langle\delta\epsilon\rangle = 0$ as expected. The autopower spectrum of the noise does not inherit the “shape” of the cross spectra from δ , δ^2 , or $\mathcal{G}_2$ due to the fact that the stochastic fields correlate only in the same spatial point

$$\langle\epsilon_i(\mathbf{x})\epsilon_j(\mathbf{x}')\rangle \sim c_{ij}\delta^D(\mathbf{x} - \mathbf{x}'). \quad (26)$$

On large scales, the noise power spectrum is approximately constant, with the amplitude approximately equal to the

Poisson prediction

$$P_{\epsilon\epsilon}(k) = \frac{c_1}{\bar{n}_g}, \quad (27)$$

where c_1 is a free dimensionless parameter of order 1. When computing higher-order n -point functions, there are several additional terms that may appear and we discuss them in the next section.

III. GALAXY POWER SPECTRUM, BISPECTRUM, AND TRISPECTRUM

Equipped with the perturbative forward model and the noise model, we are now ready to compute all relevant correlation functions. In this section, we write down the one-loop power spectrum, a simple model for the BAO damping that we will use in the BAO analysis, tree-level bispectrum, and tree-level trispectrum.

A. Power spectrum

The power spectrum of biased tracers in real space up to one-loop corrections is given by

$$\begin{aligned} P_g(k, z) = & b_1^2\alpha^2 P_{\text{lin}}(k, z) + b_1^4\alpha^4 P_{13}(k, z) + b_1^2\alpha^4 P_{22}(k, z) - b_1^2c_s^2\alpha^2 k^2 P_{\text{lin}}(k, z) \\ & + b_1b_2\alpha^4 \mathcal{I}_{\delta^2}(k, z) + \frac{b_2^2}{4}\alpha^4 \mathcal{I}_{\delta^2\delta^2}(k, z) + b_2^2\alpha^4 \mathcal{I}_{\mathcal{G}_2\mathcal{G}_2}(k, z) + \frac{b_2b_{\mathcal{G}_2}}{2}\alpha^4 \mathcal{I}_{\delta^2\mathcal{G}_2}(k, z) \\ & + \left(2b_1b_{\mathcal{G}_2} + \frac{4}{5}b_1b_{\Gamma_3}\right)\alpha^4 \mathcal{F}_{\mathcal{G}_2}(k, z) + 2b_1b_{\mathcal{G}_2}\alpha^4 \mathcal{I}_{\mathcal{G}_2}(k, z) + \frac{c_1}{\bar{n}_g}, \end{aligned} \quad (28)$$

where the explicit form of different contributions can be found in [3,55]. All nuisance parameters in this expression are functions of time, but for the fixed redshift they can be treated as constants. This form of the power spectrum is particularly convenient since all nuisance and cosmological parameters enter the model just as simple multiplicative factors. At the level of the linear theory, b_1 and α are perfectly degenerate and only the combination $b_1\alpha$ can be measured with very high precision. This degeneracy can be, in principle, broken only by the loop contributions, but those have much smaller signal-to-noise ratio, suppressed by the variance of the density field at the maximum wave number used in the analysis. In addition, since there are several new nuisance parameters at the loop level, the amplitude α is expected to be poorly constrained from the P alone. We will verify this explicitly below.

Let us stress that the correct model for the galaxy power spectrum contains the infrared resummation, which correctly accounts for the broadening of the BAO peak. However, we will not use it in our baseline analysis for several reasons. First, since we have shown that Eulerian and Lagrangian fields are almost the same up to $k_{\text{max}} \sim 0.1 \text{ h/Mpc}$, we know that the infrared resummation

cannot significantly affect the power spectrum on scales and redshifts that we use. Second, we will test this simple model against the large-volume PT challenge simulations suite and explicitly show the validity of our approximation. Finally, even if this approximation was to miss small, percent-level effects, that would be irrelevant for the conclusions regarding the expected errors from a joint P + B and P + B + T analyses.

Clearly, we can neglect the infrared resummation only on very large scales. On small scales, there is an additional information about the amplitude of the linear density field in the broadening of the BAO wiggles. While this is not included in the analysis of [35] and it is not the main goal of this paper, we explore how much information on α can be extracted from small scales using BAO. We next give a simple model of the IR resummation at small scales appropriate for our forecast.

B. BAO wiggles

One nontrivial way in which the amplitude α affects the BAO is through the broadening of the BAO peak. This non-linear effect is very well understood and can be described to very high precision in perturbation theory [46–48,56].

First, one can decompose the linear power spectrum into the smooth and wiggly parts [55,57–59],

$$P_{\text{lin}}(k, z) = P_{\text{lin}}^{\text{nw}}(k, z) + P_{\text{lin}}^{\text{w}}(k, z). \quad (29)$$

The bulk flows induced by the long wavelength fields affect the features such as BAO, leading to the damping of the oscillations. The full expression for the infrared-resummed galaxy power spectrum is schematically given by

$$P_g^{\text{IR}}(k, z) = P_g^{\text{nw}}(k, z) + e^{-\alpha^2 \Sigma_\Lambda^2(z) k^2} (1 + \alpha^2 \Sigma_\Lambda^2(z) k^2) \times \alpha^2 b_1^2 P_{\text{lin}}^{\text{w}}(k, z) + e^{-\alpha^2 \Sigma_\Lambda^2(z) k^2} P_{g,1\text{-loop}}^{\text{w}}(k, z), \quad (30)$$

where P_g^{nw} is the standard EPT power spectrum up to one loop computed using $P_{\text{lin}}^{\text{nw}}$, and $P_{g,1\text{-loop}}^{\text{w}}$ is the wiggly part of the EPT one-loop power spectrum. The BAO damping factor is given by [48]

$$\Sigma_\Lambda^2(z) \equiv \frac{1}{6\pi^2} \int_0^\Lambda dq P_{\text{lin}}(q, z) \times [1 - j_0(q\ell_{\text{BAO}}) + 2j_2(q\ell_{\text{BAO}})], \quad (31)$$

where the BAO scale is $\ell_{\text{BAO}} \approx 110 \text{ h/Mpc}$, $j_n(x)$ are spherical Bessel functions of order n , and Λ is the scale up to which the long-wavelength displacements are taken into account. The typical value used in practice is $\Lambda = 0.2 \text{ h/Mpc}$, but the result is almost insensitive to this choice [48,55,60]. Given the numerical values of $\Sigma_\Lambda^2(z)$ for relevant redshifts, the damping is important only for high k . Note that the formula for the wiggly part of the galaxy power spectrum is nonperturbative and almost perfectly describes the BAO. Therefore, unlike the broadband, it can be safely used at all scales.

It is clear from Eq. (30) how the parameter α can be measured from the BAO wiggles. Apart from the small loop corrections controlled by the perturbation theory parameter $\Delta^2(k)$, α multiplies the large parameter Σ_Λ^2 which controls the broadening of the BAO peak. In other words, the amplitude of the linear density field can be simply estimated from how much the nonlinear evolution damps the BAO wiggles. This is easy to intuitively understand in real space. On large scales, one can approximate the two-point correlation function using only the linear theory. From the smooth part, one can measure the combination $b_1\alpha$.

Then, focusing on the BAO peak only, the broadening is controlled by $\alpha^2 \Sigma_\Lambda^2$ which has only one free parameter. This is the essence of the statement that large displacements can be used to measure α without strong sensitivity on nonlinear galaxy bias if a feature is present in the initial conditions. Crucially, this is true in Fourier space only for high- k part of the BAO wiggles, beyond $k_{\text{max}} \approx 0.2 \text{ h/Mpc}$.

It is possible to test how much information on α comes from the BAO wiggles using the formulas above. While one can use Eq. (30) directly, just to keep the discussion as simple as possible, we will use in our forecast the following simplified model for the IR-resummed wiggly part of the galaxy power spectrum,

$$P_g^{\text{w}}(k, z) = e^{-\alpha^2 \Sigma_\Lambda^2(z) k^2} \alpha^2 b_1^2 P_{\text{lin}}^{\text{w}}(k, z) + c_w^2 \alpha^2 b_1^2 k^2 P_{\text{lin}}^{\text{w}}(k, z). \quad (32)$$

This equation phenomenologically captures all features of the exact result and in a Fisher forecast it is simpler to implement and interpret. We introduce a single parameter c_w^2 that encompasses all nonlinear bias and possible other nonlinear corrections and it is of order $\mathcal{O}(1)$. The wiggly part of the galaxy power spectrum at high k can then be analyzed separate from the broadband (see, for example, [61–64]).

We will consider two different setups.

- (i) **Optimistic:** In this setup, we set $c_w = 0$ and $\Lambda \rightarrow \infty$. This is the best case scenario, since we neglect all nonlinear corrections with bias parameters and we replace Σ_Λ^2 with Σ_∞^2 , to approximately take into account the total broadening from the displacement and loop corrections. This makes the effect as large as possible and the constraints as good as possible.
- (ii) **Realistic:** In this setup, we set $\Lambda = 0.2 \text{ h/Mpc}$ and leave c_w^2 as a free parameter with a Gaussian prior of width equal to 1. In this way, the information on α comes only from the part of the broadening induced with large displacements and it is to some extent degenerate with other nonlinearities.

C. Bispectrum and trispectrum

Next, we describe the higher-order n -point functions that can be consistently computed using the cubic model. The tree-level bispectrum in real space is given by [54]

$$B_g(k_1, k_2, k_3, z) = 2b_1^2 \alpha^4 K_2(\mathbf{k}_1, \mathbf{k}_2) P_{\text{lin}}(k_1, z) P_{\text{lin}}(k_2, z) + 2 \text{ perms.} + \frac{c_2}{\bar{n}_g} b_1 \alpha^2 P_{\text{lin}}(k_1, z) + 2 \text{ perms.} + \frac{c_3}{\bar{n}_g^2}. \quad (33)$$

Note that there are two additional coefficients in the noise model compared to the power spectrum. These three free parameters c_1 , c_2 , and c_3 exactly match the noise parameters used in the analysis of [35].

The only additional n -point correlation function that can be fully computed using cubic-order Eulerian kernels is the connected tree-level trispectrum, which is given by

$$\begin{aligned}
T_g^c(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4, z) = & 4b_1^2 \alpha^6 P_{\text{lin}}(k_1, z) P_{\text{lin}}(k_2, z) (K_2(\mathbf{k}_1, -\mathbf{k}_{13}) K_2(\mathbf{k}_2, \mathbf{k}_{13}) P_{\text{lin}}(k_{13}, z) \\
& + K_2(\mathbf{k}_1, -\mathbf{k}_{14}) K_2(\mathbf{k}_2, \mathbf{k}_{14}) P_{\text{lin}}(k_{14}, z)) + 5 \text{ perm.} \\
& + 6b_1^3 \alpha^6 K_3(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) P_{\text{lin}}(k_1, z) P_{\text{lin}}(k_2, z) P_{\text{lin}}(k_3, z) + 3 \text{ perm.} \\
& + \frac{c_2}{\bar{n}_g} B_{\text{m}gg}(\mathbf{k}_{12}, \mathbf{k}_3, \mathbf{k}_4, z) + 5 \text{ perm.}, \tag{34}
\end{aligned}$$

where we have defined

$$\begin{aligned}
B_{\text{m}gg}(\mathbf{k}_{12}, \mathbf{k}_3, \mathbf{k}_4, z) \equiv & \langle \delta_m(\mathbf{k}_{12}, z) \delta_g(\mathbf{k}_3, z) \delta_g(\mathbf{k}_4, z) \rangle' \\
= & 2b_1^2 \alpha^4 F_2(\mathbf{k}_3, \mathbf{k}_4) P_{\text{lin}}(k_3, z) P_{\text{lin}}(k_4, z) + 2b_1 \alpha^4 K_2(\mathbf{k}_{12}, \mathbf{k}_4) P_{\text{lin}}(k_4, z) P_{\text{lin}}(k_{12}, z) \\
& + 2b_1 \alpha^4 K_2(\mathbf{k}_{12}, \mathbf{k}_3) P_{\text{lin}}(k_3, z) P_{\text{lin}}(k_{12}, z). \tag{35}
\end{aligned}$$

Note that we keep only the largest noise contribution which at leading order has the same parameter c_2 as the bispectrum. We do this for two reasons: (a) in order to be able to have a fair comparison to the field-level analysis of [35], which does not use the most general noise model, and (b) in order to get the most optimistic results for additional constraining power of the trispectrum. Other (subdominant) noise contributions to the trispectrum can be derived from the noise model in Eq. (25) and we do this in Appendix B. Even though their impact is not expected to be large (given reasonable priors on the amplitude of the noise), they would lead to slightly larger error bars in the trispectrum analysis.

IV. LIKELIHOODS

In this section we present the likelihoods for the power spectrum, bispectrum, trispectrum, and BAO wiggles that we use in our analyses. Given a very simple form of the likelihood in the EPT we also provide a simple mode counting argument for the expected joint power spectrum and bispectrum analysis.

A. Power spectrum

The power spectrum likelihood on large scales can be approximated by the well-known equation

$$\ln \mathcal{L}_P = -\frac{1}{2} \sum_{k=k_{\min}}^{k=k_{\max}} \frac{(P_g(k, \theta) - \hat{P}_g(k))^2}{\text{Var}_P(k)}. \tag{36}$$

Note that we assume the Gaussian form of the likelihood that is appropriate on large scales. The sum runs over all k bins up to a given $k_{\max}$. The set of variables θ contains all nuisance and cosmological parameters in the analysis, while the measured power spectrum in a given k bin is $\hat{P}_g(k)$. Note that for simplicity we neglect the explicit redshift dependence. Furthermore, we neglect the cosmological parameter dependence in the covariance matrix, since it carries negligible signal in the limit when the

number of Fourier modes is much larger than the number of bins in the data vector. We will use this approximation for all other statistics as well.

From the way we write the power spectrum likelihood, it is clear that we assume a simple diagonal covariance matrix, whose elements are given by

$$\text{Var}_P(k) = \frac{4\pi^2}{V \Delta k k^2} P_g^2(k, \theta_{\text{fid}}), \tag{37}$$

where V is the volume of the survey, Δk is the width of the k bins, and θ_{fid} is the fiducial value of cosmological and nuisance parameters. On large scales used in our analysis, it is sufficient to use the linear theory in the computation of the covariance. All nonlinear corrections and off-diagonal elements in the covariance matrix are small on large scales and suppressed by the variance of the density field $\Delta^2(k_{\max})$. Furthermore, since we are interested in the posteriors of cosmological parameters after marginalization over all nuisance parameters, the final variance of cosmological parameters is not impacted only by the data covariance, but also the effective covariance due to the marginalization. This was explicitly demonstrated in some simple examples in [65]. This additional uncertainty is usually much bigger than any mistake made in the data covariance. For this reason, one recovers identical error bars on cosmological parameters on large scales in analyses based on perturbation theory, regardless of the choice of the data covariance. The explicit demonstration of this fact can be found in [66].

B. Bispectrum

Similar to the power spectrum, we assume a simple Gaussian likelihood with the diagonal covariance matrix for the bispectrum as well

$$\ln \mathcal{L}_B = -\frac{1}{2} \sum_{\mathcal{T}} \frac{(B_g(k_1, k_2, k_3, \theta) - \hat{B}_g(k_1, k_2, k_3))^2}{\text{Var}_B(k_1, k_2, k_3)}, \tag{38}$$

where the sum now runs over all triangles in Fourier space. In the Gaussian approximation only the same shape triangles correlate and the diagonal elements of the covariance matrix are given by

$$\text{Var}_B(k_1, k_2, k_3) = \frac{(2\pi)^6}{VV_{123}} s_{123} P_g(k_1, \theta_{\text{fid}}) P_g(k_2, \theta_{\text{fid}}) \times P_g(k_3, \theta_{\text{fid}}), \quad (39)$$

where s_{123} is a symmetry factor and it is equal to 6, 2, and 1 for equilateral, isosceles, and scalene triangles, respectively, and V_{123} is the volume of a given triangle bin which in the thin shell approximation reads

$$V_{123} = 8\pi^2 k_1 k_2 k_3 \Delta k^3. \quad (40)$$

Note that we assume ordered momenta, $k_3 \leq k_2 \leq k_1$, in order to avoid the double counting of equivalent triangles.

For reasons similar to the power spectrum case, one can justify the simple approximation for the bispectrum likelihood in perturbation theory. Corrections to the Gaussian approximation are expected to be of order $\Delta^2(k_{\text{max}})$ and cannot significantly change the results in the full analysis. We will comment more on this and the cross-covariance with the power spectrum below.

C. Trispectrum

While the form of the power spectrum and bispectrum likelihoods in the Gaussian approximation is well known, the trispectrum is much less used and explored in the comparisons to simulations and data (for some recent progress, see Refs. [67,68]). First of all, the connected trispectrum is a function of six scalar variables. It can be parametrized in different ways, but we will use four magnitudes of the wave numbers $k_1, \dots, k_4$ and two magnitudes of the diagonals D_1 and D_2 , defined as

$$D_1 \equiv -k_1 - k_2 \quad \text{and} \quad D_2 \equiv -k_2 - k_3. \quad (41)$$

The full data vector that depends on all momenta and diagonals is very long and has very large dimensional covariance matrix. For these reasons the compressed trispectrum is sometimes used in practice. One such example is the integrated trispectrum, where in the estimator one averages over all D_1 and D_2 for the fixed momenta $k_1, \dots, k_4$. Using this estimator significantly reduces the length of the data vector and allows for a simple evaluation of the covariance matrix. The integrated trispectrum has been used both in real and redshift space, for comparison of perturbation theory predictions to simulations [67,68].

In this paper, we want to estimate the impact of the trispectrum on analyses on large scales and find the most optimal possible improvement. For this reason we cannot

use the simplified estimators such as integrated trispectrum, where some information is lost due to integration over the diagonals. This is particularly relevant for understanding the impact of the trispectrum on constraining all cubic bias parameters, where the only hope to do it comes from using the full shape keeping all dependence on different angles. Given this, we will use the full connected trispectrum and the following likelihood:

$$\ln \mathcal{L}_T = -\frac{1}{2} \sum_{\mathcal{Q}} \frac{(T_g^c(\mathcal{Q}, \theta) - \hat{T}_g^c(\mathcal{Q}))^2}{\text{Var}_T(\mathcal{Q})}, \quad (42)$$

where the sum runs over all nonequivalent quadrilaterals $\mathcal{Q}$ defined by the six sides $\mathcal{Q} \equiv \{k_1, k_2, k_3, k_4, D_1, D_2\}$. We are again using the Gaussian approximation, in which only the same shape quadrilaterals correlate. Note that, in order to avoid the double counting, the sides of the quadrilaterals are ordered as $k_4 \leq k_3 \leq k_2 \leq k_1$, while the diagonals must satisfy the relations $|k_1 - k_2| \leq D_1 \leq |k_1 + k_2|$ and $|k_2 - k_3| \leq D_2 \leq |k_2 + k_3|$.

The diagonal elements of the connected trispectrum covariance matrix can be computed for the general $\mathcal{Q}$ and we give the derivation in Appendix C. They are given by

$$\text{Var}_T(\mathcal{Q}) = \frac{(2\pi)^9 s_{1234}}{VV_{1234}} P_g(k_1, \theta_{\text{fid}}) P_g(k_2, \theta_{\text{fid}}) P_g(k_3, \theta_{\text{fid}}) \times P_g(k_4, \theta_{\text{fid}}), \quad (43)$$

where $s_{1234} = n!$ and $0 \leq n \leq 4$ is the number of momenta $k_1, \dots, k_4$ with the same magnitude. The nontrivial result is the volume of the quadrilateral shell V_{1234} , which reads

$$V_{1234} = \frac{8}{3} \pi^2 \frac{k_1 k_2 k_3 k_4 D_1 D_2}{V_{\text{tetra}}} \Delta k^6. \quad (44)$$

In this expression V_{tetra} is the volume of the tetrahedron $\mathcal{Q}$ (for details, see Appendix C). This is a very simple formula, which allows for an easy evaluation of the trispectrum likelihood.

We notice that some of the considered tetrahedra have the six sides which lie on the same plane, resulting in a degenerate shape and in a null volume. Since their number is very small ($\sim 10^1$) compared to the total ($\sim 10^5$), we exclude them from our analysis.

D. BAO wiggles

In addition to the baseline analysis of the correlation functions up to a given k_{max} , we will also consider information on the amplitude of the linear density field in the broadening of the BAO wiggles. While one can easily extend the full analysis of the correlation functions beyond $k_{\text{max}} \approx 0.1 \, h/\text{Mpc}$, we choose to focus on the question of how much information is in the wiggles only. This is

important, since it gives an estimate of how much additional signal on the amplitude α one may ideally recover from the displacement field on smaller scales. The likelihood in this case has the same form as for the power spectrum, except that the signal and the model contain the wiggly part of the galaxy power spectrum only,

$$\ln \mathcal{L}_w = -\frac{1}{2} \sum_{k=k_{\max}}^{\bar{k}} \frac{(P_g^w(k, \theta) - \hat{P}_g^w(k))^2}{\text{Var}_P(k)}. \quad (45)$$

In our implementation we choose $\bar{k} = 0.3 \text{ h/Mpc}$ since the wiggles are essentially completely damped beyond this scale. Note that the covariance matrix is the same as for the power spectrum. Clearly, the SNR on the amplitude α is suppressed by the amplitude of the wiggly part of the power spectrum which is smaller than $\mathcal{O}(10\%)$.

E. Theoretical error

In addition to the data covariance, it is often useful in perturbation-theory-based analysis to consider theoretical error covariance. Theoretical error takes into account realistic uncertainties in the perturbative modeling which serves two purposes. One is to prevent the overfitting by choosing aggressive data cuts. The other is to prevent biases in cosmological inference. Details of the formalism and applications to galaxy clustering and lensing can be found in [65,69–72].

The elements (ij) of the theoretical error covariance matrix have the following form:

$$(C_e)_{ij} = E_i E_j \rho_{ij}, \quad (46)$$

where the index i labels a bin in the data vector, which can be a k bin or a bispectrum or trispectrum configuration. Note that indices on the right-hand side are not summed over. The quantity E_i is the measure of the absolute error for the theoretical prediction in a given bin i and can be obtained from estimating the higher-order perturbation theory contributions neglected in the calculation. The matrix ρ_{ij} measures the correlation between the two different bins and it is crucial to ensure that one effectively marginalizes only over sufficiently smooth theoretical predictions [69]. For the power spectrum, it is chosen to be

$$\rho_{ij}^P = \exp \left[-\frac{1}{2} \frac{(k_i - k_j)^2}{\Delta k^2} \right], \quad (47)$$

where Δk is the parameters that determines the smoothness scale, over which the theoretical prediction vary very smoothly. In particular, we choose $\Delta k = 0.1 \text{ h/Mpc}$. If the two k bins are separated by more than Δk they are effectively uncorrelated and the theoretical prediction can vary within the theoretical uncertainty given by E_i and E_j . If the bins are within Δk , they are constrained to effectively

fluctuate coherently. In the case of the bispectrum, the correlation matrix is a bit more complicated and takes into account the shape of the triangle as follows:

$$\rho_{ij}^B = \exp \left[-\frac{1}{2} \frac{(k_{i,1} - k_{j,1})^2}{\Delta k^2} \right] \exp \left[-\frac{1}{2} \frac{(k_{i,2} - k_{j,2})^2}{\Delta k^2} \right] \times \exp \left[-\frac{1}{2} \frac{(k_{i,3} - k_{j,3})^2}{\Delta k^2} \right]. \quad (48)$$

In conclusion, the theoretical error covariance matrix has two parameters. One is the coherence length Δk and the other one is the estimate of the absolute theoretical error in perturbation theory. Following [69] we use the following estimates for the two-loop power spectrum:

$$E_i^{P,2\text{-loop}} = D^4(z) P_g(k_i, z) \left(\frac{k_i}{0.45 \text{ h/Mpc}} \right)^{3.3}, \quad (49)$$

and the one-loop bispectrum

$$E_i^{B,1\text{-loop}} = D^2(z) B_g(k_{i,1}, k_{i,2}, k_{i,3}, z) \times \left(\frac{(k_{i,1} + k_{i,2} + k_{i,3})/3}{0.3 \text{ h/Mpc}} \right)^{1.8}. \quad (50)$$

Note that requiring the theoretical error to be of the order of a given loop size is the other way to impose the perturbativity prior, where the total amplitude of fluctuations in the galaxy number density on large scales in always small.

Once the theoretical error covariance matrix is computed, it can be straightforwardly included in the analysis by computing the Gaussian likelihood with the total covariance

$$C_{ij} = C_{ij}^{\text{data}} + (C_e)_{ij}, \quad (51)$$

where C_{ij}^{data} is the usual data covariance that contains cosmic variance and the noise. Note that the new total covariance matrix is not diagonal and one has to generalize the formulas for the likelihood above accordingly. Nevertheless, the likelihood is still Gaussian.

F. Total likelihood

Assuming the Gaussian likelihood (with or without the theoretical error) for the power spectrum, bispectrum, trispectrum, and possibly BAO wiggles, we will combine them in our analysis in a simple way as follows:

$$\ln \mathcal{L} = \ln \mathcal{L}_P + \ln \mathcal{L}_B + \ln \mathcal{L}_T + \ln \mathcal{L}_w. \quad (52)$$

At first sight, this assumption may seem too optimistic, since we neglect all cross-covariance terms that may be relevant. In what follows we justify this choice.

We can start from the BAO wiggles. When we use the wiggly part of the power spectrum, we always use only scales with $k > k_{\max}$. This separation of scales justifies a simple combination of wiggles with the rest of the data vector.

For the correlation functions at $k < k_{\max}$, the simple form of the total likelihood in Eq. (52) is justified by arguments in perturbation theory. Since we are focusing on very large scales, the nonlinear contributions to the covariance and cross-covariance among different observables is small. In particular, these contributions are smaller than the effective error on cosmological parameters that comes from marginalization over many nuisance parameters [73]. This has been explicitly checked and verified for the power spectrum and bispectrum analyses on relevant scales (e.g., [73–77]). For example, the power spectrum analysis based on perturbation theory leads to identical results regardless of the choice of the covariance (Gaussian, nonlinear, or simulation based), in a volume similar to the one we use in this paper [28,66]. The same conclusion holds in the joint P + B analysis, as shown in many works that studied this combination in various setups (see, for example, [78–81] and references therein).

Finally, regarding the trispectrum, we expect that the similar arguments hold and that Eq. (52) is a very good approximation to the truth. Even if the analysis was optimistic, this is in line with our previous choices that maximize the possible information content of the trispectrum. One important point to make is that the trispectrum has also the cross-covariance with the power spectrum already at the Gaussian level. This can be, in principle, significant. However, in practice, this is not the case. In Appendix C 1 we show that less than 0.1% of the total number of quadrilaterals correlates with the power spectrum at leading order, making such contributions to the total signal negligible. Therefore, they can be safely neglected in the analysis.

G. Simple mode counting argument for the joint power spectrum and bispectrum analysis

Before we move on to the verification of our model and approximations in simulations and the main analysis, let us first give a simple estimate of the results that we expect.

The main advantage of using Eulerian perturbation theory is that the model is a simple polynomial in nuisance parameters and the amplitude of the linear density field. This means that the logarithm of the full likelihood in Eq. (52) takes particularly simple form and one can easily estimate the precision with which different parameters can be measured.

First of all, the highest signal-to-noise ratio is in the power spectrum. On large scales, the most dominant term is the linear theory contribution proportional to $b_1^2 \alpha^2$. Therefore, the galaxy power spectrum in real space measures the combination $b_1 \alpha$ very precisely. The relative error can be estimated as follows:

$$\frac{\sigma(b_1 \alpha)}{b_1 \alpha} = \frac{1}{\sqrt{2N_{\text{pix}}}}, \quad (53)$$

where the number of pixels, in practice, can be estimated as $N_{\text{pix}} \approx V k_{\max}^3 / (2\pi)^3$. Note that this error is different from the error on the amplitude of the power spectrum, since the amplitude of the power spectrum is proportional to α^2 . In practice, this is the most well-measured number in the real space analysis and we can essentially treat it as fixed. The degeneracy between the linear bias and α is broken at the level of the power spectrum only very mildly by the loop contributions. However, given several free higher-order bias parameters, the amplitude α is essentially unconstrained by the power spectrum alone.

Moving to higher-order statistics, assuming $b_1 \alpha$ is fixed, the amplitude of the tree-level bispectrum is proportional to a single power of α . While the bispectrum also contains quadratic biases as free parameters, the angular dependence allows for degeneracy with α to be partially broken. Therefore, we can assume that the error on the amplitude of the linear field from the joint power spectrum and bispectrum analysis is given by the bispectrum SNR. We can write

$$\frac{\sigma(\alpha)}{\alpha} \approx \text{few} \times \frac{1}{\Delta(k_{\max}) \sqrt{N_{\text{pix}}}}, \quad (54)$$

where we have used Eq. (2) and the numerical factor of a few depends on how much information is lost by marginalizing over b_2 and $b_{\mathcal{G}_2}$, which can be easily obtained using the Fisher matrix.

Using $V = (2 \text{ Gpc}/h)^3$ and $k_{\max} = 0.1 \text{ h}/\text{Mpc}$, we can estimate

$$\frac{\sigma(b_1 \alpha)}{b_1 \alpha} \approx 0.004 \quad \text{and} \quad \frac{\sigma(\alpha)}{\alpha} \approx \text{few} \times 0.02. \quad (55)$$

These numbers give an idea of how well the P + B analysis in real space should perform. In reality, the constraints are a bit weaker, due to the noise and nonlinearities we neglected in these arguments. However, we can already see that the baseline expectation for the error on α is roughly $\mathcal{O}(5\% - 10\%)$, which agrees with the field-level result of [35]. Furthermore, the simple estimates above give us two additional important predictions.

First, since there are no as-strong degeneracies in the bispectrum as the $b_1 - \alpha$ degeneracy in the power spectrum, the trispectrum is not expected to significantly improve the errors. This is due to much smaller SNR in the trispectrum as well as existence of new cubic bias parameters that have to be measured as well. We expect $\Delta^2(k_{\max}) \approx \mathcal{O}(10\%)$ improvements from inclusion of the trispectrum in the analysis. Second, in a P + B analysis we get a scaling of the error with $k_{\max}$ which is given by

$$\frac{\sigma(\alpha)}{\alpha} \sim \frac{k_{\max}^{-3/2}}{\Delta(k_{\max})}. \quad (56)$$

Note that the typical amplitude of fluctuations $\Delta(k_{\max})$ grows with $k_{\max}$, which implies that the error on α shrinks faster than the naive $k_{\max}^{-3/2}$ scaling. This makes sense, since the information comes from the bispectrum whose SNR grows as the field becomes more non-Gaussian on large scales, in addition to the simple mode counting. Note that the higher-order n -point functions would lead to different scaling with $\Delta(k_{\max})$. This is yet another way to estimate where the information comes from, by carefully measuring the dependence of the inferred errors on $k_{\max}$.

In what follows, we will show that all three predictions (the size of the error bars, importance of the trispectrum, and scaling with $k_{\max}$) are confirmed in a real analysis.

V. VALIDATION ON SIMULATIONS

The goal of this section is to validate our model and assumptions using large-volume numerical simulations. Our results are not new; this has been already done in a large body of work on the power spectrum and bispectrum and comparison to large-volume simulations with a similar setup [75,78,79,82–86]. In particular, we use the set of ten boxes of the PT challenge simulation suite [40] (we refer the reader to this reference for details on the simulations). With the total volume of 566 (Gpc/h)^3 at redshift $z = 0.6$, these simulations allow for very precise comparison of data and theory even on very large scales. We use the measured power spectrum and bispectrum of dark matter halos in real space, with number density and redshifts similar to the one analyzed in [35].

In particular, we want to test the following:

- (i) Eulerian perturbation theory predicts the power spectrum and the bispectrum on large scales well enough and leads to unbiased inference of α .
- (ii) Large displacements on large scales do not produce any effect beyond the one captured by Eulerian perturbation theory. In particular, one does not have to do the infrared resummation. The difference between the infrared-resummed and baseline models appears only on smaller scales, beyond $k_{\max}$ used in our analysis.
- (iii) Typical deviation of the best-fit model and the data for $k > k_{\max}$ are within the theoretical error estimate. This means that the realistic field of biased tracers has small fluctuations in agreement with perturbativity prior. Correspondingly, any combination of higher-order bias and nonlinear correction has to be small on the scales considered, even though all bias coefficients may be of order $\mathcal{O}(1)$.
- (iv) The error bars on the amplitude α are roughly given by the simple estimate in Eq. (54). The error bars do not change significantly when the theoretical error is included in the analysis, indicating that the higher bias and loop corrections are not very relevant on very large scales used in the analysis.

In order to do this, we will use the measurements from the full volume of 566 (Gpc/h)^3 , but fit it with the covariance matrix rescaled to match the box of 8 (Gpc/h)^3 which has the same volume as in [35]. We will also use the same two cuts, $k_{\max} = 0.1$ and $k_{\max} = 0.12 \text{ h/Mpc}$. We perform the joint power spectrum and bispectrum analysis using the theoretical model and likelihoods described above. We vary the following parameters:

$$\theta = \{\alpha, b_1, b_2, b_{\mathcal{G}_2}, b_{\Gamma_3}, c_s^2, c_1, c_2, c_3\}, \quad (57)$$

which include all the necessary bias parameters to compute the one-loop power spectrum and the tree-level bispectrum. It is important to emphasize that we also make an appropriate binning of the model, integrating the theoretical prediction in each bin with the appropriate weight to take into account the number of Fourier modes for each momentum k . This is important for the lowest k bins, but it does not significantly affect the final best fit and the error bars, since they are predominantly determined by the highest k modes in the analysis.

The main results are shown in Figs. 5–7 and summarized in Table I. The first two figures show the residual plots for the power spectrum and the bispectrum obtained from the joint analysis up to $k_{\max} = 0.1 \text{ h/Mpc}$. In all plots, the statistical error bars are shown for the full PT challenge volume, which is roughly 70 times larger than the volume of $V = 8 \text{ (Gpc/h)}^3$ that we are interested in. For this reason, all error bars are extremely small. For instance, the relative statistical errors in each k bin in the power spectrum are much smaller than 1%. This exquisite precision allows us to do the stringent test of the model.

Let us first focus on the power spectrum residuals shown in Fig. 5. The left panel shows the residuals together with the estimate of the two-loop theoretical error. We can see that the best fit reproduces the data well within the theoretical uncertainties of perturbation theory. The two points with statistically significant deviation (2σ and 3σ) are likely a random fluctuation, since they cannot be fitted by a sufficiently smooth curve. To verify this, we also fit the same data with the nonlinear PYBIRD code [87] and find almost identical results.⁸ Importantly, the higher k bins beyond $k_{\max} = 0.1 \text{ h/Mpc}$ used in the analysis, lay within the estimate of the two-loop theoretical error, indicating that the total galaxy field indeed follows expectations from perturbation theory. It is instructive to do the extrapolation of the best-fit model with $k_{\max} = 0.1 \text{ h/Mpc}$ to smaller scales. This is shown in the right panel of Fig. 5. The residuals remain within the theoretical error which grows very fast with k . More importantly, one can see that using the Eulerian model which is the baseline for our analysis is

⁸Equivalent approaches for large-scale structure correlators have been implemented in other publicly available codes, such as CLASS-PT [55], VELOCILEPTORS [88], and CLASS-ILOOP [89].

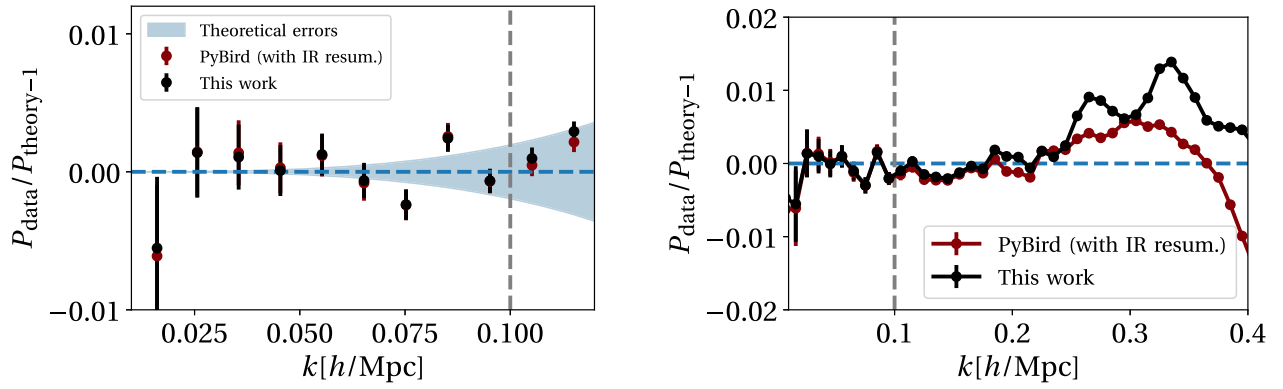


FIG. 5. Comparison between the data and the best-fit model in a joint power spectrum and bispectrum analysis up to $k_{\max} = 0.1 \, h/\text{Mpc}$ at $z = 0.6$ for the galaxy power spectrum. Left: the shaded region represents the two-loop theoretical error of Eq. (49). An independent fit obtained using the PYBIRD code which accounts for the IR resummation is also shown. Statistical error bars on data correspond to the full PT challenge volume.

insufficient on small scales to capture the BAO wiggles appropriately. However, the residual BAO wiggles appear only for $k \gtrsim 0.2 \, h/\text{Mpc}$, which is never used in our baseline analysis. This is an important confirmation that large displacements on large scales are captured well by Eulerian perturbation theory.

Similar residual plots for the bispectrum are shown in Fig. 6 for two different triangle configurations. One can see that the model fits the data well and that the residuals for $k > k_{\max}$ are within the estimate of the one-loop theoretical error.

It is important to emphasize once again that the statistical errors in these two figures are roughly 8 times smaller than the case of interest corresponding to the volume $V = 8 \, (\text{Gpc}/h)^3$. This implies that, for our main analysis, the theoretical model presented above is perfectly adequate to capture all relevant nonlinearities and it should lead to unbiased answers. This can be verified using the covariance matrix rescaled to the volume of the baseline analysis, still fitting the PT challenge data. The corresponding triangle plots for α , b_1 , $b_1\alpha$ in this analysis for two different choices of $k_{\max}$ are shown in Fig. 7. One can see that the inference of the amplitude and b_1 is unbiased, indicating that the simple model presented in Sec. III guarantees a sufficient

level of accuracy in describing the galaxy overdensity field on large scales and for the volume considered.

It is also interesting to look at the behavior of errors in Fig. 7 and Table I as a function of $k_{\max}$. We can see the improvement on the combination $b_1\alpha$ is smaller than the mode counting expectation. This is due to the fact that marginalization over higher-order biases effectively introduces an error that makes the constraints weaker. On the other hand, the errors on b_1 and α improve significantly with the increasing $k_{\max}$. This is in line with the expectation from the information content of the bispectrum, where relative errors improve faster than $k_{\max}^{-3/2}$ in agreement with Eq. (56). In conclusion, the joint P + B analysis of the PT challenge simulation boxes shows that the theoretical model presented in the previous section fits the data well and leads to errors in agreement with basic arguments based on perturbation theory.

It is interesting to ask how this picture changes when the theoretical error is included in the analysis. As we discussed, the theoretical error is an additional correlated covariance whose purpose is to mimic marginalizing over neglected higher-order corrections in the perturbation theory calculations. We repeat the same analysis of the PT challenge but with the two-loop power spectrum and one-loop bispectrum theoretical errors included in the covariance. The results are shown in Fig. 8. The triangle plots for α , b_1 , $b_1\alpha$ are shown with and without theoretical errors in the covariance, for two different values of $k_{\max}$. The inclusion of the theoretical error makes the constraints on α worse by 10% and 30% at $k_{\max} = 0.1$ and $k_{\max} = 0.12 \, h/\text{Mpc}$, respectively. Such results are expected. On larger scales, the impact of theoretical error is much smaller, given that the model is more precise. Nevertheless, the inclusion of the theoretical error does not change the error bars dramatically. In particular, they cannot explain the difference of roughly a factor of 3 between the result we find here and the SBI results for the

TABLE I. 1σ constraints and best-fit parameter for the PT challenge P + B analysis using the covariance with the volume of $V = (2 \, \text{Gpc}/h)^2$.

		P + B
$k_{\max} = 0.1 \, h/\text{Mpc}$		
α		0.960 ± 0.085
$b_1/b_{1,\text{fid}}$		1.051 ± 0.094
$k_{\max} = 0.12 \, h/\text{Mpc}$		
α		1.026 ± 0.052
$b_1/b_{1,\text{fid}}$		0.979 ± 0.051

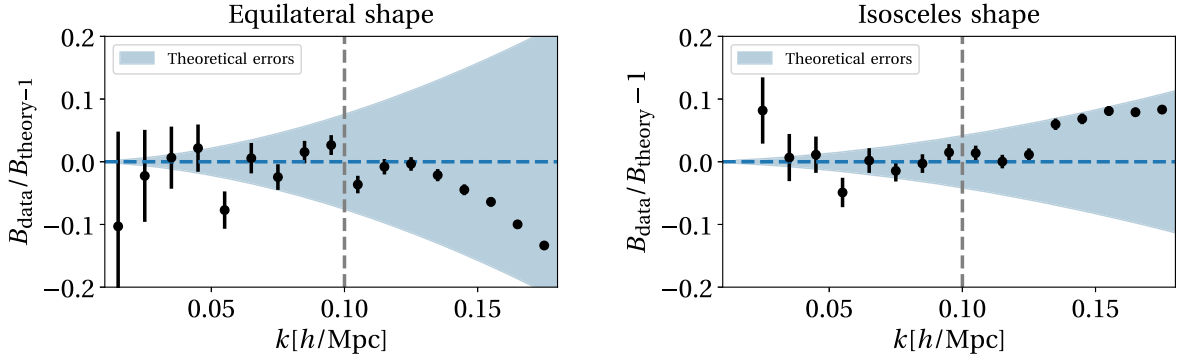


FIG. 6. Comparison between the data and the best-fit model in a joint power spectrum and bispectrum analysis up to $k_{\max} = 0.1 \text{ h/Mpc}$ at $z = 0.6$ for the equilateral and isosceles (i.e., $k_1 = 0.015 \text{ h/Mpc}$, $k_2 = k_3$) shapes of galaxy bispectrum. Statistical error bars on data correspond to the full PT challenge volume. The shaded region represents the one-loop theoretical error of Eq. (50).

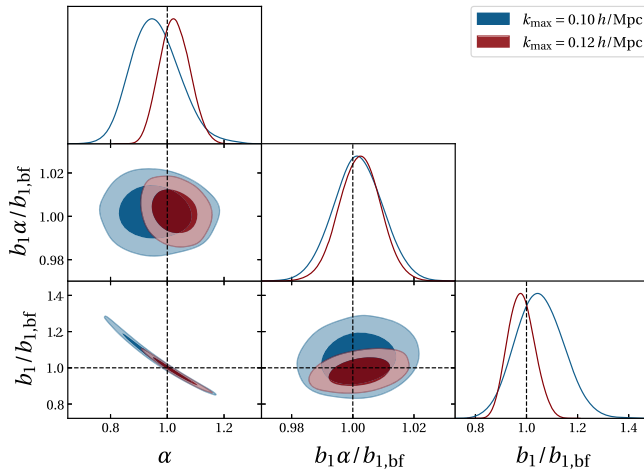


FIG. 7. P + B analysis of PT challenge data. The covariance used in the analysis is computed using the reduced volume $V = (2 \text{ Gpc}/h)^3$.

joint P + B analysis in [35]. We will come back to this point in Sec. VII.

VI. RESULTS

After extensive tests of the simple EPT model on large scales using large-volume PT challenge simulations and verifying that it is adequate for the joint P + B analysis, we are finally ready to present the main results. Our goals are to estimate the error bars of the conventional P + B analysis in a box of the volume $V = 8 \text{ (Gpc}/h)^3$ for tracers similar to the one studied in [35,36] and estimate the possible improvements coming from including the galaxy trispectrum in the analysis.

We will not analyze the real simulation data due to unavailability of the trispectrum that is an important part of our results. Instead, using the baseline model (for which we know to match the reality on large scales), we generate the mock data vector containing P, B, and T and the wiggly part

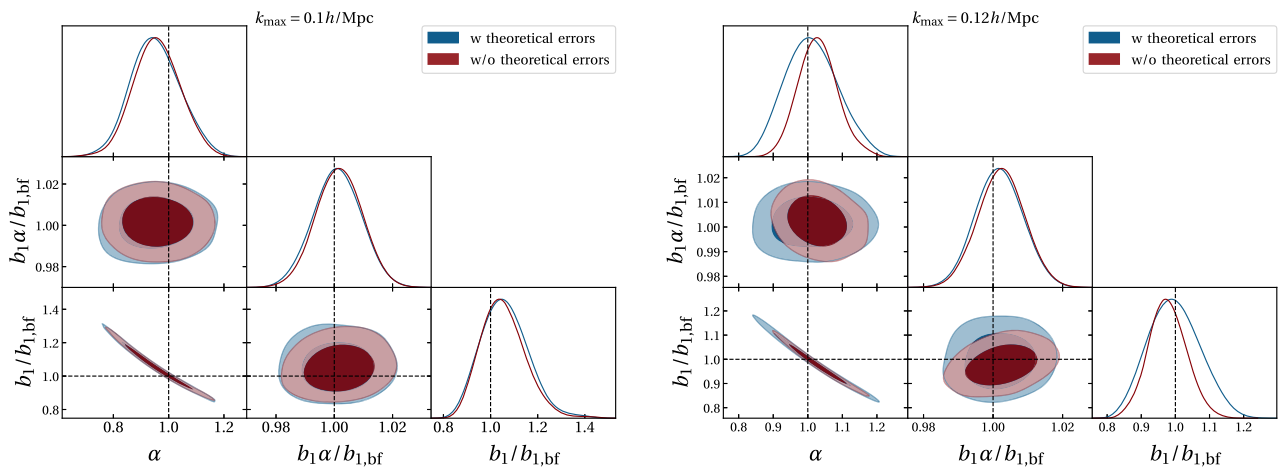


FIG. 8. P + B analysis of the PT challenge data with and without theoretical errors for (left) $k_{\max} = 0.1 \text{ h/Mpc}$ and (right) $k_{\max} = 0.12 \text{ h/Mpc}$. The covariance used in the analysis is computed using the reduced volume $V = (2 \text{ Gpc}/h)^3$.

of the power spectrum for $k > k_{\max}$ and perform the analysis of this data vector. In order to do that, for simplicity, we use the same baseline cosmology as the PT challenge which is very close to the standard Λ CDM universe. We also have to specify some fiducial values for the following set of parameters:

$$\theta = \{\alpha, b_1, b_2, b_{\mathcal{G}_2}, b_3, b_{\mathcal{G}_3}, b_{\mathcal{G}_2\delta}, b_{\Gamma_3}, c_s^2, c_w^2, c_1, c_2, c_3\}. \quad (58)$$

Obviously $\alpha_{\text{fid}} = 1$. For other parameters, we use two different sets of fiducial values depending on the tracer.

First, we chose the linear bias $b_{1,\text{fid}} = 1.7$. We also set the noise to be given by the number density of the dark matter halos in real space equal to $\bar{n} = 1.3 \times 10^{-3} \text{ (Mpc}/h)^3$ at fixed redshift $z = 0.5$. The linear bias, noise, and redshift are chosen to be similar to the SNG sample studied in [35]. The fiducial values of other parameters are far less important, and we verify that explicitly in our analysis. Since the sample is dense, the nonlinear bias parameters are small. We choose the following values: $b_{2,\text{fid}} = b_{\mathcal{G}_2,\text{fid}} = b_{\Gamma_3,\text{fid}} = b_3 = b_{\mathcal{G}_3,\text{fid}} = b_{\delta\mathcal{G}_2,\text{fid}} = 0.1$. The fiducial counterterms are given by $c_{s,\text{fid}}^2 = 1$, $c_{w,\text{fid}}^2 = 0$ and we choose that all fiducial noise-free parameters are equal to 1, $c_{1,\text{fid}} = c_{2,\text{fid}} = c_{3,\text{fid}} = 1$. When running the MCMC chains we use the following priors:

$$\begin{aligned} \mathcal{P}(\alpha) &= \mathcal{U}(0.5, 1.5), \\ \mathcal{P}(b_1) &= \mathcal{N}(1.0, 5.0), \\ \mathcal{P}(b_2) &= \mathcal{P}(b_{\mathcal{G}_2}) = \mathcal{P}(b_{\Gamma_3}) = \mathcal{P}(b_3) = \mathcal{P}(b_{\mathcal{G}_3}) \\ &= \mathcal{P}(b_{\delta\mathcal{G}_2}) = \mathcal{N}(0., 1.0), \\ \mathcal{P}(c_s^2) &= \mathcal{N}(0, 5.0), \\ \mathcal{P}(c_w^2) &= \mathcal{N}(0, 1.0), \\ \mathcal{P}(c_1) &= \mathcal{P}(c_2) = \mathcal{P}(c_3) = \mathcal{N}(1, 1). \end{aligned} \quad (59)$$

The variance of each parameter is chosen to match the choice in [35].

The second set of fiducial nuisance parameters is chosen to describe the sample at $z = 1$ used in the beyond two-point function challenge paper [36]. From [36], we estimate that $b_1 \approx 2.5$ and $\bar{n} \sim 4.5 \times 10^{-4} \text{ (h/Mpc)}^3$. For the fiducial values of the higher-order bias parameters, we use the measurements for a similar linear bias in [90,91]. This is not exact, but it provides a good estimate of the nonlinear biases which for less dense samples can be large. Labeling with a tilde the biases in [90,91], the rotation to our basis is

$$b_1 = \tilde{b}_1, \quad b_{\mathcal{G}_2} = \tilde{b}_{K^2} \quad b_2 = \tilde{b}_2 + \frac{4}{3}\tilde{b}_{K^2}, \quad (60)$$

for linear and quadratic bias and

$$\begin{aligned} b_{\Gamma_3} &= \tilde{b}_{\text{td}}, \quad b_{\mathcal{G}_3} = -\tilde{b}_{K^3}, \quad b_{\mathcal{G}_2\delta} = \tilde{b}_{\delta K^2} + \frac{1}{2}\tilde{b}_{K^3} - \frac{8}{21}\tilde{b}_{\text{td}}, \\ b_3 &= \tilde{b}_3 + 4\tilde{b}_{\delta K^2} + \frac{4}{3}\tilde{b}_{K^3} \end{aligned} \quad (61)$$

for cubic biases. Relying on the measurements in [90,91], we estimate

$$\begin{aligned} b_1 &\approx 2.5, \quad b_2 \approx 0.2, \quad b_{\mathcal{G}_2} \approx -0.6, \quad b_{\Gamma_3} \approx 1.5, \\ b_{\mathcal{G}_2\delta} &\approx 0.5, \quad b_{\mathcal{G}_3} \approx -1, \quad b_3 \approx -2. \end{aligned} \quad (62)$$

The other parameters are the same as before. In particular, $c_{s,\text{fid}}^2 = c_{1,\text{fid}} = c_{2,\text{fid}} = c_{3,\text{fid}} = 1$. We also use the same priors as in Eq. (59).

We checked that a different choice of fiducial values of the nonlinear bias parameters in both cases does not significantly impact our results, beyond small 10% corrections. More generally, all our results should be correct only up to 10%–20%, given several approximations that we use in the likelihood and the uncertainty in the fiducial parameters. This is acceptable, since we do not aim to have correct errors to better precision. We want to answer the question whether the error bars in FBI differ from P + B very significantly (by more than a factor of 2). Any mistake of the order of 10% is not very important for answering this question.

A. Joint power spectrum and bispectrum analysis

We begin by presenting the analysis performed on the mock data for a sample in [35] generated as described above, using the power spectrum and the bispectrum. The results are shown in Fig. 9 and Table II. There are three important points to make regarding the joint P + B analysis:

- (i) First, the error bars are in rough agreement with the mode counting estimate in Eq. (54). This is the confirmation that the main degeneracy $\alpha - b_1$ is broken by the bispectrum and that the remaining degeneracies of the amplitude α with the nonlinear biases b_2 and $b_{\mathcal{G}_2}$ are much milder (even though they degrade the errors by a factor of a few compared to the ideal case where all biases are known).
- (ii) Second, the error bars from a joint P + B analysis are much closer to the FBI results of [35] and roughly a factor of 3 smaller than the SBI errors in the same paper. This is a very interesting result and we will discuss it in more detail in the next section.
- (iii) Finally, the error bars follow the estimate for the scaling with $k_{\max}$ given in Eq. (56), providing another confirmation for our estimates. In particular, going from $k_{\max} = 0.1$ to $k_{\max} = 0.12 \text{ h/Mpc}$, we expect roughly a 30% improvement, in agreement with our analysis on the mock data.

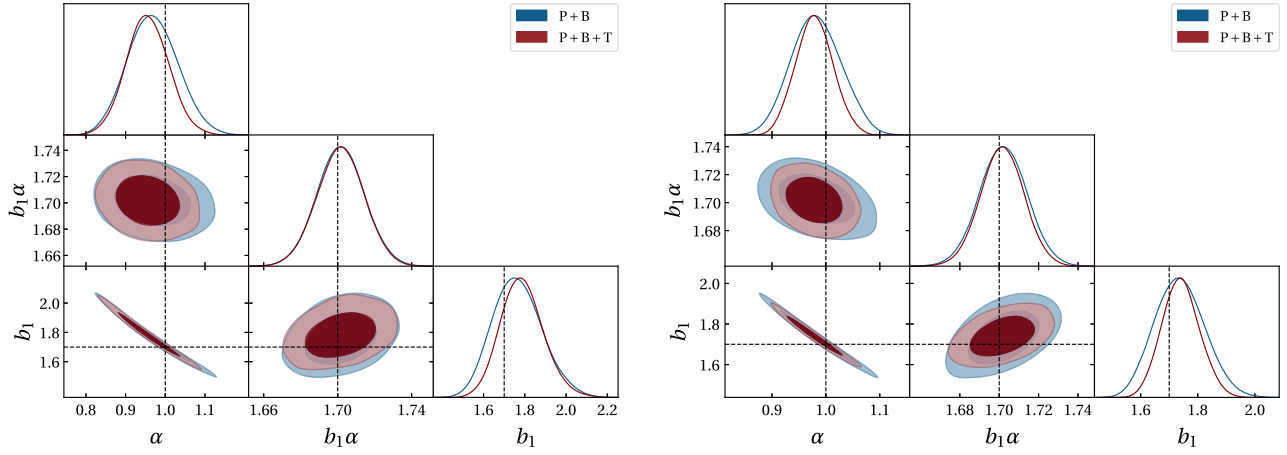


FIG. 9. Comparison between P + B analysis and P + B + T using mock data for $k_{\max} = 0.1 \text{ h/Mpc}$ (left) and $k_{\max} = 0.12 \text{ h/Mpc}$ (right).

Having obtained the error bars from the joint analysis of power spectrum and bispectrum, we proceed to investigate additional possible sources of information on top of this baseline analysis.

B. Adding the trispectrum

The three-level four-point function is the first higher-order statistic that can be consistently constructed when the galaxy overdensity field is evolved up to third order in perturbation theory. In fact, the trispectrum completes the set of correlation functions that can be reliably predicted with the cubic model. In this section we present the joint P + B + T analysis and investigate the impact of the trispectrum on constraining the amplitude α .

The results of the joint P + B + T analysis are shown in Fig. 9 and Table II. The two main findings are:

- (i) Adding the trispectrum to the P + B analysis reduces the error bars on the amplitude α by $\sim 20\%$ and $\sim 30\%$ at $k_{\max} = 0.1$ and $k_{\max} = 0.12 \text{ h/Mpc}$, respectively. This is an improvement in agreement with our SNR estimates. It indicates that there is no further major degeneracy breaking, which is expected given that there are several new free parameters in the trispectrum model. Nevertheless, even a

$\sim 20\%$ reduction of error bars is interesting and it is important to see if the impact of the trispectrum is the same in a more realistic setup (e.g., including the full nonlinear covariance or doing the analysis in redshift space).

- (ii) The error bars on α at redshift $z = 0.5$ are 0.054 and 0.035 for $k_{\max} = 0.10$ and $k_{\max} = 0.12 \text{ h/Mpc}$, respectively. This is in good agreement with the FBI analysis of [35] (given our uncertainty in the results of 10%–20%). Such result is not obvious, and it is rather nontrivial. The two approaches are using different perturbation theory models and measure α in two different ways. Therefore, assuming the absence of accidental cancellations, we interpret this agreement as a strong indication that the full complexity of the field-level analysis on large scales can be captured by the leading n -point functions computed in the simple Eulerian model, in agreement with expectations based on perturbation theory.

C. Adding the BAO wiggles

As explained in Sec. IV, the broadening of the BAO peak can help to measure the amplitude α . This can improve the power-spectrum-only analysis where α is measured only from the loop effects. The improvement is due to the fact that the damping factor Σ_{Λ}^2 is proportional to α only (and not to the combination $b_1\alpha$) such that the amplitude of fluctuations can be estimated from the broadening of the BAO peak. This effect is only due to the displacements and can be used to estimate how much information can be, in principle, recovered from the bulk flows that do not come with new free parameters.

As previously discussed, we want to consider two limiting situations. In the first case (that we indicate as “realistic”) Λ is chosen to include only large-scale displacements (i.e., $\Lambda = 0.2 \text{ h/Mpc}$ as in [55]) and one introduces a new free parameter c_w^2 which plays the role

TABLE II. 1σ constraints and best-fit parameter for P + B and P + B + T analyses using mock data. These errors are comparable to the FBI analysis of [35].

	P + B	P + B + T
$k_{\max} = 0.1 \text{ h/Mpc}$		
α	0.968 ± 0.064	0.966 ± 0.054
b_1	1.764 ± 0.118	1.767 ± 0.100
$k_{\max} = 0.12 \text{ h/Mpc}$		
α	0.981 ± 0.046	0.979 ± 0.035
b_1	1.736 ± 0.086	1.741 ± 0.065

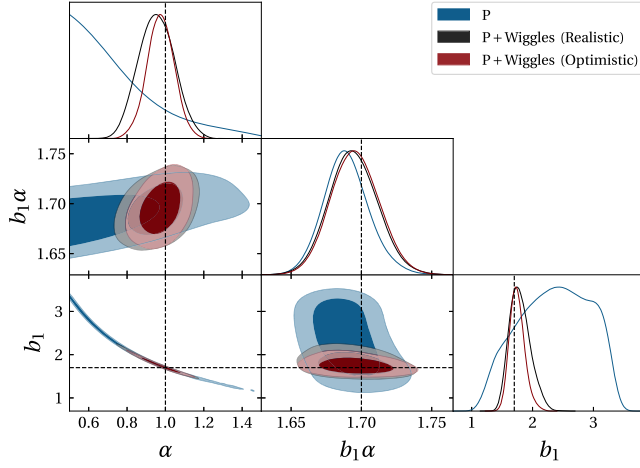


FIG. 10. Comparison between constraints obtained using only the broadband fitted until $k_{\max} = 0.1 \text{ h/Mpc}$ and when adding also wiggles (with $k_{\max,w} = 0.3 \text{ h/Mpc}$).

of a counterterm absorbing all possible higher-order contributions to the wiggly power spectrum. The second case instead represents the best-case scenario (“optimistic”) in which one sends $\Lambda \rightarrow +\infty$ and sets $c_w^2 = 0$. In this case, one assumes a perfect knowledge of all nonlinear effects induced by large bulk flows and the full broadening of the BAO peak can be described without new parameters.

In Fig. 10 and Table III we show the constraints on b_1 and α obtained for the power spectrum up to $k_{\max} = 0.1 \text{ h/Mpc}$ with and without adding the information from the wiggles with $k_{\max} < k \leq \bar{k} = 0.3 \text{ h/Mpc}$. We present the constraints for the realistic and optimistic scenarios. The main findings are as follows:

- (i) The information on the amplitude α in the wiggles is significantly higher than in the power spectrum at $k < k_{\max}$. Even though the wiggles have small amplitude ($\sim 10\%$ of the power spectrum), given the robustness of the BAO broadening, the constraints from the large displacements are stronger than the one-loop power spectrum.
- (ii) Even though adding the wiggles improves the results, they are still not as good as adding the bispectrum on large scales. In principle, it is not obvious that this had to be the case. We conclude that, in a ΛCDM -like cosmology, in order to get

TABLE III. 1σ constraints and best-fit parameter for broadband analysis of the power spectrum and when including also the wiggles in the realistic and optimistic scenarios.

	P	P + wiggles (realistic)	P + wiggles (optimistic)
α	0.743 ± 0.185	0.95 ± 0.099	0.975 ± 0.072
b_1	2.402 ± 0.535	1.801 ± 0.191	1.748 ± 0.124

optimal constraints, one has to include the higher-order n -point functions.

- (iii) The results do not change dramatically going from the optimistic to realistic scenario. This implies that the constraints from the BAO wiggles are rather robust. Note that this information is automatically included in the full-shape analyses of the power spectrum that use $k_{\max} \sim 0.3 \text{ h/Mpc}$.

D. Comparison to “beyond 2 pt challenge” results

We now apply the same methodology to a sample that resembles the one analyzed in [36]. Note that the simulation volume is the same $V = (2 \text{ Gpc}/h)^3$ but at slightly higher redshift $z = 1$. Using the exact same pipeline, we perform four different analyses. The results are shown in Figs. 11 and 12 and Table IV.

- (i) First, we perform a joint P + B analysis with $k_{\max,P} = 0.3$ and $k_{\max,B} = 0.15 \text{ h/Mpc}$. We do two versions of this analysis. In one of them, we fix b_{Γ_3}, c_2, c_3 to their fiducial values. This setup is exactly the same as the EFT P + B restricted analysis in [36] (see Fig. 22), with the same choice of prior. We recover very similar error bar on α of 3%. In the other version, we also vary b_{Γ_3}, c_2, c_3 within their usual priors. This corresponds to the baseline EFT P + B analysis in [36]. In this case, we find a $\sim 4\%$ error, which is smaller than what is reported in [36]. However, given that our results are uncertain at the $\mathcal{O}(20\%)$ level, we achieve a relatively good agreement with the EFT P + B analyses in [36]. This serves as a cross-check that our pipeline works well and that our choice of fiducial parameters is reasonable.

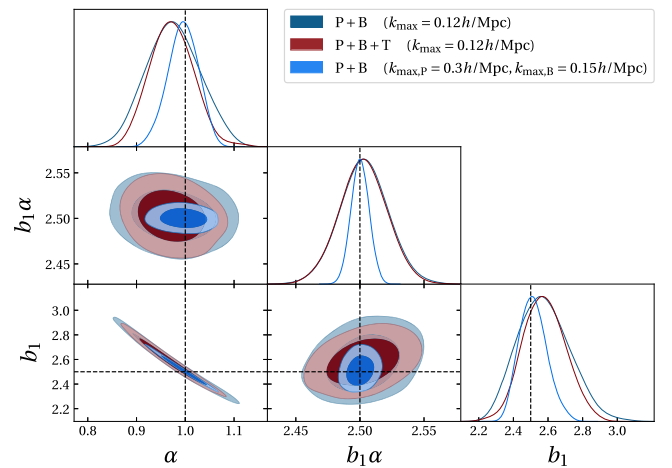


FIG. 11. P + B and P + B + T analyses of the sample of halos similar to [36]. P + B results for $k_{\max,P} = 0.3$ and $k_{\max,B} = 0.15 \text{ h/Mpc}$ fixing b_{Γ_3}, c_2, c_3 to their fiducial values are also shown.

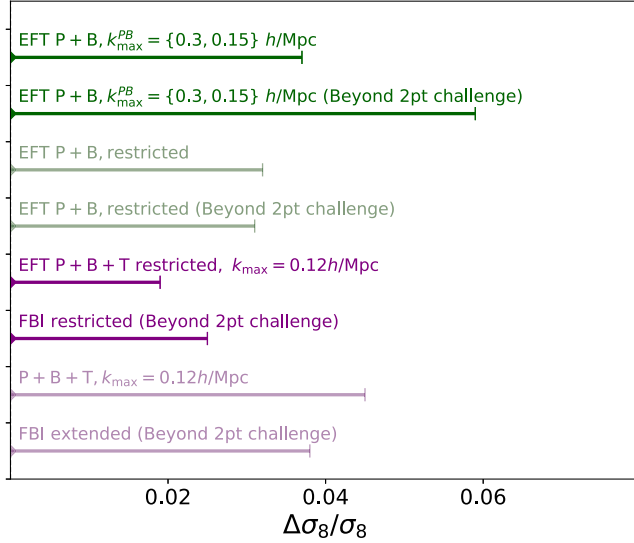


FIG. 12. Error bars on the amplitude of the density field obtained in different analyses described in the main text. While our error bars are different at $\mathcal{O}(20\%)$ level, the general trend for all examples is the same.

- (ii) Second, we perform the joint P + B analysis with $k_{\max} = 0.12 \text{ h/Mpc}$ using priors given in Eq. (59). This scale cut corresponds to $k_{\max} = 0.1 \text{ h/Mpc}$ of [36] where the cubic filter in k space is used. Since we use the spherical momentum shells, the two choices have a comparable number of Fourier modes (for more details, see Ref. [36]). The error on α in this case is 5.5%, approximately 30% larger than in the FBI “extended” analysis of [36]. This is similar to the results we found in other cases and indicates the broad agreement between the FBI and P + B analyses.
- (iii) Third, we add the trispectrum and perform the joint P + B + T analysis with the same cut $k_{\max} = 0.12 \text{ h/Mpc}$. Similar to what happens for other tracers, also here we find an improvement of roughly 20% compared to the P + B analysis. The joint P + B + T result is in rough agreement with the FBI error of 4% obtained in [36].
- (iv) Finally, we run a P + B + T analysis at $k_{\max} = 0.12 \text{ h/Mpc}$ but this time fixing all cubic biases and the bispectrum noise terms. In other words, the only free parameters remaining are the linear and quadratic biases, the dark matter counterterm, and the power spectrum shot noise. We choose this setup to

be as close as possible to the FBI restricted analysis in [36]. In this case, we find a significantly smaller error on α of 2%. This matches nicely the FBI restricted analysis in [36].

The last example is particularly interesting. If we compare the restricted P + B + T and P + B analyses at the same $k_{\max} = 0.12 \text{ h/Mpc}$, we find that the addition of the trispectrum when cubic biases are fixed makes the constraints better by more than a factor of 2. Note that the impact of fixing cubic biases in the P + B analysis is minimal. This is a very significant improvement. Does this mean that our intuition for the relevance of the trispectrum is wrong? We think not. It is still true that the trispectrum has much lower SNR than the power spectrum and the bispectrum. However, since the trispectrum in this particular setup comes with no new free parameters by design, even this small signal is very useful to break the residual degeneracy between α and b_2 and $b_{\mathcal{G}_2}$ in the bispectrum. This degeneracy breaking is what leads to much smaller error bar on α . Clearly, this example is somewhat artificial and designed to maximize the impact of n -point functions beyond the bispectrum. For a real tracer where the cubic and higher-order biases are not zero, such approach is inappropriate, as evident from the biased result of the FBI restricted analysis in [36]. Pushing this to the extreme, one can always design a setup or find a cosmological parameter such that the relevant information is only in the higher-order correlation functions. For example, if we wanted to measure the primordial non-Gaussianity parameter τ_{NL} , the FBI would be much better than the P + B analysis. However, these special situations are not what we have in mind when we discuss the comparison of the field-level inference and the standard analyses. In a generic setup, our arguments are still valid.

In order to summarize our forecasts and visually compare them to analyses in [36], we produce Fig. 12 which is supposed to be compared to Fig. 22 in [36]. While all our results are uncertain at the $\mathcal{O}(20\%)$ level, we find an overall good agreement with results of [36] and we recover all features and trends. These results are reassuring and provide additional evidence for the simple relation between the n -point functions and the field-level inference on large scales.

In conclusion, all our results are in agreement with expectations from perturbation theory. For dark matter halos in real space, where the only cosmological parameter is the amplitude of the linear field α , most of the

TABLE IV. 1σ constraints and best-fit parameters for P + B and P + B + T with $k_{\max} = 0.12 \text{ h/Mpc}$ (first and second columns) and P + B with $k_{\max, \text{P}} = 0.3$ and $k_{\max, \text{B}} = 0.15 \text{ h/Mpc}$ (third column).

	P + B	P + B + T	P + B ($k_{\max, \text{P}} = 0.3, k_{\max, \text{B}} = 0.15 \text{ h/Mpc}$)
α	0.975 ± 0.055	0.973 ± 0.045	0.995 ± 0.031
b_1	2.575 ± 0.151	2.578 ± 0.125	2.514 ± 0.0810

information comes from the bispectrum, except in some artificial examples where the higher-order correlation functions can dominate the SNR. Even though the power spectrum has the highest signal-to-noise ratio, due to the exact degeneracy between b_1 and α in linear theory, the amplitude can be measured from the one-loop power spectrum only with a modest precision of roughly 20% in a box of volume $V = 8 \text{ (Gpc/h)}^3$ and with $k_{\text{max}} = 0.12 \text{ h/Mpc}$. On the other hand, adding the bispectrum improves these errors to approximately 5%–6% (the exact number slightly depends on the sample). Addition of the trispectrum further reduces the error to roughly 4%–5%. These errors are in good agreement with the measurement of α using FBI [35,36].

For simplicity, in this section we focused on only one of the two samples studied in [35]. However, we also checked that, for the other sample at $z = 1$ and with even larger number density, we find very similar results as the FBI analysis and all our conclusions remain the same in that case as well.

VII. DISCUSSION

The results presented in the previous section indicate that the conventional analyses based on correlation functions and the FBI analysis lead to the same results in the perturbative regime, as expected based on theoretical arguments [38,39]. This seems to be in contradiction with the findings in [35] where the FBI leads to up to 5 times better error bars than the SBI P + B analysis. The new puzzle is why the SBI and the standard P + B analyses differ so much. In this section, we try to provide some possible explanations, even though the conclusive arguments are still lacking and more investigation is needed to fully clarify this issue.

The key to understanding the different results is to understand the differences in the settings of the SBI analysis of [35] and the settings in this paper. There are three major differences. One is the theoretical model used. SBI analysis of [35] is based on LPT, while we use the Eulerian theory. As we argued in detail in this work and as has been checked in the literature before, the two models are practically indistinguishable on the relevant scales and we do not think that this can be the source of the disagreement. The second major difference is the likelihood used in the analyses. SBI does not assume any likelihood (even though it has been shown that a Gaussian likelihood is an excellent approximation [28]) and uses a large number of realizations to estimate the relevant scatter in the bins for the data vector. On the other hand, we use a simple Gaussian likelihood with the Gaussian covariance, neglecting the cross-covariance between P, B, and T. We have also justified this choice in the paper and there is numerous evidence in the literature that this is a good approximation. Finally, the third major difference is the choice of bias parameters in the analyses

and the treatment of priors on these nuisance parameters. We believe that this can be the main source of disagreement.

In this work, we follow the usual logic used in perturbation theory and analyses based on the n -point functions. First, given the statistical error bars of the data, we estimate to what perturbative order we have to compute observables in order to be unbiased. In our examples, such estimate indicates that the one-loop power spectrum and the tree-level bispectrum are sufficient. Once the order in PT is fixed, the bias parameters that have to be used in the analysis are also determined. It is important to emphasize that here we do not truncate the bias expansion; we are only truncating the loop expansion of the correlation functions, given the total SNR that we expect for a given dataset. For example, in the case where we compute power spectrum up to one loop and the bispectrum at tree level, even though we use the full cubic bias model, only one of the cubic biases appears explicitly in the one-loop power spectrum (b_{r_3}). The other cubic biases either give zero contribution at this order in PT or just renormalize b_1 . The complexity of the nonlinear field at this order in PT and for the simplest observables such as P and B is just not sufficient for other bias parameters to appear explicitly. Obviously, cubic biases would appear if we had gone to higher order in the loop expansion or higher n -point functions. However, since we estimate that they are not necessary to fit the data, we can safely drop these refinements of the basic perturbative prediction. To conclude, even though our model is based on the full bias expansion up to the cubic order, 3 out of 4 cubic bias parameters do not appear in the analysis. This is a consequence of bias renormalization and the loop expansion and not restricting the number of bias parameters that are used in the analysis.

On the other hand, the practical implementation of SBI in [35] uses the nonlinear field computed at cubic order with all relevant bias parameters to predict P and B. This is similar in spirit to the field-level inference, where one can only truncate the forward model to some order in the bias expansion, without the possibility to explicitly restrict the likelihood to certain correlation functions or number of loops. Therefore, the estimated power spectrum and bispectrum contain higher-order contributions beyond what we use in our model. For example, the power spectrum has a part of the two-loop terms, such as various P_{33} contributions. The bispectrum has a part of the one-loop contributions (diagrams usually labeled as B_{123} and B_{222}) and even some two-loop contributions such as B_{233} . It is interesting to explore whether this difference leads to a large difference in the error bars in the two P + B analyses.

Naively, one may think that this difference is irrelevant, since the higher-order terms are small in PT. However, this is not obvious. While the variance of the dark matter field is small for k_{max} used in the analysis, the fact that there are many contributions from the bias expansion means that

there are combinations of bias parameters that might produce a large effect. With all bias parameters uncorrelated and of order 1, such combinations are naturally explored in the MCMC chains. Marginalizing over such parameters can then make the error bars on α significantly higher than the standard expectation. We do not have clear evidence that this is the main origin of the disagreement, but it is certainly an effect to be carefully taken into account in any comparison.

In order to illustrate this point better, let us imagine that we keep going to higher and higher orders in PT hoping to improve the precision of our inference. If we keep $k_{\max}$ fixed, this means that we fit the same data with a theory that has a rapidly increasing number of free parameters. Without careful choice of the priors, even though the individual loop diagrams may be small, summing a large number of such terms would inevitably lead to large effects. In such case the total loop contribution stops being small and the perturbativity in this sense is lost. The details of when this happens for a given $k_{\max}$ and the given observable depend on the sample and the priors, but one always has to check if this is the case. The hints of this behavior were already noticed in the analyses of the one-loop bispectrum [37,92,93]. On the other hand, for the consistency of PT, adding higher-order terms should, at worst, leave the constraints unchanged; it should never make them worse.

What are the usual ways to mitigate this problem? One possible procedure is to add informative and physically motivated priors on bias parameters that make them correlated. These correlations can be estimated from some external analysis (see, for example, [94–98] for halo occupation distribution informed priors obtained from a simulation-based approach and from analytical predictions, respectively). Another possible solution is to use the so-called perturbativity priors [37], which we already mentioned in the Introduction. These theoretically motivated priors do not rely on any galaxy evolution physics and are imposed in order to control the size of the loop corrections predicted by PT (see Ref. [37] for more details). It is important to emphasize that these two procedures lead to similar results, indicating that the real tracers indeed never produce large fluctuations due to bias nonlinearities. Perturbativity priors are also one implicit assumption that we always make for perturbation theory to make sense. Note that requiring that fluctuations in the galaxy density field are small on large scales still allows that all biases are of order 1, but imposes correlations among them, similar to the ones found in halo occupation distribution or hydrodynamical simulations.

If one accepts the perturbativity prior, what are the consequences for the SBI analysis? In principle, one would have to restrict the sampling only to those values of biases that do not produce large loop corrections. In this case, the expectation is that the cubic terms in the bias expansion would be much less relevant and the SBI will be similar to the standard analysis of the tree-level bispectrum presented

in this paper. This intuition provides a simple way to test the relevance of cubic terms. For example, one can set all cubic parameters to zero in the SBI analysis and try to recover the standard results. The other easy check is to plot the one-loop bispectrum amplitudes as a function of scale for various shapes (similar to Fig. 6) for all samples in the chains in the SBI and see if the envelope they create coincides with the estimate of the theoretical error.

If we assume that the wide priors are the origin of the disagreement between the SBI and the standard analysis, one may wonder why the FBI still gives good results with presumably large fluctuations in the galaxy density field. The answer is that the FBI is always optimal if the model is close enough to the data. For the combination of biases that do not respect perturbativity, the model would contain a large fluctuation in the galaxy density field with the particular template given by PT. FBI compares such model with the whole field and assigns a small likelihood to such set of bias parameters. Therefore, there is no contradiction in having small errors in the FBI approach and large errors from $P+B$ if the nuisance parameters are degenerate enough in the $P+B$ combination.

In conclusion, one possible origin of disagreement can be that including higher-order loop contributions on very large scales where they are supposed to be negligible can lead to artificial increase in the variance of the inferred cosmological parameters. In order to avoid this, one can either do the lower-order calculation as we did in this paper or impose some perturbativity priors in the inference. If one does either of these, the results obtained using n -point functions at fixed loop order and the SBI analysis should lead to similar constraints. Luckily, this can be relatively easily tested and we leave this for future work.

VIII. CONCLUSIONS

In this paper we studied a simple setup of dark matter halos in real space on very large scales where the only free cosmological parameter is the amplitude of the linear density field α . We computed the errors on α in the joint analyses of $P+B$ and $P+B+T$. We confirmed that most of the information comes from the bispectrum that breaks the $b_1 - \alpha$ degeneracy and that the addition of the trispectrum improves the error only by 20%–30%. Our results for the $P+B$ joint analysis are in agreement with previous literature. This indicates that the information on large scales is saturated by a few leading-order n -point functions and dominated by the bispectrum (see also [99]).

These results are important in the context of comparison to alternative data analysis techniques, such as the field-level inference of cosmological parameters. Our work suggests that the error bars in the $P+B+T$ analysis are in good agreement with the recent FBI results. In the absence of some accidents and coincidences, this implies the clear connection between the FBI and the conventional analyses and confirms the theoretical expectation that the

field-level analysis in perturbative forward modeling is ultimately equivalent to measuring the n -point correlation functions computed at the same perturbative order. Of course, FBI is by construction optimal and, assuming everything else the same, it will always lead to slightly better errors than the correlation-function-based analyses. However, this improvement for perturbative forward modeling should be small and certainly not as significant as a factor of 2 or more.

One open problem that remains to be solved is the difference of the PT-based P + B analysis presented in this paper and the SBI P + B analysis of [35]. While we do not have any definitive conclusions, we gave some possible explanations for the origin of the discrepancy in the inferred error bars. It would be very important to resolve this problem since it touches upon very important questions of how to apply PT, in practice, and how to choose appropriate priors for nuisance parameters. All these questions will be progressively more important as we move on toward more complicated observables and higher-order computations in PT.

While throughout most of this paper we focused on the typical galaxy sample where the FBI is expected to be equivalent to the standard correlation-function-based analysis, it is important to emphasize that there are situations where the FBI is supposed to perform better. This includes the BAO reconstruction [26,29,100] or a very high number density tracer, such as neutral hydrogen [17,101,102]. In such examples, there are new large parameters, beyond the variance of the density field, that suppress the signal in the conventional analysis either by erasing the information (BAO) or making the covariance large (dense tracers). It would be interesting to repeat our analysis for these particular setups and show the difference compared to FBI more explicitly.

Apart from the comparison with the field-level inference, our work is also important since it provides a simple analytical model for the trispectrum of biased tracers in real space and its covariance matrix. This extends previous works that have used some compressed observables, such as the integrated trispectrum [67,68]. Our formalism can be used to get quick estimates of the error bars for the signals specific to the four-point function, such as models that produce parity violation (for example, [103]) or nonconventional shapes of primordial non-Gaussianities (for example, [104]). Furthermore, our likelihood can be used to explore if there is a compression of the trispectrum that preserves most of the information while significantly reducing the dimensionality of the data vector [105–109]. Such compression would significantly simplify realistic data analyses.

At the end, let us point out that all our results are derived for the real space dark matter halos and a single cosmological parameter α . While this is an important step to conceptually understand the relation between the field-level inference and conventional analyses based on the n -point

functions, it is unclear to what extent the lessons we learn transfer to the more realistic case of redshift space observables and other cosmological parameters. For instance, the $b_1 - \alpha$ degeneracy in redshift space is partially broken by the power spectrum quadrupole. For this reason, in the absence of strong observational bias effects [110], the results would be different than for the real space. For example, it is found, in practice, that the addition of the bispectrum in the analysis of the redshift space data does not significantly improve the error on α . Therefore, the contributions to the SNR from different n -point functions are different in redshift space and it would be interesting to repeat our analysis there. Similarly, other cosmological parameters have different degeneracies with biases and it would be interesting to explore to what extent our general arguments apply in that case as well.

ACKNOWLEDGMENTS

We thank Kazuyuki Akitsu, Giovanni Cabass, Michele Liguori, Massimo Pietroni, Alvise Raccanelli, and Beatriz Tucci for useful discussions. We also thank Nhat-Minh Nguyen, Fabian Schmidt, and Beatriz Tucci for many useful comments on the manuscript. We thank Takahiro Nishimichi for providing the data of the PT challenge simulations. M.M. acknowledges support from MUR Progetti di Ricerca di Rilevante Interesse Nazionale (PRIN) Bando 2022—Grant No. 20228RMX4A, funded by the European Union—Next Generation EU, Mission 4, Component 1, CUP C53D23000940006.

DATA AVAILABILITY

The data that support the findings of this article are not publicly available. The data are available from the authors upon reasonable request.

APPENDIX A: CROSS-CORRELATION BETWEEN EPT AND LPT

The cross-correlation between the leading-order EPT and LPT fields is given by

$$\begin{aligned} & \langle \delta_{\text{EPT}}^{\text{LO}}(\mathbf{k}_1) \delta_{\text{LPT}}^{\text{LO}}(\mathbf{k}_2) \rangle \\ &= \int d^3\mathbf{q}_1 d^3\mathbf{q}_2 \langle \delta_1(\mathbf{q}_1) e^{-i\mathbf{k}_2 \cdot \boldsymbol{\psi}_1(\mathbf{q}_2)} \rangle e^{-i\mathbf{k}_1 \cdot \mathbf{q}_1 - i\mathbf{k}_2 \cdot \mathbf{q}_2} \\ &= i \frac{d}{d\lambda} \int d^3\mathbf{q}_1 d^3\mathbf{q}_2 \langle e^{-i(\lambda\delta_1(\mathbf{q}_1) + \mathbf{k}_2 \cdot \boldsymbol{\psi}_1(\mathbf{q}_2))} \rangle e^{-i\mathbf{k}_1 \cdot \mathbf{q}_1 - i\mathbf{k}_2 \cdot \mathbf{q}_2} \Big|_{\lambda=0}. \end{aligned} \quad (\text{A1})$$

The expectation value can be computed using the cumulant theorem. For a random variable X , we have

$$\langle e^{-iX} \rangle = \exp \left(\sum_{n=0}^{\infty} \frac{(-i)^n}{n!} \langle X^n \rangle_c \right), \quad (\text{A2})$$

where the cumulants $\langle X^n \rangle_c$ in our case represent the connected n -point functions. Since at LO all fields in the exponent are Gaussian, the only nontrivial cumulant is the power spectrum. After taking the derivative and setting $\lambda = 0$ we get

$$\begin{aligned} i \frac{d}{d\lambda} \langle e^{-i(\lambda \delta_1(\mathbf{q}_1) + \mathbf{k}_2 \cdot \psi_1(\mathbf{q}_2))} \rangle \Big|_{\lambda=0} &= i \frac{d}{d\lambda} e^{-\frac{1}{2}(\lambda^2 \langle \delta_1^2(\mathbf{q}_1) \rangle + 2\lambda \mathbf{k}_2 \cdot \langle \delta_1(\mathbf{q}_1) \psi_1(\mathbf{q}_2) \rangle + k_2^i k_2^j \langle \psi_{1i}(\mathbf{q}_2) \psi_{1j}(\mathbf{q}_2) \rangle)} \Big|_{\lambda=0} \\ &= -i \mathbf{k}_2 \cdot \langle \delta_1(\mathbf{q}_1) \psi_1(\mathbf{q}_2) \rangle e^{-\frac{1}{2} k_2^2 \sigma_v^2}. \end{aligned} \quad (\text{A3})$$

Using this in Eq. (A1) gives the following exact result for the LO cross-correlation:

$$\langle \delta_{\text{EPT}}^{\text{LO}}(\mathbf{k}) \delta_{\text{LPT}}^{\text{LO}}(-\mathbf{k}) \rangle' = P_{\text{lin}}(k) e^{-\frac{1}{2} k^2 \sigma_v^2}. \quad (\text{A4})$$

Assuming that at LO all power spectra are approximately equal to the linear power spectrum on large scales, we get the following LO cross-correlation coefficient:

$$r^{\text{LO}}(k) \approx e^{-\frac{1}{2} k^2 \sigma_v^2}, \quad (\text{A5})$$

which is Eq. (7) in the main text.

The NLO cross-correlation can be derived in a similar way. It is clear from the generic structure of the LPT expressions that it must be exponentially suppressed on

small scales, but keeping the cubic terms in the Eulerian perturbation theory guaranties partial cancellation of large displacement effects on large scales.

APPENDIX B: TRISPECTRUM SHOT NOISE TERMS

Here we show all the possible contributions to the trispectrum shot noise. The total galaxy field can indeed be expressed as a sum of the perturbative solution and the stochastic contribution

$$\delta_g(\mathbf{x}) = \delta_g^{\text{PT}}(\mathbf{x}) + \epsilon(\mathbf{x}). \quad (\text{B1})$$

We expand the noise field using

$$\epsilon(\mathbf{x}) = \epsilon_0(\mathbf{x}) + \epsilon_1(\mathbf{x}) \delta(\mathbf{x}) + \frac{1}{2} \epsilon_2(\mathbf{x}) (\delta^2(\mathbf{x}) - \langle \delta^2 \rangle) + \epsilon_{\mathcal{G}_2}(\mathbf{x}) \mathcal{G}_2(\mathbf{x}) + \dots, \quad (\text{B2})$$

and, since it is well-known that the spectra of the noise in the Poissonian limit satisfy

$$B_\epsilon = P_\epsilon^2 = \frac{1}{\bar{n}_g^2}, \quad (\text{B3})$$

we define the correlation among different noise operators as

$$\langle \epsilon_i(\mathbf{x}) \epsilon_j(\mathbf{x}') \rangle \equiv \frac{c_{ij}}{\bar{n}_g} \delta^D(\mathbf{x} - \mathbf{x}'), \quad (\text{B4})$$

$$\langle \epsilon_i(\mathbf{x}_1) \epsilon_j(\mathbf{x}_2) \epsilon_k(\mathbf{x}_3) \rangle \equiv \frac{c_{ijk}}{\bar{n}_g^2} \delta^D(\mathbf{x}_1 - \mathbf{x}_2) \delta^D(\mathbf{x}_2 - \mathbf{x}_3), \quad (\text{B5})$$

$$\langle \epsilon_i(\mathbf{x}_1) \epsilon_j(\mathbf{x}_2) \epsilon_k(\mathbf{x}_3) \epsilon_l(\mathbf{x}_4) \rangle \equiv \frac{c_{ijkl}}{\bar{n}_g^3} \delta^D(\mathbf{x}_1 - \mathbf{x}_2) \delta^D(\mathbf{x}_2 - \mathbf{x}_3) \delta^D(\mathbf{x}_3 - \mathbf{x}_4). \quad (\text{B6})$$

As a warm-up, we compute the power spectrum and bispectrum stochastic terms: the power spectrum of Eq. (B1) is

$$P_{gg}(\mathbf{k}) = \langle \delta_g(\mathbf{k}) \delta_g(\mathbf{k}') \rangle' + \langle \epsilon(\mathbf{k}) \epsilon(\mathbf{k}') \rangle', \quad (\text{B7})$$

where we have exploited $\langle \delta \epsilon \rangle = 0$. The leading stochastic contribution for the galaxy power spectrum is then

$$\langle \epsilon_0(\mathbf{k}) \epsilon_0(\mathbf{k}') \rangle' = \frac{c_{00}}{\bar{n}_g}, \quad (\text{B8})$$

with c_{00} being defined in Eq. (B4). Notice that additional terms in the expansion could arise that will eventually renormalize the c_{00} coefficient, e.g., the term $\langle \epsilon_1(\mathbf{x}) \delta(\mathbf{x}) \epsilon_1(\mathbf{x}') \delta(\mathbf{x}') \rangle'$. For the bispectrum, the leading shot noise terms are

$$\langle \delta_g^{\text{PT}}(\mathbf{k}_1) \epsilon(\mathbf{k}_2) \epsilon(\mathbf{k}_3) \rangle' + 2 \text{ perms.} = \frac{c_{01}}{\bar{n}_g} b_1 (P(k_1) + P(k_2) + P(k_3)), \quad (\text{B9})$$

and

$$\langle \epsilon(\mathbf{k}_1) \epsilon(\mathbf{k}_2) \epsilon(\mathbf{k}_3) \rangle' = \frac{c_{000}}{\bar{n}_g^2}. \quad (\text{B10})$$

For the trispectrum, we can write all the possible terms in the perturbative expansion as

$$\begin{aligned} T_g(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) &= \langle \delta_g(\mathbf{k}_1) \delta_g(\mathbf{k}_2) \delta_g(\mathbf{k}_3) \delta_g(\mathbf{k}_4) \rangle' \\ &= \langle \delta_g^{\text{PT}}(\mathbf{k}_1) \delta_g^{\text{PT}}(\mathbf{k}_2) \delta_g^{\text{PT}}(\mathbf{k}_3) \delta_g^{\text{PT}}(\mathbf{k}_4) \rangle' + \langle \delta_g^{\text{PT}}(\mathbf{k}_1) \delta_g^{\text{PT}}(\mathbf{k}_2) \delta_g^{\text{PT}}(\mathbf{k}_3) \epsilon(\mathbf{k}_4) \rangle' + 3 \text{ perms.} \\ &\quad + \langle \delta_g^{\text{PT}}(\mathbf{k}_1) \delta_g^{\text{PT}}(\mathbf{k}_2) \epsilon(\mathbf{k}_3) \epsilon(\mathbf{k}_4) \rangle' + 5 \text{ perms.} + \langle \delta_g^{\text{PT}}(\mathbf{k}_1) \epsilon(\mathbf{k}_2) \epsilon(\mathbf{k}_3) \epsilon(\mathbf{k}_4) \rangle' + 3 \text{ perms.} \\ &\quad + \langle \epsilon(\mathbf{k}_1) \epsilon(\mathbf{k}_2) \epsilon(\mathbf{k}_3) \epsilon(\mathbf{k}_4) \rangle'. \end{aligned} \quad (\text{B11})$$

The first line of Eq. (B11) is simply the tree-level perturbative expression for the galaxy trispectrum. The second line is automatically zero, since $\langle \epsilon \rangle = 0$. The third line gives rise to a number of possible stochastic terms (up to one-loop order),

$$\begin{aligned} \langle \delta_g^{\text{PT}}(\mathbf{k}_1) \delta_g^{\text{PT}}(\mathbf{k}_2) \epsilon(\mathbf{k}_3) \epsilon(\mathbf{k}_4) \rangle' &= \langle b_1 \delta(\mathbf{k}_1) b_1 \delta(\mathbf{k}_2) \epsilon_0(\mathbf{k}_3) \epsilon_0(\mathbf{k}_4) \rangle' + \langle b_1 \delta(\mathbf{k}_1) b_1 \delta(\mathbf{k}_2) \epsilon_0(\mathbf{k}_3) \int_p \epsilon_1(\mathbf{p}) \delta(\mathbf{k}_4 - \mathbf{p}) \rangle' \\ &\quad + \langle b_1 \delta(\mathbf{k}_1) b_1 \delta(\mathbf{k}_2) \epsilon_0(\mathbf{k}_3) \int_p \epsilon_2(\mathbf{p}) \delta^2(\mathbf{k}_4 - \mathbf{p}) \rangle' + \langle b_1 \delta(\mathbf{k}_1) b_1 \delta(\mathbf{k}_2) \epsilon_0(\mathbf{k}_3) \int_p \epsilon_{\mathcal{G}_2}(\mathbf{p}) \mathcal{G}_2(\mathbf{k}_4 - \mathbf{p}) \rangle' \\ &\quad + \langle b_1 \delta(\mathbf{k}_1) b_1 \delta(\mathbf{k}_2) \int_p \epsilon_1(\mathbf{p}) \delta(\mathbf{k}_3 - \mathbf{p}) \int_p \epsilon_1(\mathbf{p}) \delta(\mathbf{k}_4 - \mathbf{p}) \rangle' + \langle b_1 \delta(\mathbf{k}_1) b_2 \delta^2(\mathbf{k}_2) \epsilon_0(\mathbf{k}_3) \epsilon_0(\mathbf{k}_4) \rangle' \\ &\quad + \langle b_1 \delta(\mathbf{k}_1) b_{\mathcal{G}_2} \mathcal{G}_2(\mathbf{k}_2) \epsilon_0(\mathbf{k}_3) \epsilon_0(\mathbf{k}_4) \rangle' + \langle b_1 \delta(\mathbf{k}_1) b_2 \delta^2(\mathbf{k}_2) \epsilon_0(\mathbf{k}_3) \int_p \epsilon_1(\mathbf{p}) \delta(\mathbf{k}_3 - \mathbf{p}) \rangle' \\ &\quad + \langle b_1 \delta(\mathbf{k}_1) b_{\mathcal{G}_2} \mathcal{G}_2(\mathbf{k}_2) \epsilon_0(\mathbf{k}_3) \int_p \epsilon_1(\mathbf{p}) \delta(\mathbf{k}_3 - \mathbf{p}) \rangle', \end{aligned} \quad (\text{B12})$$

that eventually give the $\propto \bar{n}_g^{-1}$ contribution

$$\langle \delta_g^{\text{PT}}(\mathbf{k}_1) \delta_g^{\text{PT}}(\mathbf{k}_2) \epsilon(\mathbf{k}_3) \epsilon(\mathbf{k}_4) \rangle' = \frac{c_{01}}{\bar{n}_g} B_{\text{mgg}}(\mathbf{k}_{12}, \mathbf{k}_3, \mathbf{k}_4) + \frac{2b_1^2}{\bar{n}_g} (c_{02} + c_{0\mathcal{G}_2} \sigma^2(\mathbf{k}_1, \mathbf{k}_2) + c_{11}) P(k_1) P(k_2). \quad (\text{B13})$$

The fourth line of Eq. (B11) produces contributions of order $\sim \bar{n}_g^2$,

$$\langle \delta_g^{\text{PT}}(\mathbf{k}_1) \epsilon(\mathbf{k}_2) \epsilon(\mathbf{k}_3) \epsilon(\mathbf{k}_4) \rangle' \simeq \frac{b_1 c_{001}}{\bar{n}_g^2} P(k_1) + 3 \text{ perms.} \quad (\text{B14})$$

at leading perturbative order. The last contribution is given by the noise trispectrum

$$\langle \epsilon(\mathbf{k}_1) \epsilon(\mathbf{k}_2) \epsilon(\mathbf{k}_3) \epsilon(\mathbf{k}_4) \rangle' = \frac{c_{0000}}{\bar{n}_g^3}. \quad (\text{B15})$$

APPENDIX C: ESTIMATORS AND COVARIANCE MATRICES

In this appendix, we derive the expression for the trispectrum covariance matrix. The trispectrum configurations are labeled by quadrilaterals Q defined by the four momenta and the two diagonals $Q = \{k_1, k_2, k_3, k_4, D_1, D_2\}$. In order to avoid double counting and take into account all possible quadrilaterals, we have to order the momenta $k_1 \geq k_2 \geq k_3 \geq k_4$, but we leave the diagonals free, subject to the constraint that the first diagonal is always the sum of the two largest momenta $D_1 + k_1 + k_2 = 0$ and the second diagonal is always the sum of second and third largest momenta $D_2 + k_2 + k_3 = 0$. The estimator for the connected trispectrum can be then written as

$$\hat{T}_c(k_1, k_2, k_3, k_4, D_1, D_2) = \frac{1}{VV_{1234}} \int_{\mathbf{q}_i \in k_i} \int_{\mathbf{p}_i \in D_i} \hat{\delta}_g(\mathbf{q}_1) \hat{\delta}_g(\mathbf{q}_2) \hat{\delta}_g(\mathbf{q}_3) \hat{\delta}_g(\mathbf{q}_4) \delta^D(\mathbf{q}_{1234}) \delta^D(\mathbf{q}_{12} + \mathbf{p}_1) \delta^D(\mathbf{q}_{23} + \mathbf{p}_2) - \text{disconnected terms}, \quad (\text{C1})$$

where the “disconnected terms” are chosen such that the estimator is unbiased $\langle \hat{T}_c \rangle = T_c$. Their explicit form is not of interest in the Gaussian likelihood, since they cancel in the covariance matrix.

Let us first compute the volume for the thin momentum shell containing all quadrilaterals with given momenta and diagonals. It is defined as

$$V_{1234} = \int_{\mathbf{p}_i \in D_i} d^3 \mathbf{p}_i \int_{\mathbf{q}_j \in k_j} d^3 \mathbf{q}_j \delta^D(\mathbf{q}_1 + \mathbf{q}_2 + \mathbf{q}_3 + \mathbf{q}_4) \delta^D(\mathbf{q}_1 + \mathbf{q}_2 + \mathbf{p}_1) \delta^D(\mathbf{q}_2 + \mathbf{q}_3 + \mathbf{p}_2). \quad (\text{C2})$$

Note that the symmetry among six edges of the tetrahedron is broken the moment these particular δ functions are chosen. What complicates the evaluation of the integral is that $\mathbf{q}_2$ appears three times as an argument of δ functions. When doing angular integrals, this will lead to Wigner 3j symbols. Therefore, it is more convenient to rewrite the previous formula as follows:

$$V_{1234} = \int_{\mathbf{p}_i \in D_i} d^3 \mathbf{p}_i \int_{\mathbf{q}_j \in k_j} d^3 \mathbf{q}_j \delta^D(\mathbf{q}_1 + \mathbf{q}_4 - \mathbf{p}_2) \delta^D(\mathbf{q}_1 + \mathbf{q}_2 + \mathbf{p}_1) \delta^D(\mathbf{q}_2 + \mathbf{q}_3 + \mathbf{p}_2), \quad (\text{C3})$$

which we can always do using the δ functions. This form is simpler, since any of the momenta enter the arguments of the δ functions at most twice. Next, we can write

$$V_{1234} = \int_{\mathbf{p}_i \in D_i} d^3 \mathbf{p}_i \int_{\mathbf{q}_j \in k_j} d^3 \mathbf{q}_j \int \frac{d^3 \mathbf{r}}{(2\pi)^3} \frac{d^3 \mathbf{x}}{(2\pi)^3} \frac{d^3 \mathbf{y}}{(2\pi)^3} e^{-i\mathbf{r} \cdot (\mathbf{q}_1 + \mathbf{q}_4 - \mathbf{p}_2) - i\mathbf{x} \cdot (\mathbf{q}_1 + \mathbf{q}_2 + \mathbf{p}_1) - i\mathbf{y} \cdot (\mathbf{q}_2 + \mathbf{q}_3 + \mathbf{p}_2)}. \quad (\text{C4})$$

Angular integration over $\mathbf{p}_1$, $\mathbf{q}_4$, and $\mathbf{q}_3$ can be done easily,

$$V_{1234} = \frac{1}{8\pi^6} \int_{\mathbf{p}_i \in D_i} p_i^2 dp_i \int_{\mathbf{q}_j \in k_j} q_j^2 dq_j \int_0^\infty r^2 dr \int_0^\infty x^2 dx \int_0^\infty y^2 dy j_0(p_1 x) j_0(q_4 r) j_0(q_3 y) \times \int d\Omega_{p_2} \int d\Omega_{q_1} \int d\Omega_{q_2} \int d\Omega_r \int d\Omega_x \int d\Omega_y e^{-i\mathbf{r} \cdot (\mathbf{q}_1 - \mathbf{p}_2) - i\mathbf{x} \cdot (\mathbf{q}_1 + \mathbf{q}_2) - i\mathbf{y} \cdot (\mathbf{q}_2 + \mathbf{p}_2)}. \quad (\text{C5})$$

In order to perform the other angular integrals, we will use the expansion of plane waves in spherical Bessel functions and spherical harmonics,

$$e^{\pm i\mathbf{r} \cdot \mathbf{q}} = 4\pi \sum_{\ell m} (\pm i)^\ell j_\ell(qr) Y_{\ell m}(\hat{\mathbf{r}}) Y_{\ell m}^*(\hat{\mathbf{q}}). \quad (\text{C6})$$

It follows that

$$\int d\Omega_{q_1} e^{-i\mathbf{r} \cdot \mathbf{q}_1 - i\mathbf{x} \cdot \mathbf{q}_1} = (4\pi)^2 \sum_{\ell_1 m_1} \sum_{\ell_2 m_2} (-i)^{\ell_1 + \ell_2} j_{\ell_1}(q_1 r) j_{\ell_2}(q_1 x) Y_{\ell_1 m_1}(\hat{\mathbf{r}}) Y_{\ell_2 m_2}^*(\hat{\mathbf{x}}) \int d\Omega_{q_1} Y_{\ell_1 m_1}^*(\hat{\mathbf{q}}_1) Y_{\ell_2 m_2}(\hat{\mathbf{q}}_1). \quad (\text{C7})$$

By orthonormality of the spherical harmonic, this integral leads to

$$\int d\Omega_{q_1} e^{-i\mathbf{r} \cdot \mathbf{q}_1 - i\mathbf{x} \cdot \mathbf{q}_1} = (4\pi)^2 \sum_{\ell m} (-1)^\ell j_\ell(q_1 r) j_\ell(q_1 x) Y_{\ell m}(\hat{\mathbf{r}}) Y_{\ell m}^*(\hat{\mathbf{x}}). \quad (\text{C8})$$

We can to the other angular integrals in the similar way. In particular,

$$\int d\Omega_{q_2} e^{-i\mathbf{x} \cdot \mathbf{q}_2 - i\mathbf{y} \cdot \mathbf{q}_2} = (4\pi)^2 \sum_{\ell m} (-1)^\ell j_\ell(q_2 x) j_\ell(q_2 y) Y_{\ell m}(\hat{\mathbf{x}}) Y_{\ell m}^*(\hat{\mathbf{y}}), \quad (\text{C9})$$

$$\int d\Omega_{p_2} e^{+i\mathbf{r} \cdot \mathbf{p}_2 - i\mathbf{y} \cdot \mathbf{p}_2} = (4\pi)^2 \sum_{\ell m} j_\ell(p_2 r) j_\ell(p_2 y) Y_{\ell m}(\hat{\mathbf{y}}) Y_{\ell m}^*(\hat{\mathbf{r}}). \quad (\text{C10})$$

Once all the angular integrals are done, the volume of the shell is a complicated radial integral of product of spherical Bessel functions,

$$V_{1234} = 8^3 \int_{p_i \in D_i} p_i^2 dp_i \int_{q_j \in k_j} q_j^2 dq_j \int_0^\infty r^2 dr \int_0^\infty x^2 dx \int_0^\infty y^2 dy \sum_{\ell} (2\ell + 1) \times j_0(p_1 x) j_0(q_4 r) j_0(q_3 y) j_{\ell}(q_1 r) j_{\ell}(q_1 x) j_{\ell}(q_2 x) j_{\ell}(q_2 y) j_{\ell}(p_2 r) j_{\ell}(p_2 y). \quad (C11)$$

Luckily, the radial integrals can be also done in a relatively straightforward way. First, we can use

$$\int_0^\infty r^2 dr j_0(q_4 r) j_{\ell}(q_1 r) j_{\ell}(p_2 r) = \frac{\pi}{4q_1 q_4 p_2} \mathcal{P}_{\ell} \left(\frac{q_1^2 + p_2^2 - q_4^2}{2q_1 p_2} \right), \quad (C12)$$

where $\mathcal{P}_{\ell}(x)$ are the Legendre polynomials. The volume of the shell then becomes

$$V_s = (2\pi)^3 \int_{p_i \in D_i} p_i^2 dp_i \int_{q_j \in k_j} q_j^2 dq_j \sum_{\ell} (2\ell + 1) \times \frac{1}{q_1 q_4 p_2} \mathcal{P}_{\ell} \left(\frac{q_1^2 + p_2^2 - q_4^2}{2q_1 p_2} \right) \frac{1}{q_1 q_2 p_1} \mathcal{P}_{\ell} \left(\frac{q_1^2 + q_2^2 - p_1^2}{2q_1 q_2} \right) \frac{1}{q_2 q_3 p_2} \mathcal{P}_{\ell} \left(\frac{q_2^2 + p_2^2 - q_3^2}{2q_2 p_2} \right). \quad (C13)$$

Finally, the sum over ℓ can be done using

$$\sum_{\ell} (2\ell + 1) \mathcal{P}_{\ell}(\mu_1) \mathcal{P}_{\ell}(\mu_2) \mathcal{P}_{\ell}(\mu_3) = \frac{2}{\pi} \frac{1}{\sqrt{1 + 2\mu_1 \mu_2 \mu_3 - \mu_1^2 - \mu_2^2 - \mu_3^2}}. \quad (C14)$$

This equation has a nice geometric interpretation. It is related to the volume of the tetrahedron with three edges q_1 , q_2 , and p_2 ,

$$V_{\text{tetra}} = \frac{q_1 q_2 p_2}{6} \sqrt{1 + 2\mu_1 \mu_2 \mu_3 - \mu_1^2 - \mu_2^2 - \mu_3^2}, \quad (C15)$$

where μ_i are cosines of plane angles between three edges meeting in a common vertex. In our formula for V_{1234} , we have exactly the same form after closer inspections of all angles. Therefore, we can finally write (assuming that all bins are Δk wide)

$$V_s = \frac{8}{3} \pi^2 \frac{k_1 k_2 k_3 k_4 D_1 D_2}{V_{\text{tetra}}} \Delta k^6. \quad (C16)$$

The volume V_{tetra} can be computed in terms of six edges and this expression is manifestly symmetric in $k_1, \dots, k_4$ and D_1, D_2 .

Now we are ready to compute the trispectrum covariance matrix as follows:

$$\begin{aligned} \text{Var}(k_1, k_2, k_3, k_4, D_1, D_2) &= \langle \hat{T}(k_1, k_2, k_3, k_4, D_1, D_2) \hat{T}(k_a, k_b, k_c, k_d, D_a, D_b) \rangle \\ &\quad - \langle \hat{T}(k_1, k_2, k_3, k_4, D_1, D_2) \rangle \langle \hat{T}(k_a, k_b, k_c, k_d, D_a, D_b) \rangle \\ &= \frac{(2\pi)^{12}}{V^2 V_{1234} V_{abcd}} P(k_1) P(k_2) P(k_3) P(k_4) \prod_{i=1,a}^{4,d} \prod_{j=1,a}^{2,b} \int_{\mathbf{q}_i \in \mathbf{k}_i} d^3 q_i \int_{\mathbf{p}_j \in \mathbf{D}_j} d^3 p_j \\ &\quad \times \delta_D(\mathbf{q}_{1234}) \delta_D(\mathbf{q}_{12} + \mathbf{p}_1) \delta_D(\mathbf{q}_{23} + \mathbf{p}_2) \delta_D(\mathbf{q}_{abcd}) \delta_D(\mathbf{q}_{ab} + \mathbf{p}_a) \delta_D(\mathbf{q}_{bc} + \mathbf{p}_b) \delta_D(\mathbf{q}_{1a}) \delta_D(\mathbf{q}_{2b}) \\ &\quad \times \delta_D(\mathbf{q}_{3c}) \delta_D(\mathbf{q}_{4d}) + 23 \text{ perm.} \\ &= \frac{(2\pi)^9}{V V_{1234} V_{abcd}} P(k_1) P(k_2) P(k_3) P(k_4) \delta_{\mathbf{k}_1, \mathbf{k}_a}^K \delta_{\mathbf{k}_2, \mathbf{k}_b}^K \delta_{\mathbf{k}_3, \mathbf{k}_c}^K \delta_{\mathbf{k}_4, \mathbf{k}_d}^K \delta_{\mathbf{p}_1, \mathbf{p}_a}^K \delta_{\mathbf{p}_2, \mathbf{p}_b}^K \prod_{i=a}^d \prod_{j=a}^b \int_{\mathbf{q}_i \in \mathbf{k}_i} d^3 q_i \\ &\quad \times \int_{\mathbf{p}_j \in \mathbf{D}_j} d^3 p_j \delta_D(\mathbf{q}_{abcd}) \delta_D(\mathbf{q}_{ab} + \mathbf{p}_a) \delta_D(\mathbf{q}_{bc} + \mathbf{p}_b) + 23 \text{ perm.} \\ &= \frac{(2\pi)^9}{V V_{1234}} P(k_1) P(k_2) P(k_3) P(k_4) \delta_{\mathbf{k}_1, \mathbf{k}_a}^K \delta_{\mathbf{k}_2, \mathbf{k}_b}^K \delta_{\mathbf{k}_3, \mathbf{k}_c}^K \delta_{\mathbf{k}_4, \mathbf{k}_d}^K \delta_{\mathbf{p}_1, \mathbf{p}_a}^K \delta_{\mathbf{p}_2, \mathbf{p}_b}^K + 23 \text{ perm.} \\ &= \frac{(2\pi)^9 R_{1234}}{V V_{1234}} P(k_1) P(k_2) P(k_3) P(k_4), \end{aligned} \quad (C17)$$

where V_{1234} is given by Eq. (C16).

1. PT cross terms

In the Gaussian approximation some PT cross terms survive. We have

$$\begin{aligned}
& \langle T(k_1, k_2, k_3, k_4, D_1, D_2) P(k) \rangle - \langle T(k_1, k_2, k_3, k_4, D_1, D_2) \rangle \langle P(k) \rangle \\
&= \frac{1}{V^2 V_{1234} V_P} \prod_{i=1}^4 \prod_{j=1,2} \prod_{n=a,b} \int_{q_i \in k_i} d^3 q_i \int_{p_i \in D_j} d^3 p_j \int_{q_n \in k_n} d^3 q_n \delta_D(\mathbf{q}_{1234}) \delta_D(\mathbf{q}_{ab}) \delta_D(\mathbf{q}_{12} + \mathbf{p}_1) \delta_D(\mathbf{q}_{23} + \mathbf{p}_2) \\
&\quad \times \langle (\delta(\mathbf{q}_1) \delta(\mathbf{q}_2) \delta(\mathbf{q}_3) \delta(\mathbf{q}_4) \delta(\mathbf{q}_a) \delta(\mathbf{q}_b)) \rangle_c \\
&= 2 \frac{(2\pi)^9}{V^2 V_{1234} V_P} \prod_{i=1}^4 \prod_{j=1,2} \prod_{n=a,b} \int_{q_i \in k_i} d^3 q_i \int_{p_i \in D_j} d^3 p_j \int_{q_n \in k_n} d^3 q_n \delta_D(\mathbf{q}_{1234}) \delta_D(\mathbf{q}_{ab}) \delta_D(\mathbf{q}_{12} + \mathbf{p}_1) \delta_D(\mathbf{q}_{23} + \mathbf{p}_2) \\
&\quad \times \delta_D(\mathbf{q}_{1a}) \delta_D(\mathbf{q}_{3b}) \delta_D(\mathbf{q}_{24}) P(q_1) P(q_3) P(q_2) \\
&= 2 \frac{(2\pi)^6}{V V_{1234} V_P} P^2(k) P(k_2) \delta_{\mathbf{k}, \mathbf{k}_1} \delta_{\mathbf{k}, \mathbf{k}_3} \delta_{\mathbf{k}_2, \mathbf{k}_4} \prod_{i=3,4} \prod_{j=1,2} \int_{q_i \in k_i} d^3 q_i \int_{p_i \in D_j} d^3 p_j \delta_D(-\mathbf{q}_3 - \mathbf{q}_4 + \mathbf{p}_1) \delta_D(\mathbf{q}_3 - \mathbf{q}_4 + \mathbf{p}_2). \quad (\text{C18})
\end{aligned}$$

All the other contractions will give $\mathbf{p}_{1,2} = 0$, which do not fall in any bin. Now we want to evaluate

$$\begin{aligned}
& \prod_{i=3,4} \prod_{j=1,2} \int_{q_i \in k_i} d^3 q_i \int_{p_i \in D_j} d^3 p_j \delta_D(-\mathbf{q}_3 - \mathbf{q}_4 + \mathbf{p}_1) \delta_D(\mathbf{q}_3 - \mathbf{q}_4 + \mathbf{p}_2) \\
&= \frac{1}{(2\pi^2)^2} \int d^3 q_3 \int d^3 q_4 \int dp_1 p_1^2 \int dp_2 p_2^2 \int d^3 x \int d^3 y e^{i\mathbf{q}_3(\mathbf{x}-\mathbf{y})} e^{-i\mathbf{q}_4(\mathbf{x}+\mathbf{y})} j_0(p_1 y) j_0(p_2 x), \quad (\text{C19})
\end{aligned}$$

where we performed angular integrations over $\mathbf{p}_1$ and $\mathbf{p}_2$. Using Eq. (C6) leads to

$$2^6 \sum_{\ell} (2\ell + 1) \int dq_3 dq_4 dp_1 dp_2 dx dy q_3^2 q_4^2 p_1^2 p_2^2 x^2 y^2 j_{\ell}(q_3 x) j_{\ell}(q_3 y) j_{\ell}(q_4 x) j_{\ell}(q_4 y) j_0(p_1 x) j_0(p_2 y), \quad (\text{C20})$$

and using Eq. (C12) we get

$$4\pi^2 \sum_{\ell} (2\ell + 1) \int p_1 p_2 dq_3 dq_4 dp_1 dp_2 \mathcal{P}_{\ell} \left(\frac{q_3^2 + q_4^2 - p_1^2}{2q_3 q_4} \right) \mathcal{P}_{\ell} \left(\frac{q_3^2 + q_4^2 - p_2^2}{2q_3 q_4} \right). \quad (\text{C21})$$

Now it is known that

$$\sum_{\ell} \frac{(2\ell + 1)}{2} \mathcal{P}_{\ell}(\mu_1) \mathcal{P}_{\ell}(\mu_2) = \delta_D(\mu_1 - \mu_2). \quad (\text{C22})$$

So Eq. (C21) reads

$$2(4\pi^2) \delta_{D_1, D_2} \int p_1^2 dp_1 \int dq_3 \int dq_4 = 8\pi^2 \delta_{D_1, D_2} D_1^2 \Delta k^3. \quad (\text{C23})$$

So PT terms are nonzero only for quadrilaterals which have $k_1 = k_3$, $k_2 = k_4$, $D_1 = D_2$; given that we are considering quadrilaterals with ordered sides, the only possible configurations are the equilateral shapes, i.e., $k_1 = k_2 = k_3 = k_4$ and $D_1 = D_2$. They are $\sim 0.1\%$ of the total number.

- [1] D. Baumann, A. Nicolis, L. Senatore, and M. Zaldarriaga, Cosmological non-linearities as an effective fluid, *J. Cosmol. Astropart. Phys.* **07** (2012) 051.
- [2] J. J. M. Carrasco, M. P. Hertzberg, and L. Senatore, The effective field theory of cosmological large scale structures, *J. High Energy Phys.* **09** (2012) 082.
- [3] V. Assassi, D. Baumann, D. Green, and M. Zaldarriaga, Renormalized halo bias, *J. Cosmol. Astropart. Phys.* **08** (2014) 056.
- [4] L. Senatore, Bias in the effective field theory of large scale structures, *J. Cosmol. Astropart. Phys.* **11** (2015) 007.
- [5] M. Mirbabayi, F. Schmidt, and M. Zaldarriaga, Biased tracers and time evolution, *J. Cosmol. Astropart. Phys.* **07** (2015) 030.
- [6] L. Senatore and M. Zaldarriaga, Redshift space distortions in the effective field theory of large scale structures, [arXiv:1409.1225](https://arxiv.org/abs/1409.1225).
- [7] M. Pietroni, G. Mangano, N. Saviano, and M. Viel, Coarse-grained cosmological perturbation theory, *J. Cosmol. Astropart. Phys.* **01** (2012) 019.
- [8] M. M. Ivanov, Effective field theory for large-scale structure, [arXiv:2212.08488](https://arxiv.org/abs/2212.08488).
- [9] G. Cabass, M. M. Ivanov, M. Lewandowski, M. Mirbabayi, and M. Simonović, Snowmass white paper: Effective field theories in cosmology, *Phys. Dark Universe* **40**, 101193 (2023).
- [10] T. Baldauf, E. Schaan, and M. Zaldarriaga, On the reach of perturbative descriptions for dark matter displacement fields, *J. Cosmol. Astropart. Phys.* **03** (2016) 017.
- [11] T. Baldauf, E. Schaan, and M. Zaldarriaga, On the reach of perturbative methods for dark matter density fields, *J. Cosmol. Astropart. Phys.* **03** (2016) 007.
- [12] M. Schmittfull, M. Simonović, V. Assassi, and M. Zaldarriaga, Modeling biased tracers at the field level, *Phys. Rev. D* **100**, 043514 (2019).
- [13] F. Schmidt, F. Elsner, J. Jasche, N. M. Nguyen, and G. Lavaux, A rigorous EFT-based forward model for large-scale structure, *J. Cosmol. Astropart. Phys.* **01** (2019) 042.
- [14] C. Modi, M. White, A. Slosar, and E. Castorina, Reconstructing large-scale structure with neutral hydrogen surveys, *J. Cosmol. Astropart. Phys.* **11** (2019) 023.
- [15] M. Schmittfull, M. Simonović, M. M. Ivanov, O. H. E. Philcox, and M. Zaldarriaga, Modeling galaxies in redshift space at the field level, *J. Cosmol. Astropart. Phys.* **05** (2021) 059.
- [16] F. Schmidt, An n -th order Lagrangian forward model for large-scale structure, *J. Cosmol. Astropart. Phys.* **04** (2021) 033.
- [17] A. Obuljen, M. Simonović, A. Schneider, and R. Feldmann, Modeling HI at the field level, *Phys. Rev. D* **108**, 083528 (2023).
- [18] J. Stadler, F. Schmidt, and M. Reinecke, Fast, accurate and perturbative forward modeling of galaxy clustering. Part I. Galaxies in the restframe, *J. Cosmol. Astropart. Phys.* **04** (2025) 089.
- [19] J. Stadler, F. Schmidt, M. Reinecke, and M. Esposito, Fast, accurate and perturbative forward modeling of galaxy clustering part II: Redshift space, *J. Cosmol. Astropart. Phys.* **11** (2025) 055.
- [20] G. Cabass and F. Schmidt, The EFT likelihood for large-scale structure, *J. Cosmol. Astropart. Phys.* **04** (2020) 042.
- [21] F. Elsner, F. Schmidt, J. Jasche, G. Lavaux, and N.-M. Nguyen, Cosmology inference from a biased density field using the EFT-based likelihood, *J. Cosmol. Astropart. Phys.* **01** (2020) 029.
- [22] F. Schmidt, G. Cabass, J. Jasche, and G. Lavaux, Unbiased cosmology inference from biased tracers using the EFT likelihood, *J. Cosmol. Astropart. Phys.* **11** (2020) 008.
- [23] G. Cabass and F. Schmidt, The likelihood for LSS: Stochasticity of bias coefficients at all orders, *J. Cosmol. Astropart. Phys.* **07** (2020) 051.
- [24] F. Schmidt, Sigma-eight at the percent level: The EFT likelihood in real space, *J. Cosmol. Astropart. Phys.* **04** (2021) 032.
- [25] A. Kostić, N.-M. Nguyen, F. Schmidt, and M. Reinecke, Consistency tests of field level inference with the EFT likelihood, *J. Cosmol. Astropart. Phys.* **07** (2023) 063.
- [26] I. Babić, F. Schmidt, and B. Tucci, BAO scale inference from biased tracers using the EFT likelihood, *J. Cosmol. Astropart. Phys.* **08** (2022) 007.
- [27] J. Stadler, F. Schmidt, and M. Reinecke, Cosmology inference at the field level from biased tracers in redshift-space, *J. Cosmol. Astropart. Phys.* **10** (2023) 069.
- [28] B. Tucci and F. Schmidt, EFTofLSS meets simulation-based inference: σ_8 from biased tracers, *J. Cosmol. Astropart. Phys.* **05** (2024) 063.
- [29] I. Babić, F. Schmidt, and B. Tucci, Straightening the ruler: Field-level inference of the BAO scale with LEFTfield, *J. Cosmol. Astropart. Phys.* **11** (2025) 066.
- [30] M. Peron, T. Nishimichi, M. Pietroni, and A. Taruya, Renormalized perturbation theory at Field-level: The LSS bootstrap in GridSPT, *J. Cosmol. Astropart. Phys.* **10** (2025) 098.
- [31] G. D'Amico, M. Marinucci, M. Pietroni, and F. Vernizzi, The large scale structure bootstrap: perturbation theory and bias expansion from symmetries, *J. Cosmol. Astropart. Phys.* **10** (2021) 069.
- [32] L. Amendola, M. Marinucci, M. Pietroni, and M. Quartin, Improving precision and accuracy in cosmology with model-independent spectrum and bispectrum, *J. Cosmol. Astropart. Phys.* **01** (2024) 001.
- [33] M. Marinucci, K. Pardede, and M. Pietroni, Bootstrapping Lagrangian perturbation theory for the large scale structure, *J. Cosmol. Astropart. Phys.* **10** (2024) 051.
- [34] A. Ansari, A. Banerjee, S. Jain, and S. Lalsodagar, Bootstrapping LSS perturbation theory beyond third order, *J. Cosmol. Astropart. Phys.* **11** (2025) 084.
- [35] N.-M. Nguyen, F. Schmidt, B. Tucci, M. Reinecke, and A. Kostić, How much information can be extracted from galaxy clustering at the field level?, *Phys. Rev. Lett.* **133**, 221006 (2024).
- [36] E. Krause *et al.* (Beyond-2 pt Collaboration), A parameter-masked mock data challenge for beyond-two-point galaxy clustering statistics, *Astrophys. J.* **990**, 99 (2025).
- [37] D. Braganca, Y. Donath, L. Senatore, and H. Zheng, Peeking into the next decade in large-scale structure cosmology with its effective field theory, *J. Cosmol. Astropart. Phys.* **11** (2025) 025.

- [38] G. Cabass, M. Simonović, and M. Zaldarriaga, Cosmological information in perturbative forward modeling, *Phys. Rev. D* **109**, 043526 (2024).
- [39] F. Schmidt, On the connection between field-level inference and n -point correlation functions, *J. Cosmol. Astropart. Phys.* **09** (2025) 056.
- [40] T. Nishimichi, G. D’Amico, M. M. Ivanov, L. Senatore, M. Simonović, M. Takada, M. Zaldarriaga, and P. Zhang, Blinded challenge for precision cosmology with large-scale structure: results from effective field theory for the redshift-space galaxy power spectrum, *Phys. Rev. D* **102**, 123541 (2020).
- [41] A. Taruya, T. Nishimichi, and D. Jeong, Grid-based calculation for perturbation theory of large-scale structure, *Phys. Rev. D* **98**, 103532 (2018).
- [42] M. Peloso and M. Pietroni, Galilean invariance and the consistency relation for the nonlinear squeezed bispectrum of large scale structure, *J. Cosmol. Astropart. Phys.* **05** (2013) 031.
- [43] A. Kehagias and A. Riotto, Symmetries and consistency relations in the large scale structure of the universe, *Nucl. Phys. B* **873**, 514 (2013).
- [44] P. Creminelli, J. Gleyzes, M. Simonović, and F. Vernizzi, Single-field consistency relations of large scale structure. Part II: Resummation and redshift space, *J. Cosmol. Astropart. Phys.* **02** (2014) 051.
- [45] P. Creminelli, J. Noreña, M. Simonović, and F. Vernizzi, Single-field consistency relations of large scale structure, *J. Cosmol. Astropart. Phys.* **12** (2013) 025.
- [46] L. Senatore and M. Zaldarriaga, The IR-resummed effective field theory of large scale structures, *J. Cosmol. Astropart. Phys.* **02** (2015) 013.
- [47] M. Mirbabayi, M. Simonović, and M. Zaldarriaga, Baryon acoustic peak and the squeezed limit bispectrum, *arXiv:1412.3796*.
- [48] T. Baldauf, M. Mirbabayi, M. Simonović, and M. Zaldarriaga, Equivalence principle and the baryon acoustic peak, *Phys. Rev. D* **92**, 043514 (2015).
- [49] M. Tegmark and P. J. E. Peebles, The time evolution of bias, *Astrophys. J. Lett.* **500**, L79 (1998).
- [50] K. C. Chan, R. Scoccimarro, and R. K. Sheth, Gravity and large-scale non-local bias, *Phys. Rev. D* **85**, 083509 (2012).
- [51] Y. Donath, M. Lewandowski, and L. Senatore, Direct signatures of the formation time of galaxies, *Phys. Rev. D* **109**, 123510 (2024).
- [52] V. Desjacques, D. Jeong, and F. Schmidt, Large-scale galaxy bias, *Phys. Rep.* **733**, 1 (2018).
- [53] R. Angulo, M. Fasiello, L. Senatore, and Z. Vlah, On the statistics of biased tracers in the effective field theory of large scale structures, *J. Cosmol. Astropart. Phys.* **09** (2015) 029.
- [54] F. Bernardeau, S. Colombi, E. Gaztanaga, and R. Scoccimarro, Large scale structure of the universe and cosmological perturbation theory, *Phys. Rep.* **367**, 1 (2002).
- [55] A. Chudaykin, M. M. Ivanov, O. H. E. Philcox, and M. Simonović, Nonlinear perturbation theory extension of the Boltzmann code CLASS, *Phys. Rev. D* **102**, 063533 (2020).
- [56] L. Senatore and G. Trevisan, On the IR-resummation in the EFTofLSS, *J. Cosmol. Astropart. Phys.* **05** (2018) 019.
- [57] D. Baumann, D. Green, and B. Wallisch, Searching for light relics with large-scale structure, *J. Cosmol. Astropart. Phys.* **08** (2018) 029.
- [58] J. Hamann, S. Hannestad, J. Lesgourgues, C. Rampf, and Y. Y. Y. Wong, Cosmological parameters from large scale structure—geometric versus shape information, *J. Cosmol. Astropart. Phys.* **07** (2010) 022.
- [59] Z. Vlah, U. Seljak, M. Y. Chu, and Y. Feng, Perturbation theory, effective field theory, and oscillations in the power spectrum, *J. Cosmol. Astropart. Phys.* **03** (2016) 057.
- [60] D. Blas, M. Garny, M. M. Ivanov, and S. Sibiryakov, Time-sliced perturbation theory II: Baryon acoustic oscillations and infrared resummation, *J. Cosmol. Astropart. Phys.* **07** (2016) 028.
- [61] O. H. E. Philcox, M. M. Ivanov, M. Simonović, and M. Zaldarriaga, Combining full-shape and BAO analyses of galaxy power spectra: A 1.6% CMB-independent constraint on H_0 , *J. Cosmol. Astropart. Phys.* **05** (2016) 032.
- [62] P. Zhang, G. D’Amico, L. Senatore, C. Zhao, and Y. Cai, BOSS correlation function analysis from the effective field theory of large-scale structure, *J. Cosmol. Astropart. Phys.* **02** (2022) 036.
- [63] S.-F. Chen, Z. Vlah, and M. White, A new analysis of galaxy 2-point functions in the BOSS survey, including full-shape information and post-reconstruction BAO, *J. Cosmol. Astropart. Phys.* **02** (2022) 008.
- [64] H. Gil-Marín, How to optimally combine pre-reconstruction full shape and post-reconstruction BAO signals, *J. Cosmol. Astropart. Phys.* **05** (2022) 040.
- [65] A. Chudaykin, M. M. Ivanov, and M. Simonović, Optimizing large-scale structure data analysis with the theoretical error likelihood, *Phys. Rev. D* **103**, 043525 (2021).
- [66] D. Wadekar, M. M. Ivanov, and R. Scoccimarro, Cosmological constraints from BOSS with analytic covariance matrices, *Phys. Rev. D* **102**, 123521 (2020).
- [67] D. Gualdi, S. Novell-Masot, H. Gil-Marín, and L. Verde, Matter trispectrum: Theoretical modelling and comparison to N-body simulations, *J. Cosmol. Astropart. Phys.* **01** (2021) 015.
- [68] D. Gualdi and L. Verde, Integrated trispectrum detection from BOSS DR12 NGC CMASS, *J. Cosmol. Astropart. Phys.* **09** (2022) 050.
- [69] T. Baldauf, M. Mirbabayi, M. Simonović, and M. Zaldarriaga, LSS constraints with controlled theoretical uncertainties, *arXiv:1602.00674*.
- [70] M. G. Moreira, F. Andrade-Oliveira, X. Fang, H.-J. Huang, E. Krause, V. Miranda, R. Rosenfeld, and M. Simonović, Mitigating baryonic effects with a theoretical error covariance, *Mon. Not. R. Astron. Soc.* **507**, 5592 (2021).
- [71] A. Aires, N. Kokron, R. Rosenfeld, F. Andrade-Oliveira, and V. Miranda, Mitigation of nonlinear galaxy bias with a theoretical-error likelihood, *J. Cosmol. Astropart. Phys.* **04** (2025) 070.
- [72] A. Maraio, A. Hall, and A. Taylor, Mitigating baryon feedback bias in cosmic shear through a theoretical error covariance in the matter power spectrum, *Mon. Not. R. Astron. Soc.* **537**, 1749 (2025).

- [73] R. Takahashi, N. Yoshida, M. Takada, T. Matsubara, N. Sugiyama, I. Kayo, T. Nishimichi, S. Saito, and A. Taruya, Non-Gaussian error contribution to likelihood analysis of the matter power spectrum, *Astrophys. J.* **726**, 7 (2011).
- [74] D. Wadekar and R. Scoccimarro, Galaxy power spectrum multipoles covariance in perturbation theory, *Phys. Rev. D* **102**, 123517 (2020).
- [75] A. Oddo, E. Sefusatti, C. Porciani, P. Monaco, and A. G. Sánchez, Toward a robust inference method for the galaxy bispectrum: Likelihood function and model selection, *J. Cosmol. Astropart. Phys.* **03** (2020) 056.
- [76] S. Martin, P. Schneider, and P. Simon, The bispectrum covariance beyond Gaussianity: A log-normal approach, *Astron. Astrophys.* **540**, A9 (2012).
- [77] R. Takahashi, N. Yoshida, M. Takada, T. Matsubara, N. Sugiyama, I. Kayo, A. J. Nishizawa, T. Nishimichi, S. Saito, and A. Taruya, Simulations of baryon acoustic oscillations II: Covariance matrix of the matter power spectrum, *Astrophys. J.* **700**, 479 (2009).
- [78] E. Sefusatti, M. Crocce, S. Pueblas, and R. Scoccimarro, Cosmology and the bispectrum, *Phys. Rev. D* **74**, 023522 (2006).
- [79] M. M. Ivanov, O. H. E. Philcox, T. Nishimichi, M. Simonović, M. Takada, and M. Zaldarriaga, Precision analysis of the redshift-space galaxy bispectrum, *Phys. Rev. D* **105**, 063512 (2022).
- [80] K. C. Chan and L. Blot, Assessment of the information content of the power spectrum and bispectrum, *Phys. Rev. D* **96**, 023528 (2017).
- [81] S. Novell-Masot, H. Gil-Marín, and L. Verde, On approximations of the redshift-space bispectrum and power spectrum multipoles covariance matrix, *J. Cosmol. Astropart. Phys.* **06** (2024) 048.
- [82] E. Sefusatti, M. Crocce, and V. Desjacques, The matter bispectrum in N-body simulations with non-Gaussian initial conditions, *Mon. Not. R. Astron. Soc.* **406**, 1014 (2010).
- [83] A. Moradinezhad Dizgah, M. Biagetti, E. Sefusatti, V. Desjacques, and J. Noreña, Primordial non-Gaussianity from biased tracers: Likelihood analysis of real-space power spectrum and bispectrum, *J. Cosmol. Astropart. Phys.* **05** (2021) 015.
- [84] D. Alkhanishvili, C. Porciani, E. Sefusatti, M. Biagetti, A. Lazanu, A. Oddo, and V. Yankelevich, The reach of next-to-leading-order perturbation theory for the matter bispectrum, *Mon. Not. R. Astron. Soc.* **512**, 4961 (2022).
- [85] A. Eggemeier, R. Scoccimarro, R. E. Smith, M. Crocce, A. Pezzotta, and A. G. Sánchez, Testing one-loop galaxy bias: Joint analysis of power spectrum and bispectrum, *Phys. Rev. D* **103**, 123550 (2021).
- [86] A. Oddo, F. Rizzo, E. Sefusatti, C. Porciani, and P. Monaco, Cosmological parameters from the likelihood analysis of the galaxy power spectrum and bispectrum in real space, *J. Cosmol. Astropart. Phys.* **11** (2021) 038.
- [87] G. D'Amico, L. Senatore, and P. Zhang, Limits on w CDM from the EFTofLSS with the PYBIRD code, *J. Cosmol. Astropart. Phys.* **01** (2021) 006.
- [88] S.-F. Chen, Z. Vlah, E. Castorina, and M. White, Redshift-space distortions in Lagrangian perturbation theory, *J. Cosmol. Astropart. Phys.* **03** (2021) 100.
- [89] D. Linde, A. Moradinezhad Dizgah, C. Radermacher, S. Casas, and J. Lesgourgues, CLASS-OneLoop: Accurate and unbiased inference from spectroscopic galaxy surveys, *J. Cosmol. Astropart. Phys.* **07** (2024) 068.
- [90] T. Lazeyras, C. Wagner, T. Baldauf, and F. Schmidt, Precision measurement of the local bias of dark matter halos, *J. Cosmol. Astropart. Phys.* **02** (2016) 018.
- [91] T. Lazeyras and F. Schmidt, Beyond LIMD bias: A measurement of the complete set of third-order halo bias parameters, *J. Cosmol. Astropart. Phys.* **09** (2018) 008.
- [92] O. H. E. Philcox, M. M. Ivanov, G. Cabass, M. Simonović, M. Zaldarriaga, and T. Nishimichi, Cosmology with the redshift-space galaxy bispectrum monopole at one-loop order, *Phys. Rev. D* **106**, 043530 (2022).
- [93] G. D'Amico, Y. Donath, M. Lewandowski, L. Senatore, and P. Zhang, The BOSS bispectrum analysis at one loop from the effective field theory of large-scale structure, *J. Cosmol. Astropart. Phys.* **05** (2024) 059.
- [94] K. Akitsu, Mapping the galaxy-halo connection to the galaxy bias: Implication to the HOD-informed prior, [arXiv:2410.08998](https://arxiv.org/abs/2410.08998).
- [95] M. M. Ivanov, C. Cuesta-Lazaro, S. Mishra-Sharma, A. Obuljen, and M. W. Toomey, Full-shape analysis with simulation-based priors: Constraints on single field inflation from BOSS, *Phys. Rev. D* **110**, 063538 (2024).
- [96] M. M. Ivanov, A. Obuljen, C. Cuesta-Lazaro, and M. W. Toomey, Full-shape analysis with simulation-based priors: Cosmological parameters and the structure growth anomaly, *Phys. Rev. D* **111**, 063548 (2025).
- [97] H. Zhang, M. Bonici, G. D'Amico, S. Paradiso, and W. J. Percival, HOD-informed prior for EFT-based full-shape analyses of LSS, *J. Cosmol. Astropart. Phys.* **04** (2025) 041.
- [98] M. M. Ivanov, Simulation-based priors without simulations: An analytic perspective on EFT parameters of galaxies, [arXiv:2503.07270](https://arxiv.org/abs/2503.07270).
- [99] K. Akitsu, M. Simonović, S.-F. Chen, G. Cabass, and M. Zaldarriaga, Cosmology inference with perturbative forward modeling at the field level: A comparison with joint power spectrum and bispectrum analyses, [arXiv:2509.09673](https://arxiv.org/abs/2509.09673).
- [100] I. Babić, F. Schmidt, and B. Tucci, Forward vs backward: Improving BAO constraints with field-level inference, [arXiv:2505.13588](https://arxiv.org/abs/2505.13588).
- [101] S. Foreman, A. Obuljen, and M. Simonović, Improving cosmological analyses of HI clustering by reducing stochastic noise, *Phys. Rev. D* **110**, 063555 (2024).
- [102] N. Kokron, Local primordial non-Gaussianity from 'zero-bias' 21 cm radiation during reionization, [arXiv:2504.20025](https://arxiv.org/abs/2504.20025).
- [103] G. Cabass, M. M. Ivanov, and O. H. E. Philcox, Colliders and ghosts: Constraining inflation with the parity-odd galaxy four-point function, *Phys. Rev. D* **107**, 023523 (2023).
- [104] P. Creminelli, A. Joyce, J. Khoury, and M. Simonovic, Consistency relations for the conformal mechanism, *J. Cosmol. Astropart. Phys.* **04** (2013) 020.
- [105] A. Heavens, R. Jimenez, and O. Lahav, Massive lossless data compression and multiple parameter estimation from galaxy spectra, *Mon. Not. R. Astron. Soc.* **317**, 965 (2000).

- [106] R. Scoccimarro, The bispectrum: From theory to observations, *Astrophys. J.* **544**, 597 (2000).
- [107] C. Reichardt, R. Jimenez, and A. Heavens, Recovering physical parameters from galaxy spectra using moped, *Mon. Not. R. Astron. Soc.* **327**, 849 (2001).
- [108] D. J. Eisenstein, H.-j. Seo, E. Sirko, and D. Spergel, Improving cosmological distance measurements by reconstruction of the baryon acoustic peak, *Astrophys. J.* **664**, 675 (2007).
- [109] O. H. E. Philcox, M. M. Ivanov, M. Zaldarriaga, M. Simonovic, and M. Schmittfull, Fewer mocks and less noise: Reducing the dimensionality of cosmological observables with subspace projections, *Phys. Rev. D* **103**, 043508 (2021).
- [110] V. Desjacques, D. Jeong, and F. Schmidt, The galaxy power spectrum and bispectrum in redshift space, *J. Cosmol. Astropart. Phys.* **12** (2018) 035.