

Optimal marked statistics from the effective field theory of large scale structure

Marco Marinucci^{1,2,3} Michele Liguori^{2,3} Federico Semenzato^{2,3} and Benjamin D. Wandelt^{4,5}

¹*Institute for Theoretical Physics, ETH Zurich, 8093 Zurich, Switzerland*

²*Dipartimento di Fisica Galileo Galilei, Università di Padova, I-35131 Padova, Italy*

³*INFN Sezione di Padova, I-35131 Padova, Italy*

⁴*Department of Physics and Astronomy, Johns Hopkins University, 3400 North Charles Street, Baltimore, Maryland 21218, USA*

⁵*Department of Applied Mathematics and Statistics, Johns Hopkins University, 3400 North Charles Street, Baltimore, Maryland 21218, USA*



(Received 10 July 2026; accepted 10 September 2026; published 6 October 2026)

We present a simple and robust framework to assess the information content of the power spectrum of a marked dark matter density field within the effective field theory of large scale structure, providing a physically motivated regime of validity of perturbation theory for such fields. We optimize the choice of mark to maximize the constraining power on cosmological parameters, focusing on the resulting reduction of parameter uncertainties. After marginalizing over the small-scale counterterms, we find improvements in the parameter error bars by a factor of approximately 1.30–2.3, 1.3–8.3, and 1.2–2.5 for Ω_m , σ_8 , and h , respectively, depending on redshift and smoothing scale. The optimized mark exhibits a simple functional form, for which we provide a physical interpretation in terms of its response to the underlying density field. The results presented in this work, which are valid for dark matter in real space, demonstrate the potential of marked statistics as a tool to extract additional cosmological information beyond standard two-point analyses and open a pathway for further exploration of nonlinear transformations of the density field.

DOI: [10.1103/3bqt-5fyk](https://doi.org/10.1103/3bqt-5fyk)

I. INTRODUCTION

The next generation of large scale structure (LSS) surveys, including Euclid [1,2], DESI [3], SPHEREx [4], the Vera Rubin Observatory [5] and the Nancy Grace Roman Space Telescope [6], will map the three-dimensional distribution of matter across an unprecedented volume of the observable Universe. A central challenge is to extract as much cosmological information as possible from the resulting datasets, particularly on mildly nonlinear scales where the signal is rich but the theoretical modeling is demanding. Doing so will substantially sharpen constraints on the full set of cosmological parameters, including the matter density Ω_m , the amplitude of fluctuations σ_8 , the Hubble constant H_0 , neutrino masses, and primordial non-Gaussianity.

The workhorse observable of LSS analyses is the matter power spectrum $P(k)$, the two-point function of the density contrast $\delta(\mathbf{x})$ in Fourier space. At linear order, this is a near-complete statistic for Gaussian initial conditions, but

gravitational evolution generates non-Gaussianities that lock significant information into higher-order correlators. Recovering this information directly, via the bispectrum and beyond [7–13], is computationally expensive and theoretically challenging in the nonlinear regime.

A growing body of work has therefore explored alternative summaries of the density field that are cheap to compute, analytically tractable, and yet capable of compressing information from multiple higher-order correlators simultaneously. Among the most promising are marked statistics [14–28]: one constructs a nonlinearly weighted version of the density field and computes its power spectrum. The resulting marked power spectrum (MPS) is sensitive to the clustering environment and can encode higher-order information while retaining the simplicity of a two-point function.

The most widely studied mark was introduced in [15], which proposed weighting the local density field by a function of the overdensity δ_R smoothed on a scale R :

$$m(\delta_R(\mathbf{x})) = \left(1 + \frac{\delta_R(\mathbf{x})}{1 + \delta_s}\right)^{-p}. \quad (1)$$

Positive values of p up-weight underdense regions (voids), while negative p enhances overdense ones. Reference [19]

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

showed that this mark provides significant improvements on σ_8 , neutrino masses, and modifications of gravity, with the common interpretation that underdense environments are particularly sensitive to variations of these parameters.

A natural question then arises: is the mark in Eq. (1) optimal, or could alternative choices capture more of the available information, for instance on the standard parameters or on primordial non-Gaussianity? This question was initially addressed in [22], where the functional form $m(\delta_R)$ was treated as a free function and optimized to maximize the Fisher information on a given parameter set, using Gaussian processes with empirical covariances and derivatives estimated from N -body simulations. Significant improvements were found across multiple parameters. An important observation in that work is that the Fisher matrix is invariant under affine transformations of the mark function, $m \rightarrow Am + B$. As a consequence, a void-enhancing mark and an overdensity-enhancing mark carry identical information when the full data vector $\{P, C, M\}$ is used. The true driver of information gain is therefore the environmental contrast modulated by the mark, rather than the preferential weighting of voids specifically. This result relies on the assumption of a Gaussian likelihood, which is a good approximation for most standard observables in the absence of strong degeneracies in the parameter space. Significant deviations from Gaussianity could in principle modify the picture. Recent machine learning frameworks have generalized the mark optimization from a fiducial cosmology to a broad parameter prior, extending the functional form of the mark through rotationally equivariant neural architectures to isolate local morphological features of the cosmic web [29].

The MPS can also be studied analytically, using the effective field theory of large scale structure (EFTofLSS) [30–33]. References [17,18,24,26] showed that expanding the mark as a power series in δ_R and computing the marked power spectrum at one-loop order yields predictions in good agreement with N -body simulations up to $k \lesssim 0.3h \text{ Mpc}^{-1}$.

In this work we study the marked statistics of the dark matter field in real space by combining the analytical control of the EFTofLSS with the flexibility of the mark optimization program. The key idea is to expand the mark as a polynomial in powers of δ_R , as is usually done in perturbation theory, and treat the coefficients of the expansion as free parameters:

$$m(\delta_R) = c_0 + c_1\delta_R + c_2\delta_R^2 + c_3\delta_R^3 + \mathcal{O}(\delta_R^4), \quad (2)$$

and optimize them to maximize the Fisher information. This turns the EFT expansion into a flexible polynomial basis, playing a role analogous to the Gaussian process (GP) approach of [22] but within a fully analytical theoretical framework. Affine invariance of the Fisher matrix is a necessary ingredient of the optimization:

without it, the problem is degenerate, since a constant shift of the mark leaves all observables invariant up to a linear rescaling. Imposing such invariance reduces the free parameter space to $\{c_1, c_2, c_3\}$ (with $c_0 = 1$ fixed by convention). A central aspect of our framework is the restriction of the optimization domain to the perturbative regime, where higher-loop contributions remain suppressed with respect to the one-loop result [34,35]. This physically motivated constraint, which we derive and discuss in detail below, substantially restricts the space of viable marks and ensures the theoretical consistency of the optimization. The Fisher information serves as the cost function, and the optimization runs in minutes on a laptop.

We validate our optimal marks against 1000 realizations of the Quijote simulation suite [36] and find that our analytical predictions reproduce the measured power spectra at the percent level. The resulting Fisher forecasts show improvements in the parameter error bars by a factor of approximately ~ 1.30 – 2.3 on Ω_m , ~ 1.3 – 8.3 on σ_8 , and ~ 1.2 – 2.5 on h , depending on redshift and smoothing scale, in good agreement with [22]. The convergence of an analytical, simulation-free approach with a GP-based one validates both the EFTofLSS description of marked statistics and the broader picture that a polynomial expansion is a sufficient functional basis for mark optimization in the perturbative regime. We focus here on dark matter in real space as a controlled starting point for assessing the information content and potential gains from optimized marks, with the framework providing a basis for future extensions to redshift space and to galaxy tracers. Because the pipeline is fast and fully analytic, it is a natural ingredient for future MCMC analyses of actual survey data, as well as a stepping stone toward optimized marks for primordial non-Gaussianity, where simulation-based approaches are considerably more expensive.

The paper is organized as follows. In Sec. II we introduce the marked field and review its perturbative description within EFTofLSS, defining the general polynomial mark and the one-loop expressions for $P(k)$, $C(k)$, and $M(k)$. In Sec. II C we derive the perturbativity constraints on the c_i coefficients. In Sec. III A we describe the Fisher matrix formalism, the covariance model, and the optimization procedure. In Sec. IV we present our main results: the optimal marks, the comparison with Quijote simulations, and the forecasted parameter improvements. We conclude and outline future directions in Sec. V.

II. THEORETICAL MODEL

A. General mark function

In this work we will focus on the dark matter density contrast field $\delta(\mathbf{x})$, modeled in real space. Extensions to redshift space and biased tracers would bring the framework into direct contact with galaxy survey data and are left

for future work. The dark matter density field $\delta(\mathbf{x})$ is defined as

$$\delta(\mathbf{x}) = \frac{\rho(\mathbf{x})}{\bar{\rho}} - 1 \quad (3)$$

where $\rho(\mathbf{x})$ is the matter density and $\bar{\rho} \equiv \langle \rho(\mathbf{x}) \rangle$ is its mean value at a fixed time. Given a density perturbation field, one can always define its smoothed version on a comoving scale R as

$$\delta_R(\mathbf{x}) = \int d^3\mathbf{z} W_R(|\mathbf{z}|) \delta(\mathbf{x} - \mathbf{z}), \quad \delta_R(\mathbf{k}) = W_R(k) \delta(\mathbf{k}), \quad (4)$$

where W_R is a generic smoothing function that we will assume to be gaussian in Fourier space.¹ Previous works [14–27] have shown that this quantity can be exploited to construct a new, nonlinear map of the density field itself, in order to extract more information from the full dark matter distribution and environment. The technical details of the construction of the marked field are presented in Sec. III A. The mark function usually adopted in literature is [15,19,27]

$$m(\delta_R(\mathbf{x})) = \left(1 + \frac{\delta_R(\mathbf{x})}{1 + \delta_s} \right)^{-p}. \quad (6)$$

This choice of the mark (not being the only one) is particularly suited to amplify, in a very intuitive way, the specific features of the density field environment: positive values of p enhance underdense regions (voids), while overdense regions (density peaks) are magnified when $p < 0$. This functional form of the mark was introduced based on the hypothesis that low-density regions are more sensitive to physics beyond the Λ CDM model, particularly in the context of massive neutrinos [17–20], modified gravity [15,16,21], and primordial non-Gaussianities [26]. The marked field is then defined as

$$\rho_M(\mathbf{x}) = m(\delta_R(\mathbf{x})) \rho(\mathbf{x}), \quad (7)$$

and, by defining the mean mark as

$$\bar{\rho}_M = \langle \rho_M(\mathbf{x}) \rangle = \langle m(\delta_R(\mathbf{x})) \rho(\mathbf{x}) \rangle \equiv \bar{m} \bar{\rho}, \quad (8)$$

one can obtain the expression for the marked overdensity field

¹In Fourier and in configuration space it reads

$$W_R(k) = e^{-\frac{k^2 R^2}{2}}, \quad W_R(x) \equiv \int \frac{d^3\mathbf{k}}{(2\pi)^3} W_R(k) = \frac{1}{(2\pi)^{3/2}} \frac{e^{-\frac{x^2}{2R^2}}}{R^3}. \quad (5)$$

$$\delta_M(\mathbf{x}) = \frac{\rho_M(\mathbf{x}) - \bar{\rho}_M}{\bar{\rho}_M} = \frac{m(\delta_R(\mathbf{x}))}{\bar{m}} [1 + \delta(\mathbf{x})] - 1. \quad (9)$$

This mapping effectively provides a nonlinear transformation of the density field: this operation creates a new (weighted) field with additional nonlinear and non-Gaussian features in its two-point correlation function. This is the reason why one should expect the marked density field to encapsulate more information about the surrounding environment when the same scales are analyzed with the power spectrum of the unmarked density field. This is similar to the findings of [26], where it was highlighted that the combination of the standard power spectrum and bispectrum has, in general, more information on cosmological parameters compared to an analysis with the marked power spectrum. We will investigate this point further in the following sections.

A related concern is that the nonlinear mapping may produce a field with large nonlinearities, potentially compromising the applicability of perturbation theory. To preserve perturbativity, we therefore impose a set of conditions on the weighting scheme. This constitutes one of the main novelties of this work and will be discussed in more detail in Sec. II C and Appendix A.

B. EFTofLSS for the marked power spectra

Perturbative methods based on the effective field theory of large scale structure [30,31] have been shown to be flexible and accurate enough to reproduce the galaxy power spectrum and bispectrum at one-loop in redshift space with sufficient accuracy up to mildly nonlinear scales. Much effort is being put in to push the EFTofLSS to higher perturbative orders [37–41] and for higher n -point functions [35,42–44]. References [17,18,24,26] showed that this analytical procedure can be safely applied to marked correlators, producing results that are in good agreement with numerical simulations, both for dark matter and halos. Here, we briefly review the methodology and the main results obtained in this framework. Throughout the analysis, we work in real space and consider the dark matter density field. Extensions to redshift space and biased tracers are natural next steps, left for future work.

With the same spirit as previous works, which adopted the mark function in Eq. (6), we start by expanding $m(\delta_R)$ as a power series of δ_R

$$\begin{aligned} m(\delta_R(\mathbf{x})) &= \sum_{n=0}^{\infty} (-1)^n \frac{p(p+1)\dots(p+n-1)}{n!} \frac{\delta_R^n(\mathbf{x})}{(1+\delta_s)} \\ &\equiv \sum_{n=0}^{\infty} c_n(p, \delta_s) \delta_R^n(\mathbf{x}), \end{aligned} \quad (10)$$

with

$$c_n(p, \delta_s) \equiv (-1)^n \frac{p(p+1)\dots(p+n-1)}{n!(1+\delta_s)^n}. \quad (11)$$

For this series to be convergent, the first requirement is that $|\delta_R(\mathbf{x})/(1+\delta_s)| \lesssim 1$, which can be translated to the condition on the fluctuation of the smoothed field

$$\langle \delta_R^2(\mathbf{x}) \rangle \equiv \sigma_{RR}^2(z) \lesssim \frac{1+\delta_s}{D(z)}, \quad (12)$$

where $D(z)$ is the growth function for the linear matter density, $\delta_R(z) \simeq (D(z)/D(z_{\text{in}}))\delta(z_{\text{in}})$. When considering the mark function employed in [19], the c_n parameters that appear in the perturbative expansion of Eq. (10) are functions of the parameters p and δ_s . In this work, we will consider a more general form of the mark function by leaving these parameters free to vary. In practice, since we are interested in the one-loop power spectrum, we will consider the mark function expanded up to third order in δ_R

$$m(\delta_R) = c_0 + c_1\delta_R + c_2\delta_R^2 + c_3\delta_R^3 + \mathcal{O}(\delta_R^4), \quad (13)$$

with c_n 's now free to vary. We still need this expansion to be convergent and meaningful in a perturbative sense. We will investigate more about this in the following section. We can now proceed with the perturbative expansion up to the third order of the marked field in Eq. (9)

$$\delta_M(\mathbf{x}) = \sum_{n=0}^3 \delta_M^{(n)}(\mathbf{x}), \quad (14)$$

where (n) indicates the n th perturbative order, and we have defined

$$\begin{aligned} \delta_M^{(1)}(\mathbf{x}) &= c_0\delta^{(1)}(\mathbf{x}) + c_1\delta_R^{(1)}(\mathbf{x}), \\ \delta_M^{(2)}(\mathbf{x}) &= c_0\delta^{(2)}(\mathbf{x}) + c_1\delta^{(1)}(\mathbf{x})\delta_R^{(1)}(\mathbf{x}) + c_1\delta_R^{(2)}(\mathbf{x}) \\ &\quad + c_2\delta_R^{(1)}(\mathbf{x})\delta_R^{(1)}(\mathbf{x}) \\ \delta_M^{(3)}(\mathbf{x}) &= c_0\delta^{(3)}(\mathbf{x}) + c_1\delta_R^{(1)}(\mathbf{x})\delta^{(2)}(\mathbf{x}) + c_1\delta_R^{(2)}(\mathbf{x})\delta^{(1)}(\mathbf{x}) \\ &\quad + c_1\delta_R^{(3)}(\mathbf{x}) + c_2\delta_R^{(1)}(\mathbf{x})\delta_R^{(1)}(\mathbf{x})\delta^{(1)}(\mathbf{x}) \\ &\quad + 2c_2\delta_R^{(1)}(\mathbf{x})\delta_R^{(2)}(\mathbf{x}) + c_3\delta_R^{(1)}(\mathbf{x})\delta_R^{(1)}(\mathbf{x})\delta_R^{(1)}(\mathbf{x}). \end{aligned} \quad (15)$$

In order to find the perturbative solution for the dark matter field and its smoothed version, it is useful to start from the dynamics of the dark matter fluid, which can be described by the standard continuity and Euler equations in combination with the Poisson equation

$$\dot{\delta} + a^{-1}\partial_i((1+\delta)v^i) = 0, \quad (16)$$

$$\dot{v}^i + H v^i + a^{-1}v^j\partial_j v^i + a^{-1}\partial_i\Phi = -\frac{\partial_j\tau^{ij}}{a\rho}, \quad (17)$$

$$\frac{\nabla^2\Phi}{a^2H^2} = \frac{3}{2}\Omega_m(a)\delta, \quad (18)$$

where ρ is the density of the energy of dark matter, δ is the density contrast already introduced in Eq. (3), v^i is the perturbation of the velocity of dark matter, and τ_{ij} is the effective stress tensor. A dot here represents a derivative with respect to the cosmic time t , while $H = d \log a / dt$ is the Hubble rate and $\Omega_m = \bar{\rho}_m / \rho_{\text{crit}}$ is the (nonrelativistic) matter density parameter. These functions of time are analytically defined in a Λ CDM universe with a cosmological constant and no curvature as

$$H(t) = H_0 \sqrt{\Omega_{m,0}a(t)^{-3} + \Omega_{\Lambda,0}}, \quad \Omega_m(t) = \frac{\Omega_{m,0}a(t)^{-3}}{(H(t)/H_0)^2},$$

$$\Omega_{\Lambda}(t) = 1 - \Omega_m(t) \quad (19)$$

with H_0 , $\Omega_{m,0}$, and $\Omega_{\Lambda,0}$ respectively being the values of the Hubble parameter, the matter density parameter, and the cosmological constant density parameter today [45]. The right-hand side of the Euler equation in Eq. (17) represents the effective stress-energy tensor, which describes the effect of short nonlinear modes on the dynamics of the long modes. The EFTofLSS provides a theoretical framework that allows us to expand it as a perturbative series of the long wavelength gravitational potential, resulting, eventually, in the addition of the so-called *counterterms*, which parametrize the cutoff dependence of the theory, effectively marginalizing and renormalizing the uncontrolled short scales. Here we will momentarily neglect the stress-energy tensor as we will briefly discuss its effect later. It is convenient to rewrite these equations in Fourier space and consider the scale factor $a(t)$ as the new time variable, such that $dt = da/\mathcal{H}$, where we have defined the conformal Hubble rate $\mathcal{H} = aH$. Moreover, we also introduce the rescaled velocity divergence, defined as

$$\theta \equiv -\frac{\partial_i v^i}{f\mathcal{H}}, \quad (20)$$

with $f \equiv d \log D / d \ln a$ being the growth factor. Neglecting vorticity modes, the continuity and Euler equations now become

$$\begin{aligned} a \frac{d\delta(\mathbf{k}, a)}{da} - f(a)\theta(\mathbf{k}, a) \\ = f(a) \int_{\mathbf{q}_1} \int_{\mathbf{q}_2} (2\pi)^3 \delta_D(\mathbf{k} - \mathbf{q}_{12}) \alpha(\mathbf{q}_1, \mathbf{q}_2) \theta(\mathbf{q}_1, a) \delta(\mathbf{q}_2, a), \end{aligned} \quad (21)$$

$$\begin{aligned}
& a \frac{d\theta(\mathbf{k}, a)}{da} - f(a)\theta(\mathbf{k}, a) + \frac{3}{2} \frac{\Omega_m(a)}{f(a)} (\theta(\mathbf{k}, a) - \delta(\mathbf{k}, a)) \\
& = f(a) \int_{\mathbf{q}_1} \int_{\mathbf{q}_2} (2\pi)^3 \delta_D(\mathbf{k} - \mathbf{q}_{12}) \beta(\mathbf{q}_1, \mathbf{q}_2) \theta(\mathbf{q}_1, a) \theta(\mathbf{q}_2, a),
\end{aligned} \quad (22)$$

where we have already included the Poisson equation in the Euler equation and made use of the well-known evolution equation for f

$$a \frac{df(a)}{da} + f^2 + \left(2 + \frac{d \ln H}{d \ln a}\right) f - \frac{3}{2} \Omega_m(a) = 0. \quad (23)$$

The functions of momenta $\alpha(\mathbf{q}_1, \mathbf{q}_2)$ and $\beta(\mathbf{q}_1, \mathbf{q}_2)$ are the standard dark matter interaction vertices [9]

$$\begin{aligned}
\alpha(\mathbf{q}_1, \mathbf{q}_2) &= 1 + \frac{\mathbf{q}_1 \cdot \mathbf{q}_2}{q_1^2}, \\
\text{and } \beta(\mathbf{q}_1, \mathbf{q}_2) &= \frac{|\mathbf{q}_1 + \mathbf{q}_2|^2 \mathbf{q}_1 \cdot \mathbf{q}_2}{2q_1^2 q_2^2}
\end{aligned} \quad (24)$$

and we have used the notation $\int_{\mathbf{q}} \equiv \int \frac{d^3 \mathbf{q}}{(2\pi)^3}$ and $\mathbf{q}_{1\dots n} \equiv \mathbf{q}_1 + \dots + \mathbf{q}_n$. As anticipated above, we can write the dark matter density and velocity contrasts as a perturbative expansion of the form

$$\delta(\mathbf{k}, a) = \sum_n \delta^{(n)}(\mathbf{k}, a), \text{ and } \theta(\mathbf{k}, a) = \sum_n \theta^{(n)}(\mathbf{k}, a), \quad (25)$$

which allows us to solve the Eqs. (21) and (22) order by order. The perturbative solution of these equations can then be expressed as a convolution in Fourier space of n linear fields over momentum kernels

$$\begin{aligned}
\delta^{(n)}(\mathbf{k}, a) &= \int \frac{d^3 \mathbf{q}_1}{(2\pi)^3} \dots \\
&\int \frac{d^3 \mathbf{q}_n}{(2\pi)^3} F_n(\mathbf{q}_1, \dots, \mathbf{q}_n) \delta^{(1)}(\mathbf{q}_1, a) \dots \delta^{(1)}(\mathbf{q}_n, a),
\end{aligned} \quad (26)$$

$$\begin{aligned}
\theta^{(n)}(\mathbf{k}, a) &= \int \frac{d^3 \mathbf{q}_1}{(2\pi)^3} \dots \\
&\int \frac{d^3 \mathbf{q}_n}{(2\pi)^3} G_n(\mathbf{q}_1, \dots, \mathbf{q}_n) \delta^{(1)}(\mathbf{q}_1, a) \dots \delta^{(1)}(\mathbf{q}_n, a),
\end{aligned} \quad (27)$$

where the F_n and G_n are the well known kernels for matter density contrast that can be obtained by imposing the so-called Einstein-de Sitter approximation for time evolution [9], which translates to $\Omega_m = f^2$. In the same fashion, we can express the marked density field with the new

convolution kernels as

$$\begin{aligned}
\delta_M^{(n)}(\mathbf{k}, a) &= \int \frac{d^3 \mathbf{q}_1}{(2\pi)^3} \dots \\
&\int \frac{d^3 \mathbf{q}_n}{(2\pi)^3} H_n(\mathbf{q}_1, \dots, \mathbf{q}_n) \delta^{(1)}(\mathbf{q}_1, a) \dots \delta^{(1)}(\mathbf{q}_n, a),
\end{aligned} \quad (28)$$

where the expressions for the H_n kernels up to third order are given by

$$\begin{aligned}
H_1(\mathbf{q}_1) &= C_{\delta_M}(q_1) F_1(\mathbf{q}_1) = C_{\delta_M}(q_1), \\
H_2(\mathbf{q}_1, \mathbf{q}_2) &= C_{\delta_M}(q_{12}) F_2(\mathbf{q}_1, \mathbf{q}_2) \\
&\quad + C_{\delta_M^2}(q_1, q_2) F_1(\mathbf{q}_1) F_1(\mathbf{q}_2), \\
H_3(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) &= C_{\delta_M}(q_{123}) F_3(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) + 2C_{\delta_M^2}(q_1, q_{23}) \\
&\quad \times F_1(\mathbf{q}_1) F_2(\mathbf{q}_2, \mathbf{q}_3) + C_{\delta_M^3}(k_1, k_2, k_3) \\
&\quad \times F_1(\mathbf{q}_1) F_1(\mathbf{q}_2) F_1(\mathbf{q}_3),
\end{aligned} \quad (29)$$

and the third order kernel has to be symmetrized over the momenta. Moreover, we have defined the functions

$$C_{\delta_M}(q_1) = c_0 + c_1 W_R(q_1), \quad (30)$$

$$C_{\delta_M^2}(q_1, q_2) = c_2 W_R(q_1) W_R(q_2) + \frac{c_1}{2} [W_R(q_1) + W_R(q_2)], \quad (31)$$

$$\begin{aligned}
C_{\delta_M^3}(q_1, q_2, q_3) &= c_3 W_R(q_1) W_R(q_2) W_R(q_3) \\
&\quad + \frac{c_2}{3} [W_R(q_2) W_R(q_3) + W_R(q_1) W_R(q_3) \\
&\quad + W_R(q_1) W_R(q_2)],
\end{aligned} \quad (32)$$

which encode the perturbative structure of the mark function through the coefficients c_n .

With all these ingredients, we are now ready to compute the theoretical prediction for the observables we will use in this work, the matter power spectrum, the marked power spectrum and the cross power spectrum among the two fields

$$\begin{aligned}
P(k) &= \langle \delta(\mathbf{k}) \delta(\mathbf{k}') \rangle', \quad C(k) = \langle \delta(\mathbf{k}) \delta_M(\mathbf{k}') \rangle', \\
M(k) &= \langle \delta_M(\mathbf{k}) \delta_M(\mathbf{k}') \rangle',
\end{aligned} \quad (33)$$

where $\langle \dots \rangle'$ represents the correlator normalized by the usual $(2\pi)^3 \delta_D(\mathbf{k} + \mathbf{k}')$ factor, and we have left the redshift dependence implicit to avoid clutter. Inserting the perturbative expansions into these expressions, we obtain the following expressions up to one-loop order

$$\begin{aligned}
P(k) &= P_{11}(k) + P_{22}(k) + P_{13}(k), \\
\bar{m}C(k) &= C_{11}(k) + C_{22}(k) + C_{13}(k), \\
\bar{m}^2M(k) &= M_{11}(k) + M_{22}(k) + M_{13}(k). \quad (34)
\end{aligned}$$

For the standard power spectrum, the one-loop terms are the usual ones

$$\begin{aligned}
P_{11}(k) &= P_{\text{lin}}(k) \equiv \langle \delta^{(1)}(\mathbf{k}) \delta^{(1)}(\mathbf{k}') \rangle', \\
P_{22}(k) &= 2 \int_{\mathbf{q}} [F_2(\mathbf{k} - \mathbf{q}, \mathbf{q})]^2 P_{\text{lin}}(|\mathbf{k} - \mathbf{q}|) P_{\text{lin}}(q), \\
P_{13}(k) &= 6 P_{\text{lin}}(k) \int_{\mathbf{q}} F_3(\mathbf{k}, -\mathbf{q}, \mathbf{q}) P_{\text{lin}}(q), \quad (35)
\end{aligned}$$

for the cross power spectrum, we obtain

$$\begin{aligned}
C_{11}(k) &= H_1(\mathbf{k}) P_{\text{lin}}(k), \\
C_{22}(k) &= 2 \int_{\mathbf{q}} F_2(\mathbf{k} - \mathbf{q}, \mathbf{q}) H_2(\mathbf{k} - \mathbf{q}, \mathbf{q}) P_{\text{lin}}(|\mathbf{k} - \mathbf{q}|) P_{\text{lin}}(q), \\
C_{13}(k) &= 3 P_{\text{lin}}(k) \int_{\mathbf{q}} [H_1(\mathbf{k}) F_3(\mathbf{k}, \mathbf{q}, -\mathbf{q}) \\
&\quad + H_3(\mathbf{k}, \mathbf{q}, -\mathbf{q})] P_{\text{lin}}(q), \quad (36)
\end{aligned}$$

and, finally, for the marked power spectrum, we have

$$\begin{aligned}
M_{11}(k) &= H_1^2(\mathbf{k}) P_{\text{lin}}(k), \\
M_{22}(k) &= 2 \int_{\mathbf{q}} [H_2(\mathbf{k} - \mathbf{q}, \mathbf{q})]^2 P_{\text{lin}}(|\mathbf{k} - \mathbf{q}|) P_{\text{lin}}(q), \\
M_{13}(k) &= 6 H_1(\mathbf{k}) P_{\text{lin}}(k) \int_{\mathbf{q}} H_3(\mathbf{k}, -\mathbf{q}, \mathbf{q}) P_{\text{lin}}(q). \quad (37)
\end{aligned}$$

As already highlighted in [26], the one-loop terms of the marked correlators (both the cross and the auto power spectra) present additional contributions due to the presence of the smoothed field in the mark function. The additional pieces are convolutions of bispectrum- and (disconnected) trispectrumlike terms with some kernels given by the marking procedure. These terms arise because the marked field is, by definition, more non-Gaussian, and already at the level of its two-point function, it contains information from higher order correlation functions of the unmarked field. This is why it is expected to gain more constraining power using marked correlators compared to the usual power spectrum analysis. However, the information gain is not expected to surpass that from a combination of higher order functions, such as the bispectrum and the trispectrum [35]. This is particularly true if one requires the marked field to still preserve perturbativity. We leave details about this in Sec. II C and Appendix A.

We now take into account the back-reaction of small-scale physics on the long wavelength modes. For this, we need to consistently reintroduce the stress-energy tensor in

Eq. (22). This has been achieved with the EFTofLSS [30,31], which represents a systematic way to account for the UV cutoff dependence of the loop integrals. The stress-energy tensor is written as a Taylor expansion of the long-wavelength quantities, imposing the equivalence principle and solving the equations perturbatively. In practice, for the standard one loop power spectrum in real space, EFTofLSS introduces an additional term at third perturbative order, which, in Fourier space, reads $\delta^{\text{ct}}(\mathbf{k}) = -c_s^2 k^2 \delta^{(1)}(\mathbf{k})$. The $\sim k^2$ dependence can be easily understood as a consequence of the equivalence principle, while the speed of sound c_s^2 is a parameter that renormalizes the UV dependence of the theory. In data analysis, c_s^2 is treated as a free parameter of the theory that has to be fitted with simulations or observational data. References [17,18,26] showed that when marked correlators are considered, no additional counterterms are required, and the relevant terms for the cross and marked power spectra are

$$C_{\text{ct}}(\mathbf{k}) = -2c_s^2 k^2 H_1(k) P_{\text{lin}}(k) \quad (38)$$

$$M_{\text{ct}}(\mathbf{k}) = -2c_s^2 k^2 H_1^2(k) P_{\text{lin}}(k). \quad (39)$$

For the shot noise and the IR-resummation of the power spectra, we follow [26].

Summarizing, the one loop EFT model for the standard, cross, and marked power spectrum of matter in real space has the following forms²

$$P(k) = P_{11}(k) + P_{22}(k) + P_{13}(k) + P_{\text{ct}}(k), \quad (40)$$

$$\bar{m}C(k) = C_{11}(k) + C_{22}(k) + C_{13}(k) + C_{\text{ct}}(k), \quad (41)$$

$$\bar{m}^2M(k) = M_{11}(k) + M_{22}(k) + M_{13}(k) + M_{\text{ct}}(k), \quad (42)$$

which eventually contain only the free parameter c_s^2 .

C. Perturbative prior

One of the key assumptions of our work is that one can safely apply perturbation techniques to marked fields as described above. Before moving on with the comparison to simulations and with the optimization of such weighting maps, we need to accurately consider if this is the case. It is well known that in standard perturbation theory [35] the key quantity that controls nonlinearities is the variance of the density field at a given scale k , defined as

$$\Delta^2(k) = \frac{1}{2\pi^2} \int_0^k dq q^2 P_{\text{lin}}(q), \quad (43)$$

²Notice that in this work we will limit our analysis to the matter field in real space in order to compare our optimization procedure to previous results. The analysis can be easily extended to biased tracers (galaxies) in redshift space and will be investigated in future works.

where $P_{\text{lin}}(k)$ is the linear power spectrum. Crucially, for perturbation theory to be accurate, this parameter must be small when evaluated at the maximum value of the wave-number considered, k_{max} . As an example, in a Λ CDM cosmology at redshift $z = 0$, the variance of the field is $\Delta^2 \simeq 0.22$ at $k = 0.12 \text{ h/Mpc}$ and $\Delta \simeq 0.53$ at $k = 0.2 \text{ h/Mpc}$, getting smaller at higher redshift and larger for higher values of k , as expected. For perturbation theory to be valid, it is essential that not only the relative size of each individual term in perturbative calculations is of order $\sim \Delta^2(k_{\text{max}})$, but also that the *total* loop contribution to the correlation functions has the same size. This assumption may seem obvious, but it has nontrivial consequences: for a biased tracer, such as halos or galaxies, the total m -loop contributions to a given n -point function can have a large number of terms multiplied by a large number of different nuisance parameters. Assuming that the variance of the density fields controls the perturbative expansion and that it is small corresponds to the requirement that all the bias parameters and counterterms, despite being of order $\sim \mathcal{O}(1)$, cannot contribute to large loop corrections for a realistic tracer. In this work, we are considering the matter field; hence, these assumptions should always be valid, at least on large scales well above the nonlinear scale. However, the marking procedure produces a new field that is, in general, a biased and more nonlinear tracer of the underlying dark matter field. The new parameters introduced in the mark function, the c_i 's, distort the original dark matter field in a specific way, and we need to verify if the assumptions we made so far are still valid for the marked field. A way to estimate if the perturbative treatment of the n -point functions is under control is to use the following (approximate) signal-to-noise (SNR) ratio formula for a connected n -point function [35]

$$(\text{SNR})_n^2 \simeq \int^{k_{\text{max}}} \frac{d^3 \mathbf{k}_1}{(2\pi)^3} \dots \int^{k_{\text{max}}} \frac{d^3 \mathbf{k}_n \langle \delta(\mathbf{k}_1) \dots \delta(\mathbf{k}_n) \rangle_c^2}{(2\pi)^3 P(k_1) \dots P(k_n)} \sim N_{\text{pix}} \Delta^{2(n-2)}(k_{\text{max}}) [1 + \mathcal{O}(\Delta^2(k_{\text{max}}))], \quad (44)$$

where $N_{\text{pix}} \equiv V \int^{k_{\text{max}}} d^3 \mathbf{k} / (2\pi)^3$ is the number of pixels of Fourier modes in a given survey. This expression is approximate in two respects: first, the covariance is evaluated at leading (Gaussian) order, retaining only the disconnected contribution $P(k_1) \dots P(k_n)$ in the denominator; second, the final scaling assumes standard perturbation theory power counting, $\langle \delta^n \rangle_c \sim \Delta^{2(n-1)}$, which holds in the weakly nonlinear regime. Moreover, numerical prefactors of order $\sim \mathcal{O}(1)$ arising from the angular integration over the momentum structure of the correlators are not tracked. Note that the same parameter Δ^2 controls both the SNR for the n -point functions and the corrections to the SNR due to higher loop contributions. This is expected in

perturbation theory where all nonlinearities are indeed controlled by Δ^2 .³ This result shows two main things: on one hand, the signal-to-noise in each higher order n -point function is $\sim \mathcal{O}(\Delta^2)$ smaller than the previous one, meaning that the information content of the nonlinear field is almost saturated by the leading summary statistics; on the other hand, Eq. (44) shows that, analogously, the ℓ -loop contribution to the signal-to-noise ratio of the n -point function scales as $\sim \mathcal{O}(N_{\text{pix}} \Delta^{2(n+\ell-2)})$. This scaling implies that higher-loop contributions are increasingly suppressed, reflecting the perturbative nature of the nonlinear fields up to a cutoff scale k_{max} .

The same arguments must hold for the marked field since we are interested in using a perturbative approach to predict the shape of its summary statistics. Note that these assumptions are strictly needed within the EFTofLSS and are not required in other approaches, such as those employed in simulation based analyses or machine learning. For a marked field δ_M , the parameter controlling the perturbative expansion is

$$\Delta_M^2(k) = \frac{1}{2\pi^2} \int^{k_{\text{max}}} dq q^2 M_{11}(q), \quad (45)$$

with $M_{11}(k) = C_{\delta_M}^2(k) P_{\text{lin}}(k)$. We can expand Eq. (45) and obtain

$$\begin{aligned} \Delta_M^2(k) &= \frac{1}{2\pi^2} \int^{k_{\text{max}}} dq q^2 P_{\text{lin}}(q) - \frac{c_1}{\pi^2} \int^{k_{\text{max}}} dq q^2 W_R(q) P_{\text{lin}}(q) \\ &\quad + \frac{c_1^2}{2\pi^2} \int^{k_{\text{max}}} dq q^2 W_R^2(q) P_{\text{lin}}(q) \\ &\equiv \Delta^2(k) - c_1 \Delta_W^2(k) + c_1^2 \Delta_{W^2}^2(k), \end{aligned} \quad (46)$$

where we have assumed $c_0 = 1$, as we will do in the rest of the paper, and we have defined the integrals

$$\begin{aligned} \Delta_W^2(k) &\equiv \frac{1}{\pi^2} \int^{k_{\text{max}}} dq q^2 W_R(q) P_{\text{lin}}(q), \\ \Delta_{W^2}^2(k) &\equiv \frac{1}{2\pi^2} \int^{k_{\text{max}}} dq q^2 W_R^2(q) P_{\text{lin}}(q), \end{aligned} \quad (47)$$

which have values $\Delta_W^2 \simeq 0.07$ and $\Delta_{W^2}^2 \simeq 0.015$ for $k_{\text{max}} \simeq 0.25 \text{ h/Mpc}$ at redshift $z = 0$ for $R = 30 \text{ Mpc/h}$, $\Delta_W^2 = 0.26$, and $\Delta_{W^2}^2 = 0.07$ with $R = 15 \text{ Mpc/h}$. This means two things: a smaller smoothing radius R makes the theory for the marked field less perturbative; however, the variance of the marked field can still be $\Delta_M^2 \lesssim 1$ if the value of c_1 is adjusted accordingly to respect this condition.

³Notice that there are some special cases where new parameters can play a non-negligible role, causing a failure of the perturbative expansion, see the discussion in [46]. For the marked power spectrum, similar effects can emerge due to the nonlinear remapping, as described in the main text.

As a consequence of this, we expect the assumption of perturbativity to be safely satisfied by the marked field. In addition, one can see that the one loop contribution to the signal-to-noise can produce terms that remain constant even at very large scales, mimicking the effect of a shot-noiselike term [21]. This is somewhat similar to what is discussed in [46] for the case of biased tracers, where the term presenting the second order bias b_2 can give an unwanted non-negligible contribution to the nonlinear trispectrum in the squeezed limit. This kind of term is usually under perturbative control for real life tracers: in the case of marked fields, these terms may arise due to the weighting procedure and could lead to a potential breaking of the perturbation theory assumptions. We leave further details about this in Appendix A.

Imposing that the numerator of Eq. (A8) is smaller than its denominator translates into a condition for c_2 and c_1 , as shown in Fig. 1. The near-symmetry of the allowed region around an offset point is a consequence of fixing the constant coefficient to $c_0 = 1$: the perturbativity condition depends on the relative contributions of the different terms in the mark, so choosing a nonzero c_0 shifts the corresponding region in the (c_1, c_2) plane away from the origin. What is shown in this figure is somewhat expected: for larger values of the smoothing radius, the perturbative approach is more convergent, which leads to a larger allowed parameter space. It would be interesting, though, to explore the contributions coming from even higher loops than the one considered in this work. Perturbation theory requires that, for example, the two loop order is always smaller than the total power spectrum computed at one loop. This condition, usually referred to as *perturbativity*

prior [34], has been recently exploited in real data analysis [47,48] and relies on the arguments presented here. Interestingly, we could impose such perturbativity condition also for the marked power spectrum: in Appendix A, we show that terms from the two loop marked power spectrum can give a non-negligible contribution on large scales. This should not surprise us, given the one loop example of M_{22} . What is interesting is that pieces coming from the two loop include also the mark parameter c_3 and other combinations of c_2 and c_1 . In Appendix A, we also show that for scenarios in which the perturbation theory is not valid anymore, nonlinear scales significantly leak into larger scales. For all the reasons stated above, we find it more convenient to impose a prior that is easier to implement, such that each successive term in the expression of the mark function must be suppressed with respect to the previous one when evaluated at the maximum value of the smoothed overdensity field $\delta_{R,\max}$, measured from the simulations, as reported in Table I. This translates into the two conditions

$$|c_2|\delta_{R,\max} < |c_1|, \quad |c_3|\delta_{R,\max} < |c_2|. \quad (48)$$

Points in parameter space that violate either of these conditions are excluded from the scan, as they would correspond to a mark whose higher-order terms dominate the lower-order ones, invalidating the perturbative interpretation of the model. We will additionally require that the mark functions are positive in the range of δ_R sampled by the simulations. This condition is both physically and operationally motivated: a negative mark would flip the sign of the contribution of the corresponding region to the

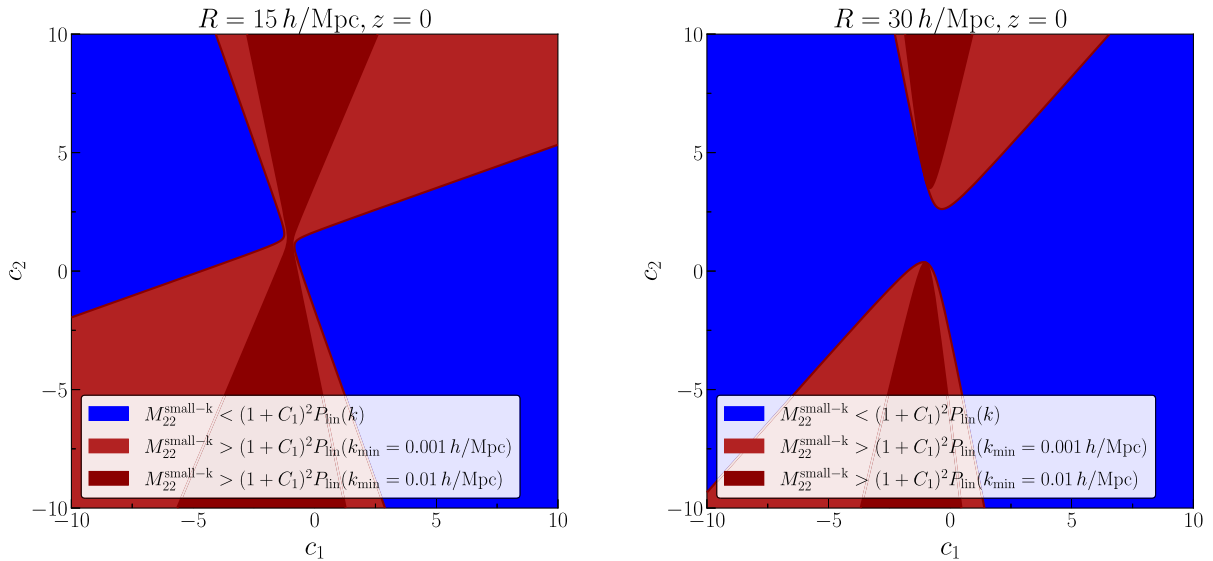


FIG. 1. Prior on the coefficients c_1 and c_2 imposed by requiring perturbativity of the marked field, shown for the two smoothing scales $R = 15 \text{ Mpc}/h$ and $R = 30 \text{ Mpc}/h$ at $z = 0$. The blue region indicates the allowed parameter space, while the red regions highlight the excluded ones for two different choices of the minimum wave number $k_{\min}$. The allowed region shrinks as the smoothing radius decreases, reflecting the more nonlinear structure probed at smaller R .

TABLE I. Minimum and maximum values of the smoothed overdensity field δ_R , measured from the Quijote simulations at $z = \{0, 0.5, 1\}$ for different smoothing radii. Density contrasts are progressively suppressed at higher redshift, while smaller smoothing scales probe significantly larger overdensities at all redshifts.

$R[h^{-1} \text{ Mpc}]$	$\delta_{R,\min}(z=0)$	$\delta_{R,\max}(z=0)$	$\delta_{R,\min}(z=0.5)$	$\delta_{R,\max}(z=0.5)$	$\delta_{R,\min}(z=1)$	$\delta_{R,\max}(z=1)$
30	-0.41	0.68	-0.34	0.49	-0.26	0.42
15	-0.66	2.22	-0.57	1.54	-0.45	1.20
10	-0.77	5.47	-0.69	3.58	-0.56	2.79

marked field, making the statistical interpretation of the marked power spectrum as a weighted summary statistic ill-defined. In practice, enforcing positivity ensures that $m(\delta_R)$ can always be interpreted as a well-defined, non-negative weight function assigned to each region of the density field.

III. MINIMIZATION PROCEDURE

The aim of this paper is to identify the optimal configuration of the mark function that maximizes the information gain from marked correlators relative to a standard power spectrum-only analysis. In particular, we explore the constraining power on a set of cosmological parameters, $\{\Omega_m, \sigma_8, h\}$, while marginalizing over the EFT parameter c_s^2 . To this end, we vary the shape of the mark function parametrized by $\{c_0, c_1, c_2, c_3\}$, fixing $c_0 = 1$ as discussed in Sec. III C. We scan a three-dimensional grid for $\{c_1, c_2, c_3\}$ in the range $[-10, 10]$ subject to two physically motivated priors. We adopt a grid-based scan as we found gradient-based methods to converge to the same optimal mark at a higher computational cost. The mark function in Eq. (13) is intended as a controlled expansion around the smoothed density field, and for this reason, we impose a perturbative prior on the c_i coefficients as discussed in II C. For each configuration, we compute the Fisher matrix using the theoretical model outlined above, including both derivatives and covariance as detailed in Secs. III A and III B. We compare the constraints obtained from the power spectrum alone with those from the combination of power spectrum, cross spectrum, and marked power spectrum, quantifying the improvement through the ratio $r(\theta) = \sigma_P(\theta)/\sigma_{P+C+M}(\theta)$ (or equivalently the FoM, see Sec. III B) for each of the selected parameters $\theta = \{\Omega_m, \sigma_8, h\}$, while marginalizing over the EFT parameter c_s^2 . After identifying the optimal mark configurations, we compute the full non-Gaussian covariance using 1000 realizations from the Quijote suite at the fiducial cosmology. This allows us to assess the actual improvement in parameter constraints when a realistic covariance is taken into account. Additional methodological details are provided in the following sections.

A. Details about the Fisher matrix

In this section, we detail the procedure adopted to perform the Fisher forecast for cosmological parameters,

combining the standard power spectrum, the cross power spectrum, and the marked power spectrum for the dark matter field. We focus on the fiducial Λ CDM cosmology adopted in Quijote and consider variations in the matter density Ω_m , the amplitude of fluctuations σ_8 , and the reduced Hubble parameter $h = H_0/100$. These parameters have been shown to drive the largest variations in large scale structure observables at low redshift and hence are the most promising in terms of constraining power from LSS. In addition, we include and marginalize over the EFT counterterm c_s^2 , which parametrizes the impact of small-scale physics on the one-loop power spectrum through a $k^2 P_{\text{lin}}(k)$ correction; we will discuss possible degeneracies among these parameters and the role played by the marked power spectrum in breaking them in Sec. IV B.

The Fisher matrix provides an extremely useful tool for forecasts in cosmology [49]. Forecast methods based on the Fisher matrix rely on the Cramér-Rao lower bound, which establishes that the covariance matrix of any unbiased estimator of the parameter set θ has a lower limit given by the inverse of the Fisher information matrix $F_{\mu\nu}$. The latter is defined as the expectation value of the Hessian of the log-likelihood

$$F_{\mu\nu} = -\left\langle \frac{\partial^2 \log \mathcal{L}}{\partial \theta_\mu \partial \theta_\nu} \right\rangle. \quad (49)$$

By assuming that, at leading order, the observables considered follow a Gaussian distribution with mean D and covariance C_D , the Fisher matrix is

$$F_{\mu\nu} = \sum_{i,j} \frac{\partial D_i}{\partial \theta_\mu} (C_D^{-1})_{ij} \frac{\partial D_j}{\partial \theta_\nu}, \quad (50)$$

where the sum runs over the data bins. The inverse Fisher matrix in Eq. (50) gives the lower limit for the error bar for a given parameter θ_μ , which satisfies $\sigma^2(\theta_\mu) \geq (F^{-1})_{\mu\mu}$. These estimates include marginalization over the other parameters considered in the forecast. In order to compute the Fisher matrix, one needs to assume fiducial parameters θ_{fid} , over which the derivatives are estimated. The observables we consider in this work are the power spectrum alone $D_P = \{P(k)\}$ and the combination of the standard dark matter power spectrum $P(k)$, the cross power spectrum among the unmarked and marked field $C(k)$, and the

power spectrum of the marked field $M(k)$, i.e., $D_{PCM} = \{P(k), C(k), M(k)\}$. We will compute the Cramér-Rao bounds in both cases and compare them to quantify the improvement in the resulting parameter constraints. As already discussed, each of these quantities encodes complementary information about the underlying density field: while $P(k)$ captures the clustering of matter on large scales, $M(k)$ weights the contribution of low- and high-density regions according to the adopted mark function, effectively probing different environmental regimes [14–27]. When LSS data are considered, the three cosmological parameters examined in this analysis are the ones that are better constrained from the shape of the power spectrum: other Λ CDM parameters such as the amount of baryonic matter Ω_b or the primordial tilt n_s are usually fixed using priors from other datasets, such as CMB data or measurements from BBN [50–52]. However, given that our approach relies on a theoretical model, including additional parameters is straightforward; see, for example, [26].⁴ The choice of the reference values for the cosmological parameters is dictated by the fiducial cosmology of the Quijote simulation suite [36]

$$\{\Omega_m, \sigma_8, h, n_s, \Omega_b\} = \{0.3175, 0.834, 0.6711, 0.9624, 0.049\}, \quad (51)$$

while the fiducial value for the parameter c_s^2 is obtained by fitting the unmarked power spectrum at fixed cosmology. The derivatives with respect to a given cosmological parameter θ_μ are computed using our analytical model using the symmetric finite-difference scheme:

$$\frac{dD_i(\theta_\mu)}{d\theta_\mu} = \frac{D_i(\theta_\mu + (\Delta\theta_\mu)/2) - D_i(\theta_\mu - (\Delta\theta_\mu)/2)}{\Delta\theta_\mu}. \quad (52)$$

All other parameters are held fixed while varying θ , ensuring that the derivative isolates the physical dependence of the observable on that specific cosmological parameter. This approach has been extensively validated in the context of Fisher forecasts for large scale structure [e.g., [15, 34, 53, 54]]. The finite steps $\Delta\theta$ are chosen to be small enough to accurately estimate the derivatives while avoiding numerical instabilities due to round-off errors: we use $\Delta\Omega_m = 0.02$, $\Delta\sigma_8 = 0.03$, and $\Delta h = 0.01$. We have verified that the derivatives computed using our analytical prescription with these values for the Δ 's agree with those obtained numerically from the Quijote suite, which provides simulations at fiducial and displaced parameter values for derivative estimation, and the agreement is stable across all the parameters, the redshift, and the range of scales considered.

⁴Notice that in Ref. [22] the authors vary only Ω_m and σ_8 due to the high computational power of their pipeline. In Sec. IV we show a more direct comparison.

B. Covariance matrices

The evaluation of the Fisher matrix for a given set of parameters requires an estimate of the covariance matrix of the observables involved in the forecasting procedure. Crucially, the covariance matrices of the cross and marked power spectra depend on the choice of the parameters c_i 's that enter the mark function as in Eq. (13): this means that the covariances have to be recomputed each time during the optimization procedure, where the c_i 's are varied, even though the cosmological parameters are kept fixed to their fiducial values. As a first step in our optimization procedure, we use perturbation theory to compute the Gaussian part of the fields involved. As already clarified in Sec. II C, non-Gaussian contributions to the covariance matrix (off-diagonal terms) are expected to be of order $\sim \Delta^2(k)$ and could be safely neglected for the redshift and the scale we are considering in this work. The expressions we will use for the covariances of the autospectra are given by [26]

$$\begin{aligned} \text{Cov}[P(k_i)P(k_j)] &= 2 \frac{(2\pi)^3}{VV_s} P^2(k_i) \delta_{ij}, \\ \text{Cov}[C(k_i)C(k_j)] &= \frac{(2\pi)^3}{VV_s} [P(k_i)M(k_i) + C^2(k_i)] \delta_{ij}, \\ \text{Cov}[M(k_i)M(k_j)] &= 2 \frac{(2\pi)^3}{VV_s} M^2(k_i) \delta_{ij}, \end{aligned} \quad (53)$$

while the cross-covariances between the different spectra are

$$\begin{aligned} \text{Cov}[P(k_i)C(k_j)] &= 2 \frac{(2\pi)^3}{VV_s} P(k_i)C(k_i) \delta_{ij}, \\ \text{Cov}[P(k_i)M(k_j)] &= 2 \frac{(2\pi)^3}{VV_s} C^2(k_i) \delta_{ij}, \\ \text{Cov}[C(k_i)M(k_j)] &= 2 \frac{(2\pi)^3}{VV_s} C(k_i)M(k_i) \delta_{ij}. \end{aligned} \quad (54)$$

Specific configurations of the mark parameters could lead to a potentially large correlation between the dark matter density perturbation field δ and the marked field δ_M : the particular case $\{c_0, c_1, c_2, c_3\} = \{1, 0, 0, 0\}$ would give a singular joint covariance matrix, manifesting itself as an ill-conditioned forecast case, whose inversion is numerically unstable. This is also true, for example, for mark configurations close to the one presented. To mitigate this, we employ the pseudoinverse of the covariance matrix, defined using the Moore-Penrose formalism, and introduce a threshold parameter $f_{\text{cond}} = 10^{-9}$ below which the eigenvalues are truncated, following [22, 55]. This regularization ensures that the Fisher matrix defined in Eq. (50) remains numerically stable and well defined. We tested the impact of the condition number threshold f_{cond} . For larger values, such as $f_{\text{cond}} = 10^{-6}$, too many eigenvalues are truncated, leading to unphysically large error bars. In the range

$f_{\text{cond}} = 10^{-9} - 10^{-10}$, the estimated uncertainties are stable, varying at the $\sim 10\%$ level, as shown in Fig. 6. Smaller values of f_{cond} instead lead to unstable and overly optimistic results. We therefore adopt $f_{\text{cond}} = 10^{-9}$ as a conservative choice, which we expect to yield more reliable results, particularly for smaller values of R . This approach follows standard practice in Fisher analyses involving correlated observables [56,57].

Given the relatively low computational power required to estimate the theoretical derivatives and the Gaussian covariances, we vary the mark functions in order to identify the optimal configuration of c_i 's, which maximizes the cosmological information content. We explore two complementary optimization criteria:

- (i) Global optimization: maximization of the figure of merit (FoM) defined as

$$\text{FoM} = \sqrt{\det(F_{\mu\nu})}, \quad (55)$$

which measures the inverse hypervolume of the confidence ellipsoid in parameter space. A large FoM indicates a smaller joint uncertainty on the parameters, corresponding to a more constraining dataset; and

- (ii) Individual optimization: minimization of the marginalized uncertainty on each parameter $\theta = \{\Omega_m, \sigma_8, h\}$ obtained using the Cramér-Rao bound

$$\sigma^2(\theta_\mu) = (F^{-1})_{\mu\mu}. \quad (56)$$

This approach highlights whether the improvement from marked statistics primarily benefits specific cosmological directions (e.g., σ_8) rather than a global reduction of degeneracies.

In practice, the FoM is sensitive to the degree of degeneracy among parameters, while the marginalized errors are more sensitive to the amplitude of the derivative signals. It is therefore not guaranteed that the mark configuration maximizing the FoM also minimizes the uncertainty on each parameter individually. We therefore explore both optimization strategies in parallel. Once the optimal configurations of the mark functions are obtained using a Gaussian covariance, we proceed to the estimate of the full non-Gaussian covariance using the Quijote simulations: we compute the variance of the different observables among 1000 different realizations at the fixed fiducial cosmology. This refined evaluation is necessary in order to have a more realistic covariance that captures possible mode couplings that are not included in the Gaussian one.

This two-stage approach follows the methodology of Cowell *et al.* [22] and ensures that our conclusions about the information gain from marked statistics are robust to the inclusion of realistic covariance effects.

The forecast is performed at the redshifts $z = 0$, $z = 0.5$, and $z = 1$, and for smoothing scales $R = 10 \text{ Mpc}/h$,

$R = 15 \text{ Mpc}/h$, and $R = 30 \text{ Mpc}/h$. This choice ensures that the derived constraints are also representative of low- z observables, where nonlinearities play an important role. The Fourier-space binning and survey volume are matched to those of the Quijote simulations [36], with $V = 1 (\text{Gpc}/h)^3$ and $\Delta k = (2\pi)/V^{1/3} \simeq 0.0063 \text{ h}/\text{Mpc}$. Furthermore, we will use a maximum wave number of $k_{\text{max}} = 0.30 \text{ h}/\text{Mpc}$. This scale cut is chosen specifically for the present dark-matter, real-space setup, for which we have explicitly verified that the one-loop marked-statistics model reproduces the simulations at the percent level up to $k_{\text{max}} \simeq 0.30 \text{ h}/\text{Mpc}$. It should therefore not be interpreted as a universal scale cut for marked statistics: its applicability to biased tracers in redshift space, with their additional bias and redshift-space contributions, will require a dedicated validation.

C. Affine invariance and optimal weights

One of the key results of [22] resides in the exploitation of the invariance of the Fisher matrix under affine transformations of the mark function, so that

$$\begin{aligned} M(\delta_R) &\rightarrow M'(\delta_R) = AM(\delta_R) + B, \\ (F_{\mu\nu})^{-1} &\rightarrow (F'_{\mu\nu})^{-1} = (F_{\mu\nu})^{-1}. \end{aligned} \quad (57)$$

In our approach, the transformation in Eq. (57) corresponds to a shift in the $\{c_n\}$ parameters

$$\{c_0, c_1, c_2, c_3\} \rightarrow \{c'_0, c'_1, c'_2, c'_3\} = \{Ac_0 + B, Ac_1, Ac_2, Ac_3\}. \quad (58)$$

One can easily see that the observables considered here, under this affine transformation, are changed as

$$P(k) \rightarrow P'(k) = P(k), \quad (59)$$

$$C(k) \rightarrow C'(k) = BP(k) + AC(k), \quad (60)$$

$$M(k) \rightarrow M'(k) = B^2P(k) + 2ABC(k) + A^2M(k), \quad (61)$$

and the covariance matrix transforms as $C \rightarrow C' = N \cdot C \cdot N^\dagger$, with

$$N \equiv \begin{pmatrix} 1 & 0 & 0 \\ B & A & 0 \\ B^2 & 2AB & A^2 \end{pmatrix}. \quad (62)$$

These transformation laws also hold for our theoretical modeling since perturbation theory only involves linear operations among the different spectra. The Fisher matrix is invariant under this affine transformation. Interestingly, the affine invariance of the Fisher information matrix also extends naturally to any complete set of n -point correlators

of the density field δ and marked field Δ (by complete, we mean including all the available auto- and cross-spectra); this is shown explicitly in Appendix B.

Equation (58) indicates that the parameter space of the mark functions can be reduced: the invariance of the Fisher matrix under a constant shift of the mark function implies that the information content is independent of c_0 , which can therefore be chosen arbitrarily. For convenience,⁵ we will fix $c_0 = 1$. The invariance under linear scaling $m(\delta_R) \rightarrow Am(\delta_R)$ leaves the ratio of the mark function unchanged for two different values of δ_R , explicitly

$$\frac{m(\delta_{R,1})}{m(\delta_{R,2})} \rightarrow \frac{m'(\delta_{R,1})}{m'(\delta_{R,2})} = \frac{m(\delta_{R,1})}{m(\delta_{R,2})}. \quad (63)$$

In other words, this implies that if $g(c_1, c_2, c_3)$ is a linear combination of the three parameters, the relation

$$g(c'_1, c'_2, c'_3) = \kappa(A)g(c_1, c_2, c_3), \quad (64)$$

holds, with κ being a generic function of A . A clear consequence of this is that what is relevant, in terms of information content, is not the absolute values of the mark parameters c_1, c_2, c_3 , but their relative size, i.e., the ratios c_2/c_1 and c_3/c_1 . Since we impose the perturbativity prior described in Eq. (48) (see also the discussion in Appendix A), which directly constrains these ratios (e.g., $|c_2| \lesssim |c_1|/\delta_{R,\max}$ and similarly for c_3), the physically relevant region of parameter space is naturally restricted. In this sense, the prior already reduces the effective redundancy associated with the overall normalization of the mark. This choice simplifies the numerical implementation and allows us to scan the space of mark shapes without explicitly fixing a normalization convention.

A consequence of affine invariance is that the mark function $m(\delta_R)$ itself does not provide physically meaningful information, since it is defined only up to an equivalence class under $m \rightarrow Am + b$. When a specific mark is found to improve constraints on a cosmological parameter, the apparent physical interpretation (for example, that information resides preferentially in underdense or overdense regions) is gauge-dependent and therefore not invariant. As the number of n -point correlators included in the data vector grows, the information becomes distributed across an increasing number of cross-correlators, making it progressively harder to attribute the information gain to any particular density environment. The physically meaningful quantity is the equivalence class $[m]$, not any particular representative.

This conclusion holds in the regime in which the likelihood is nearly gaussian and could be approximated

with the Fisher matrix: highly non-Gaussianities in the likelihood, such as those sourced by large degeneracies in the parameter space [58–62], may alter this picture and lead to different conclusions.

For explicitly rotation-invariant marks of the form introduced in [15], or like those considered in this work, which depend only on the scalar smoothed density field $m(\delta_R)$, the full invariant content reduces to two quantities: the smoothing scale R , which sets the spatial scale at which the local environment is defined, and the relative weighting between density environments. Concretely, the affine-invariant content is encoded in ratios of the form

$$\frac{|m(\delta_{R,1}) - m(\delta_{R,3})|}{|m(\delta_{R,3}) - m(\delta_{R,2})|} \quad (65)$$

for fixed density values $\delta_{R,1}, \delta_{R,2}, \delta_{R,3}$. The sign of this contrast is itself gauge-dependent, since the reflection $m \rightarrow -m$ is an affine transformation with $A = -1$, $B = 0$, mapping a void-upweighting mark onto a halo-upweighting one while leaving the Fisher information unchanged.

This analysis also suggests a possible route to interpretability: removing the matter power spectrum $P(k)$ breaks the full affine invariance, since the shift symmetry $m \rightarrow m + b$ no longer closes on the reduced data vector, rendering the zero-crossing δ_R^* of the optimal mark a genuine invariant with a physical interpretation as the optimal density threshold separating information-rich from information-poor environments. This of course comes at a price, as becomes clear when examining the Fourier-space content of the marked statistics. At mildly nonlinear scales, the marked power spectrum is equivalent to a particular projection of the matter bispectrum, where the mark function acts as a convolution kernel weighting different triangle configurations. Specifically, $M(k)$ probes squeezed bispectrum configurations $B(k, k, q)$ with $q \lesssim 1/R$, since the smoothing suppresses small-scale modes in $m(\delta_R)$, while $C(k)$ is sensitive to less-squeezed, more equilateral configurations. Removing $C(k)$ is therefore equivalent to discarding sensitivity to equilateral and folded bispectrum configurations, mostly sourced by gravitational evolution. Nevertheless, this approach is of considerable interest, as the optimal density threshold δ_R^* could provide a direct window into the relationship between the parameter-dependent information content of the field and its morphological properties. We note, however, that breaking affine invariance by removing $C(k)$ requires revisiting several assumptions underlying the present analysis: in particular, c_0 is no longer freely fixable to unity by gauge choice and becomes a physically meaningful parameter. The perturbative priors on the mark coefficients, which in the current framework constrain the ratios c_2/c_1 and c_3/c_1 independently of the overall normalization, may likewise need to be reassessed in this new context. We defer a systematic investigation of this approach to future work.

⁵We fix the value of c_0 to 1 and not, for example, to 0, to avoid instabilities in the theoretical modeling during the optimization operation.

IV. RESULTS AND DISCUSSION

We now discuss the application of the analytical optimization procedure described in the previous sections. Exploiting the availability of a fully analytical EFTofLSS prediction, we rapidly evaluate the Fisher information matrix across any mark configuration parametrized by $\{c_0, c_1, c_2, c_3\}$, scanning a three-dimensional grid of $100 \times 100 \times 100$ values for c_1, c_2 , and c_3 in the range $[-10, 10]$, subject to the perturbativity prior discussed in Sec. II C. We focus on three smoothing radii, $R = 10, 15$ and $30 \text{ Mpc}/h$. These are chosen based on the analysis performed in [27], where the usual mark functional dependence was adopted to analyze the BOSS dataset. We also investigate the redshift dependence of the optimal mark considering $z = 0$, to allow comparison with [22], $z = 0.5$ and $z = 1$, bracketing the redshift range targeted by realistic surveys such as BOSS [63], DESI [3], and Euclid [2]. The optimal mark coefficients, obtained using the Gaussian covariance approximation, are reported in Table III.

The resulting optimal mark functions are shown in Fig. 2. The picture that emerges depends significantly on the smoothing scale considered. For $R = 10 \text{ Mpc}/h$, the perturbativity prior restricts the allowed parameter space

$\{c_1, c_2, c_3\}$ considerably, and all optimal marks collapse onto a similar, smooth, monotonically decreasing function of δ_R , closely resembling a linearized expansion of the standard mark around $\delta_R \simeq 0$. In this regime, the optimization has little freedom and the resulting mark functions are largely insensitive to the choice of target parameter or redshift. For $R = 15$ and $30 \text{ Mpc}/h$, where the perturbativity prior is less restrictive, the optimal marks display a richer structure: they develop a parabolic shape over the range of δ_R sampled by the simulations (see Table I), enhancing the differential weighting between high- and low-density environments. The variation with redshift is more pronounced at $R = 15 \text{ Mpc}/h$ than at $R = 30 \text{ Mpc}/h$, which we discuss further below.

This outcome is consistent with the affine invariance of the Fisher matrix: as discussed in Sec. II, when the full data vector $\{P, C, M\}$ is adopted, any mark function that up-weights underdense regions carries exactly the same information as one that up-weights overdense regions because the two are related by an affine transformation that leaves the Fisher matrix invariant. The actual information gain in the marked field therefore does not arise from the voids themselves, as sometimes implied in the literature [19], but

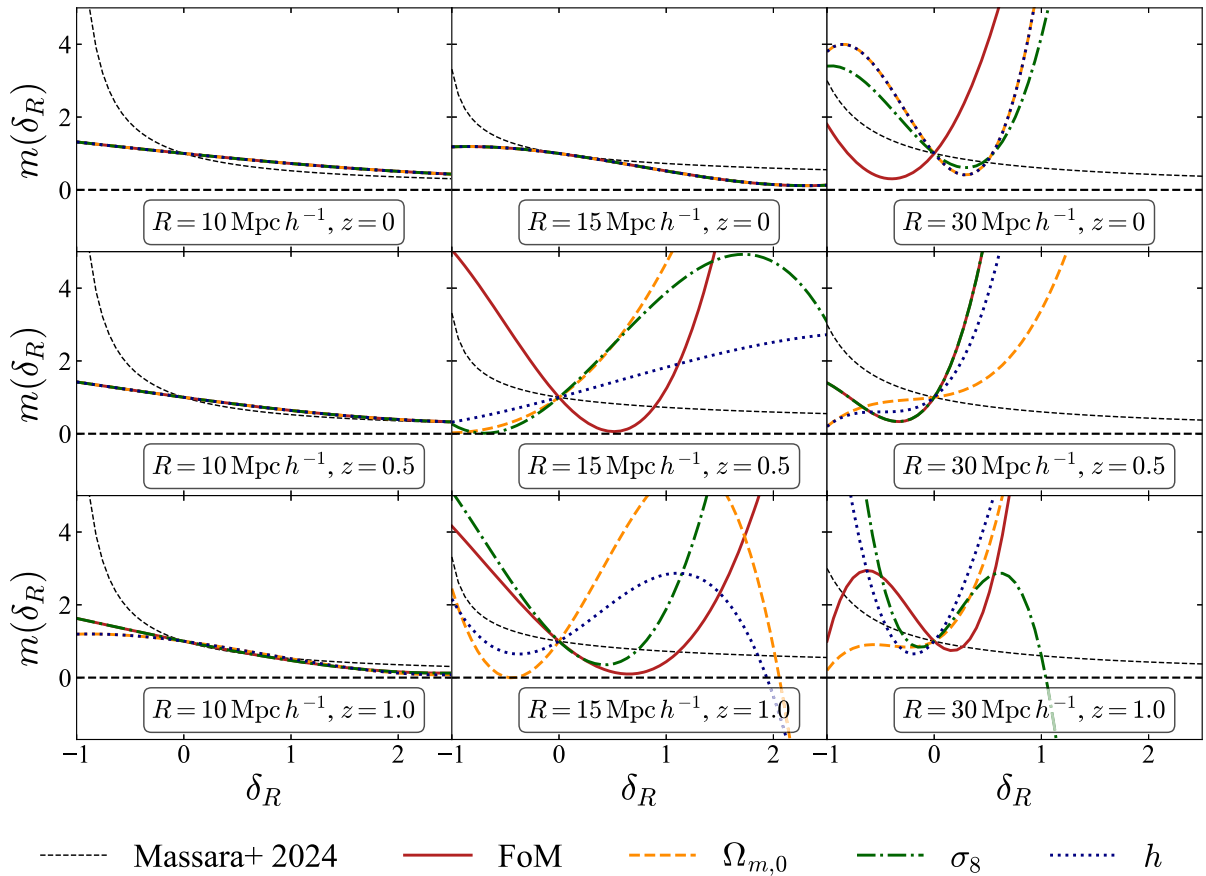


FIG. 2. Optimal mark functions for different smoothing radii and redshifts. We show the result of the optimization in different scenarios: maximizing the FoM over the three parameters $\{\Omega_{m,0}, \sigma_8, h\}$ jointly, and minimizing the uncertainty on each parameter individually. In dashed black, we also plot the mark functions adopted in the analysis [27].

rather from the environmental modulation of clustering [64,65]. This result is in agreement with [22].

This interpretation is also consistent with the findings of [26], who showed that the bispectrum combined with the standard power spectrum captures a significant fraction of the information recovered by marked statistics. The reason is straightforward: the late-time bispectrum encodes correlations between modes generated by gravitational nonlinearities, even in the absence of primordial non-Gaussianity, and is therefore a probe of environmental structure. The marked power spectrum effectively reweights two-point statistics to access part of this information in a computationally simpler way; see also the discussion in Appendix A.

The strength of the method adopted here is twofold. First, being grounded in perturbation theory, it provides a theoretically motivated and fully analytical framework that can be readily extended to additional cosmological parameters, as demonstrated in [17,18,24,26]. Second, its computational efficiency and flexibility make it a promising starting point for incorporation into a future MCMC analysis of survey data: the evaluation time for biased tracers in real space is around $t_{\text{eval}} \simeq 2.5s$ for one given cosmology and mark configuration, and $t_{\text{eval}} \simeq 6.5s$ when nonlocal primordial non-Gaussianities are included. These results can be further optimized and extended to redshift space, for which we already have an implementation that needs to be optimized. We leave this to future work.

A. Comparison with simulations

Having identified the optimal marks, we compare the theoretical predictions against power spectra measured from the Quijote simulation suite. Among all the optimal marks obtained at different smoothing scales and redshifts, we select three representative cases that display qualitatively distinct shapes, illustrating the variety of forms that the marked power spectrum can take. Figure 3 shows the matter power spectrum $P(k)$, the cross-spectrum between the marked and unmarked fields $C(k)$, and the marked power spectrum $M(k)$ for these three cases.

The marked power spectrum from each simulation box is estimated at the particle level through a two-step procedure [66]. We first assign the N -body particles to a 256^3 grid via cloud-in-cell (CIC) interpolation to build $\delta(\mathbf{x})$ [67], smooth it on the scale R to obtain δ_R , and interpolate δ_R back to each particle position. Each particle is then redeposited on the grid carrying the weight $m(\delta_R)$ evaluated at its location, which reconstructs $\rho_M(\mathbf{x})$ of Eq. (9) directly from the particles.

Across all wave numbers of interest, the analytical model reproduces the simulation measurements at the percent level, demonstrating that the EFTofLSS framework correctly describes all three spectra up to mildly nonlinear scales $k_{\text{max}} \simeq 0.30 h/\text{Mpc}$.

We next examine the covariance properties of the three spectra. The right panels of Fig. 3 show the correlation

matrices estimated from 1000 realizations at the fiducial cosmology. For $R = 10 \text{ Mpc}/h$ and $R = 30 \text{ Mpc}/h$ the correlation matrices are close to diagonal and hence well described by the Gaussian approximation, whereas for $R = 15 \text{ Mpc}/h$ the covariance displays more prominent off-diagonal structure. We show here three very different shapes for the mark function, which translate into different shapes for the power spectra.

The right panels show the correlation matrix of the joint data vector $\{P(k), C(k), M(k)\}$, estimated from the 1000 Quijote realizations. The structure of the covariance depends sensitively on the smoothing scale. For $R = 10 \text{ Mpc}/h$, the correlation matrix is close to diagonal within each block, indicating that the non-Gaussian contributions to the covariance are small. This is consistent with the fact that at this smoothing scale the optimal mark is a smooth, nearly linear function of δ_R (see Fig. 2), which limits the coupling between different Fourier modes induced by the mark. As R increases to $15 \text{ Mpc}/h$, significant off-diagonal correlations begin to appear, particularly in the M - M and M - C blocks, reflecting the more complex, nonmonotonic shape of the optimal mark at this scale and the larger non-Gaussian contributions to the density field probed. A qualitatively similar pattern is observed for $R = 30 \text{ Mpc}/h$, despite the marked power spectrum having a very different amplitude and shape compared to the $R = 15 \text{ Mpc}/h$ case.

These off-diagonal correlations, while they do not qualitatively alter the Fisher forecasts, must be properly accounted for in any rigorous likelihood analysis. In the following section we therefore present results based on the full non-Gaussian covariance estimated from simulations, and defer a detailed comparison with the Gaussian approximation to Appendix C.

B. Parameter improvements and implications

Finally, we quantify the cosmological information gain from marked statistics. Table II summarizes the forecasted improvement on the marginalized error bars of the cosmological parameters, defined as the ratio $r(\theta_\mu) \equiv \sigma_P(\theta_\mu) / \sigma_{P+C+M}(\theta_\mu)$, for the full simulation-based covariances, across three smoothing radii $R = \{10, 15, 30\} \text{ Mpc}/h$ and three redshifts $z = \{0, 0.5, 1\}$.

The results show two clear and physically motivated trends. First, the improvements are systematically larger at smaller smoothing radii: as R decreases, the mark probes increasingly nonlinear environments, capturing more of the higher-order statistical information encoded in the density field that is inaccessible to the standard power spectrum. Second, the improvement decreases with increasing redshift at fixed R , since at higher redshift the dark matter field is more linear and the nonlinear environmental information that the mark is designed to capture becomes progressively less abundant. Both trends are consistent with the expectation that the gain from marked statistics is fundamentally driven by the amount of nonlinear information available in the field.

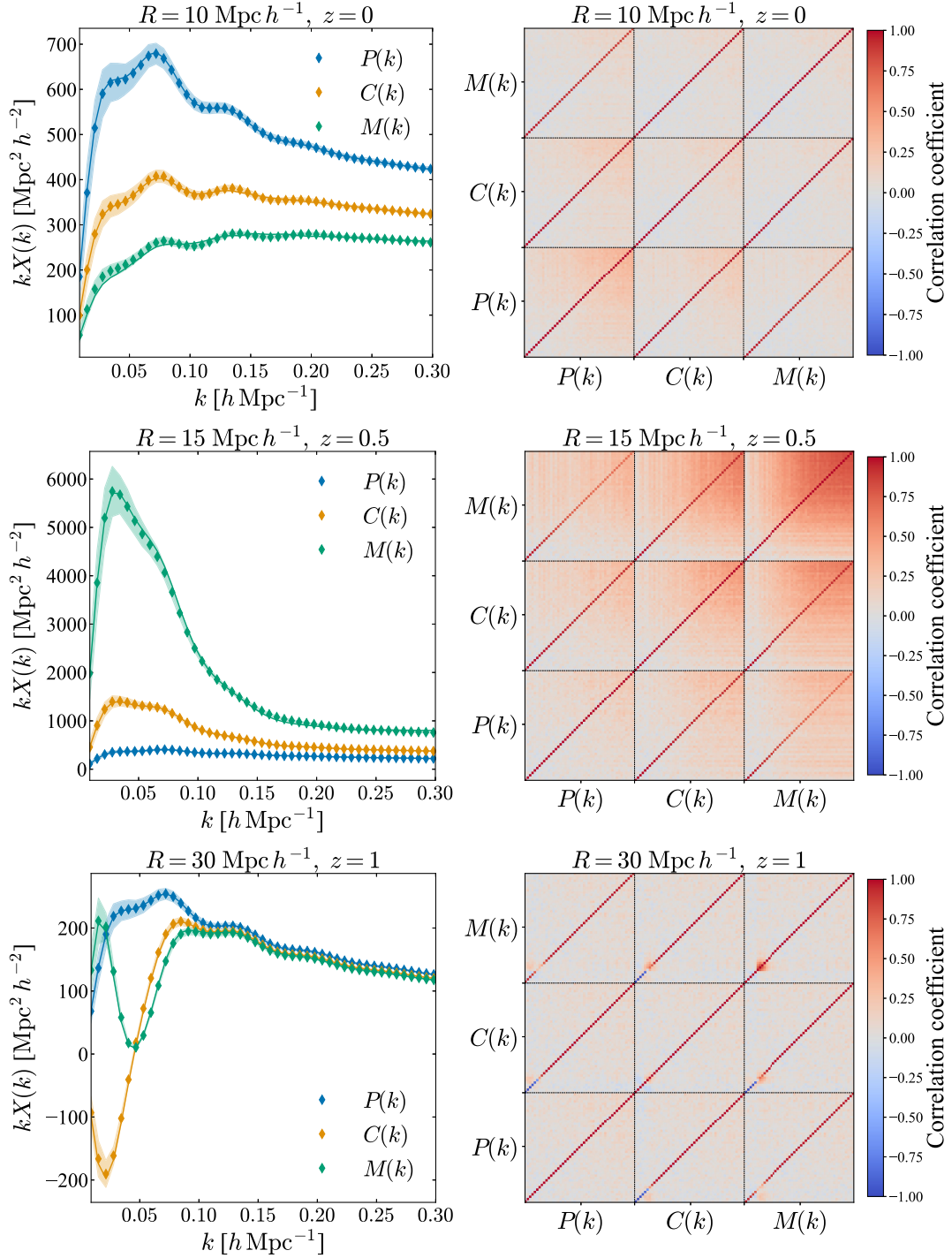


FIG. 3. Comparison of theoretical predictions with power spectra measured from the Quijote simulations for the mark with $R = 10 \text{ Mpc}/h$ at $z = 0$ optimizing over all the parameters (top), $R = 15 \text{ Mpc}/h$ at $z = 0.5$ optimizing for σ_8 (middle) and $R = 30 \text{ Mpc}/h$ at $z = 1$ optimizing the FoM (bottom). On the right columns we show the corresponding correlation matrices estimated from 1000 realizations of the Quijote suite at the fiducial cosmology.

For $R = 10 \text{ Mpc}/h$, the improvements are the largest across all configurations. At $z = 0$, the FoM-optimized mark yields $r(\sigma_8) \simeq 8.3$, $r(\Omega_m) \simeq 2.3$, and $r(h) \simeq 2.5$, representing the most significant gains found in our analysis. These large values reflect the fact that at this

smoothing scale the mark is sensitive to highly nonlinear overdensities, with $\delta_{R,\text{max}} \simeq 5.5$ at $z = 0$ (see Table I), where the density field carries substantial nonlinear information. The large improvement on σ_8 can be understood as a direct consequence of marginalizing over the EFT

TABLE II. Improvement in the cosmological parameter error bars obtained from the inclusion of the marked statistics. The improvement is quantified as the ratio $r(\theta) = \sigma_P(\theta)/\sigma_{P+C+M}(\theta)$, where σ_P denotes the marginalized uncertainty from the power spectrum alone, and σ_{P+C+M} corresponds to the joint analysis including cross- and marked-power contributions. Values larger than unity indicate a gain in constraining power relative to the standard analysis. The results shown here are based on the robust simulation-based analysis; a full comparison between Gaussian and simulation-based results is provided in the Appendix C.

Optimization of FoM			
$R = 10 \text{ Mpc} h^{-1}$			
Redshift	$r(\Omega_m)$	$r(\sigma_8)$	$r(h)$
0	2.34	8.32	2.47
0.5	1.53	3.11	1.50
1	1.43	2.06	1.38
$R = 15 \text{ Mpc} h^{-1}$			
Redshift	$r(\Omega_m)$	$r(\sigma_8)$	$r(h)$
0	1.48	5.55	1.57
0.5	1.89	2.79	1.83
1	1.46	1.97	1.32
$R = 30 \text{ Mpc} h^{-1}$			
Redshift	$r(\Omega_m)$	$r(\sigma_8)$	$r(h)$
0	1.28	1.83	1.25
0.5	1.23	1.25	1.16
1	1.48	1.72	1.32
Optimization of single parameters			
$R = 10 \text{ Mpc} h^{-1}$			
Redshift	$r(\Omega_m)$	$r(\sigma_8)$	$r(h)$
0	2.34	8.32	2.47
0.5	1.53	3.11	1.50
1	1.38	2.06	1.27
$R = 15 \text{ Mpc} h^{-1}$			
Redshift	$r(\Omega_m)$	$r(\sigma_8)$	$r(h)$
0	1.48	5.55	1.57
0.5	1.20	1.19	1.15
1	1.09	2.21	1.23
$R = 30 \text{ Mpc} h^{-1}$			
Redshift	$r(\Omega_m)$	$r(\sigma_8)$	$r(h)$
0	1.72	2.50	1.69
0.5	1.29	1.25	1.19
1	1.36	1.67	1.26

counterterm parameter c_s^2 . The correlation between these two parameters can be estimated analytically as follows. The power spectrum derivatives with respect to σ_8 and c_s^2 read

$$\frac{\partial P(k)}{\partial \sigma_8} \simeq 2\sigma_8 P_{\text{lin}}(k), \quad \frac{\partial P(k)}{\partial c_s^2} = -2\sigma_8^2 k^2 P_{\text{lin}}(k), \quad (66)$$

where in the first expression we have neglected contributions of order $\mathcal{O}(P_{\text{1loop}}/P_L)$. At leading order, the inverse covariance of the power spectrum is diagonal and given by

$$[\text{Cov}(P(k_i), P(k_j))]^{-1} = \frac{V\Delta k}{(2\pi)^2} \frac{k_i^2}{P_{\text{lin}}^2(k_i)} \delta_{ij}. \quad (67)$$

The correlation coefficient between c_s^2 and σ_8 is defined as

$$\rho_F^2(\sigma_8, c_s^2) \equiv \frac{(F_{\sigma_8 c_s^2})^2}{F_{\sigma_8 \sigma_8} F_{c_s^2 c_s^2}}, \quad (68)$$

where $F_{\alpha\beta}$ denotes the Fisher matrix element associated with parameters α and β . Substituting Eqs. (66) and (67), we obtain the compact result

$$\begin{aligned} \rho_F^2(\sigma_8, c_s^2) &= \frac{(\sum_i k_i^4)^2}{(\sum_i k_i^2)(\sum_i k_i^6)} \\ &\simeq \frac{21}{25} \left[1 + \mathcal{O}\left(\left(\frac{k_{\min}}{k_{\max}}\right)^3, c_s^2 k_{\max}^2\right) \right] \simeq 0.85, \end{aligned} \quad (69)$$

where the higher-order corrections amount to only a few percent. Remarkably, this correlation is nearly independent of $k_{\max}$. This behavior follows from the fact that the logarithmic derivative of the power spectrum with respect to σ_8 is approximately scale-independent, while the one with respect to c_s^2 exhibits a simple power-law dependence on k . The same argument does not apply to Ω_m and h , whose power spectrum derivatives encode nontrivial k -dependence through the transfer function, the BAO features, and the matter-radiation equality scale.

The correlation in Eq. (69) drives the degradation of the forecasted error on σ_8 when only the power spectrum is considered. We find $\sigma(\sigma_8)|_{P\text{-only}} \simeq 0.02$ for fixed c_s^2 , which degrades to $\sigma(\sigma_8)|_{P\text{-only}} \simeq 0.08$ once c_s^2 is marginalized over, with a degradation by a factor of ~ 4 . Considering a restricted Fisher matrix in which only σ_8 and c_s^2 are varied, we have

$$\sigma(\sigma_8)|_{c_s^2 \text{ fixed}} = (F_{\sigma_8 \sigma_8})^{-1/2}, \quad (70)$$

while the marginalized error reads

$$\sigma(\sigma_8)|_{c_s^2 \text{ marg.}} = [F_{\sigma_8 \sigma_8} (1 - \rho_F^2(\sigma_8, c_s^2))]^{-1/2} \simeq 2.5 \sigma(\sigma_8)|_{c_s^2 \text{ fixed}}, \quad (71)$$

in agreement with our numerical findings and confirming the robustness of our results. Including the cross-spectrum and the marked power spectrum in the analysis breaks this degeneracy through their distinct k -dependence. This is

consistent with the findings of [26], where the marked power spectrum was shown to alleviate degeneracies between cosmological parameters and the EFTofLSS bias and counterterm coefficients, acting in a role traditionally played by the bispectrum. This mechanism provides a useful indication of the potential of marked statistics for more realistic analyses, although its quantitative impact for biased tracers in redshift space cannot be inferred directly from the present results. In that case, the enlarged parameter space associated with galaxy bias, redshift-space counterterms, and stochastic contributions introduces additional directions that can partially absorb the scale dependence responsible for the degeneracy breaking. We therefore expect the complementarity of P , C , and M to remain relevant, but leave a quantitative assessment of the resulting information gain and its dependence on the full nuisance-parameter space to future work.

At $z = 0.5$ the improvements reduce to $r \sim 1.5$ – 3 , and at $z = 1$ further to $r \sim 1.4$ – 2 , confirming the expected suppression with redshift. We note that for $R = 10$ Mpc/ h the FoM-based and single-parameter optimizations provide identical results due to the reduced parameter space after imposing perturbativity.

For $R = 15$ Mpc/ h , the simulation-based results show a clear and interesting pattern: the improvements obtained from the FoM-optimized mark are usually larger than those from the single-parameter optimization. At $z = 0$, the FoM-optimized mark yields $r(\sigma_8) \simeq 5.6$, $r(\Omega_m) \simeq 1.5$, and $r(h) \simeq 1.6$, while the single-parameter optimization gives more modest values, $r \sim 1.1$ – 1.2 for Ω_m and h , with the exception of σ_8 which reaches $r \simeq 5.6$ in both cases. This behavior is somewhat unexpected, since one would naively expect the single-parameter optimization to provide larger improvements on each individual parameter by construction. We attribute this to the fact that the mark functions selected by the single-parameter optimization tend to have larger non-Gaussian contributions to the covariance matrix, which are not captured by the Gaussian approximation used during the optimization but are properly accounted for in the simulation-based covariance. As a result, the gains forecasted during optimization are partially washed out when the more realistic covariance is employed. The FoM-based optimization, by contrast, appears to select mark functions that are less sensitive to these non-Gaussian corrections, making it a more robust figure of merit for the purposes of mark optimization at this smoothing scale. At $z = 0.5$ the improvements are $r \sim 1.8$ – 2.8 for the FoM optimization, slightly larger than at $z = 1$ where they reduce to $r \sim 1.3$ – 2.0 , consistent with the general redshift trend discussed above.

For $R = 30$ Mpc/ h , the improvements are the most modest across all three radii, as expected given that at this smoothing scale the density field is closer to the linear regime and the mark captures less nonlinear information. At $z = 0$, the FoM-optimized mark show improvements

$r(\sigma_8) \simeq 1.8$, $r(\Omega_m) \simeq 1.3$, and $r(h) \simeq 1.3$, while the single-parameter optimization gives somewhat larger values, up to $r(\sigma_8) \simeq 2.5$ and $r(\Omega_m) \simeq 1.7$. The fact that σ_8 and Ω_m benefit more than h from the addition of marked statistics has a natural explanation in the context of our parameter set $\{\Omega_m, \sigma_8, h\}$ applied to the matter field in real space. Without a linear halo bias parameter b_1 , the amplitude of the power spectrum is directly sensitive to σ_8 , but the marked power spectrum effectively provides a weighted integral over bispectrum triangles, and this richer environmental information is particularly efficient in breaking the remaining parameter degeneracies. At $z = 0.5$ the improvements are more modest, $r \sim 1.2$ – 1.3 , while at $z = 1$ they recover slightly to $r \sim 1.3$ – 1.7 .

Taken together, these results confirm that marked statistics provide a robust and significant gain in cosmological constraining power across all smoothing scales and redshifts considered, with the largest improvements achieved at small R and low z , where the nonlinear signal is richest. These findings are broadly consistent with recent data analyses combining the one-loop power spectrum and bispectrum, such as [47], where improvements of 30%, 18%, and 13% are reported on σ_8 , h , and Ω_m , respectively. A direct comparison is not straightforward, since that analysis is performed on galaxy data in redshift space and accounts for a range of observational effects not considered here; nevertheless, the overall trends are recovered. Furthermore, recent power-spectrum-only analyses extended to two-loop order [40] report improvements of 35%, 20%, and 15% on Ω_m , h , and A_s , respectively, which translate to an improvement of $\sim 7.5\%$ on σ_8 . This suggests that extending the power spectrum analysis to higher $k_{\max}$ primarily tightens constraints on shape parameters such as Ω_m , whereas the inclusion of higher-order statistics has a proportionally larger impact on amplitude parameters such as σ_8 and A_s . The results presented here are consistent with this picture: the complementary information captured by the marked power spectrum is primarily of higher-order origin, as already argued in [26].

To gain further insight into how the mark functions modify the density field, we show in Fig. 4 the density field together with its smoothed counterpart at redshift $z = 0$, while the corresponding mark functions are displayed in Fig. 5 for the two cases $R = 10$ and 30 Mpc/ h . For the case $R = 10$ Mpc/ h , upper row of Fig. 5, our method does not appear to significantly enhance the contrast between the under- and overdense regions. This behavior becomes particularly evident when compared to the mark proposed [27], with $p = 1$ and $\delta_s = 0.1$. A plausible explanation is that our approach substantially restricts the allowed region of the parameter space $\{c_0, c_1, c_2, c_3\}$ through the perturbativity prior, which induces a much smoother mark function. Indeed, as one can see from Fig. 2, the mark obtained by optimizing the FoM closely resembles the functional form of [27], albeit with a more limited dynamic

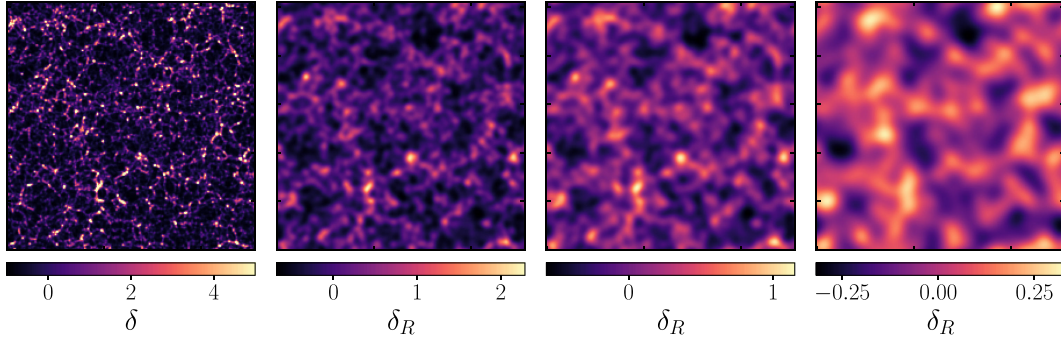


FIG. 4. Density field for a slice of one realization of the Quijote simulations at $z = 0$ (left plot). We also show the smoothed density field at the same redshift for the radii used in this work $R = 10 \text{ Mpc}/h, 15 \text{ Mpc}/h$ and $30 \text{ Mpc}/h$.

range. This represents a limitation of our approach: for small values of the smoothing radius R , the theory becomes less convergent. As a consequence of this, we find that using the mark adopted in [27] gives $\sim 2\text{--}5\%$ better constraints on the three parameters, confirming the intuition that a larger contrast between environmental density contrasts yields better cosmological constraints.

For a large $R = 30 \text{ Mpc}/h$, such perturbative limitation is no longer present: our optimal mark generates now a stronger contrast than the standard mark function with $p = 1$ and $\delta_s = 0.5$ does as shown in the lower plot of Fig. 5. These results indicate once again that, increasing the weight function for regions where there is more structure, such as voids and nodes or filaments, is optimal for parameter constraints. When compared to the usual mark functions, with parameters $p = 1$ and $\delta_s = 0.5$, we find that our procedure provides 40%–60% better constraints on the cosmological parameters, and a similar trend is observed for $R = 15 \text{ Mpc}/h$.

These results should be viewed in their proper context: we are working with the dark matter field in real space, a particularly clean setting with well-controlled theory and minimal systematics. This choice provides a controlled baseline in which the information content of marked statistics and the corresponding gains from optimization can be isolated before introducing the additional degrees of freedom associated with biased tracers and redshift-space distortions. On realistic tracers such as galaxies in redshift space, Ref. [26] showed that the information gain from $P + C + M$ is comparable to that from $P + B$. Nevertheless, all the observables considered here are two-point functions that can be measured from galaxy catalogs at minimal extra computational cost, and whose covariance matrices are far easier to estimate and model than those of the bispectrum or higher-order statistics. A complete assessment of the information gain in this more general setting, including the impact of the additional bias, counterterm, and stochastic parameters, requires extending the present framework to redshift space and galaxy tracers and is left to future work. The marked power spectrum therefore offers a valuable compromise: it captures part of the non-Gaussian

information encoded in the bispectrum while retaining the simplicity of a two-point analysis.

We note that we also investigated whether our polynomial mark framework can reproduce the optimal mark shapes found in [22] via Gaussian processes. First, we repeat our optimization varying only the two parameters Ω_m and σ_8 for $R = 30 \text{ Mpc}/h$ at $z = 0$ in order to have a direct comparison with [22]. We find the mark parameters to be $\{c_0, c_1, c_2, c_3\} = \{1.0, -1.36, -0.65, -0.65\}$ when the FoM is minimized and $\{c_0, c_1, c_2, c_3\} = \{1.0, -1.56, -0.75, 0.25\}$ and $\{c_0, c_1, c_2, c_3\} = \{1.0, -3.37, -1.46, 1.66\}$ when the error bar on Ω_m and σ_8 are minimized, respectively. When a Gaussian covariance is considered, the improvement on the parameters are in excellent agreement with the findings of [22], $r(\Omega_m) = 4.45$ and $r(\sigma_8) = 1.20$, both for FoM and single optimization. By including the full non-Gaussian covariance obtained from 1000 realizations, $r(\Omega_m) = 1.66/1.73$ and $r(\sigma_8) = 2.09/1.50$ for FoM and single optimization respectively. We observe that qualitatively these results agree with previous findings: in particular while the improvement for Ω_m is degraded by non-Gaussian contribution to the covariance, the improvement on σ_8 is enhanced. However, the slight quantitative disagreement with the improvement obtained by [22], see their Table II, can be understood in terms of the estimate of the non-Gaussian covariance: while we use $N_{\text{sims}} \gg N_{\text{bins}}$ in order to have a correct estimate of the full covariance, they used ~ 10 times less simulations to estimate it, possibly underestimating the non-Gaussian contribution [68].

Moreover, we identified combinations of $\{c_0, c_1, c_2, c_3\}$ that closely match the GP-derived mark functions of [22]; the corresponding coefficients are $\{c_0, c_1, c_2, c_3\} = \{-0.37, 2.11, 11.43, -43.20\}$ for the case with $R = 30 \text{ Mpc}/h$ at $z = 0$. This demonstrates the viability of our polynomial basis as an approximator for the GP result, and establishes a formal connection between the two fitting procedures. However, this particular coefficient set violates the perturbativity prior described in Sec. II C, and we therefore do not use it for actual parameter forecasts in this work. The two approaches are in this sense complementary:

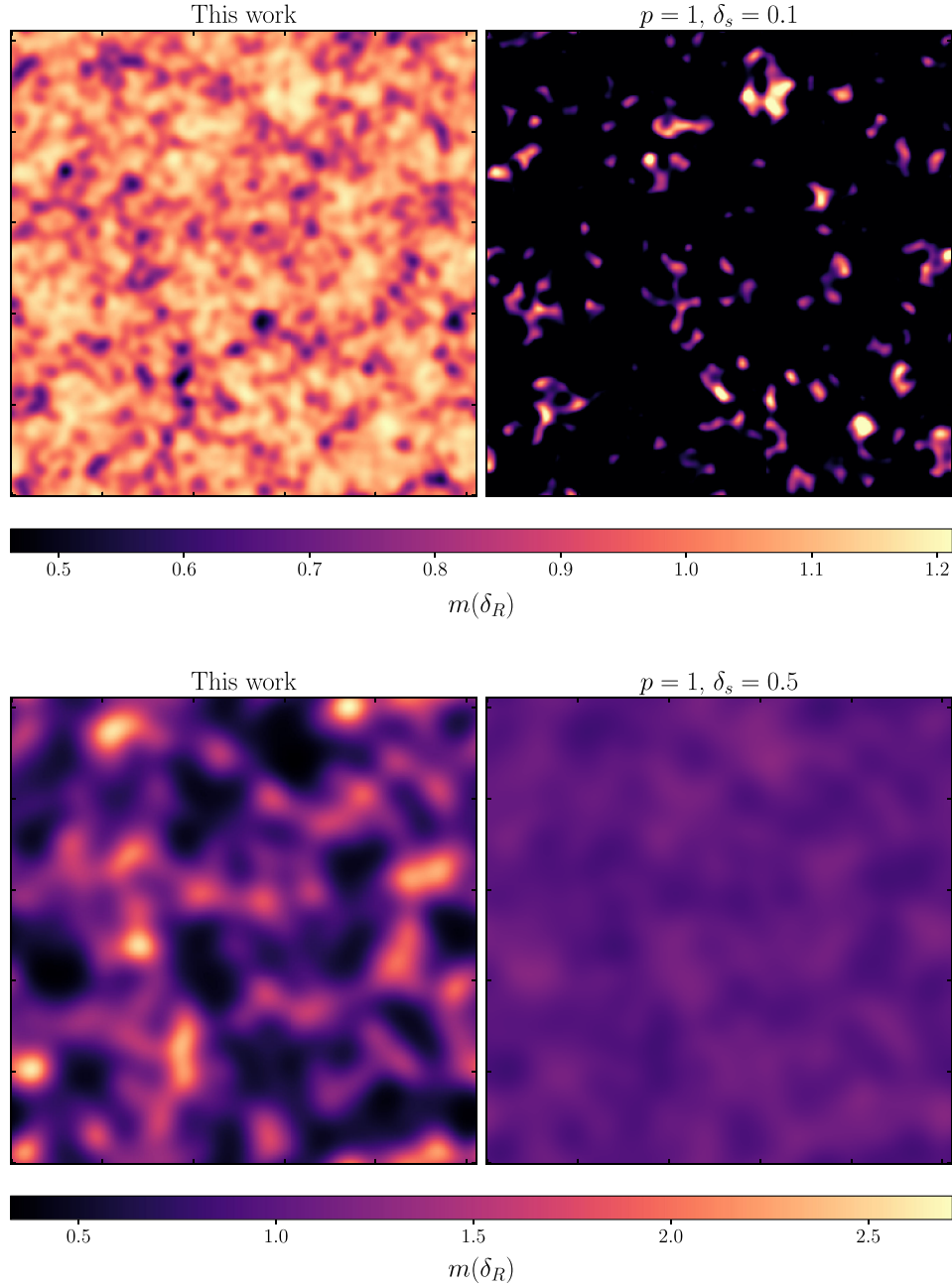


FIG. 5. Mark functions $m(\delta_R)$ obtained with our optimization scheme at $z = 0$, for $R = 10$ Mpc/ h (upper) and $R = 30$ Mpc/ h (lower). We also plot the result for the mark function from Eq. (6), with the values used in the analysis of [27].

the GP method has greater freedom to explore mark shapes beyond the perturbative regime, while ours offers full analytical control within it. It is nonetheless encouraging that, under comparable conditions (Gaussian covariance, as adopted in [22]), both coefficient sets yield similar parameter improvements, further validating the physical picture of environmental contrast as the driver of information gain.

V. OUTLOOK AND CONCLUSIONS

In this work, we have presented a fully analytical, perturbative framework to optimize the marked power

spectrum, or more precisely, the joint statistics $\{P, C, M\}$ combining the standard power spectrum, the cross-spectrum, and the marked power spectrum at mildly nonlinear scales, working to one-loop order in the EFTofLSS. The mark function is written as a polynomial in the smoothed density field δ_R , with free coefficients $\{c_1, c_2, c_3\}$ that are varied to maximize the Fisher information on any chosen set of cosmological parameters. This approach produces a polynomial fitting function for the optimal mark that is directly analogous to the Gaussian process ansatz of [22], but is

obtained entirely analytically and without any N -body simulations.

A major practical advantage of the analytical approach is computational speed. Evaluating the full Fisher matrix grid over the mark parameter space takes minutes on a standard laptop, as opposed to the large sets of N -body simulations required to train a GP or a machine-learning model. The only step in our pipeline that requires simulations is the estimation of the non-Gaussian covariance matrix at the fiducial cosmology, needed to go beyond the Gaussian approximation; this requires a single set of realizations at a fixed cosmology, rather than the many simulation runs at varied cosmology that data-driven optimization methods typically demand. This makes our pipeline a natural candidate for incorporation into a future MCMC analysis of actual survey data, where the mark coefficients could be jointly varied alongside the cosmological parameters.

A second advantage of our approach is that marked statistics help break degeneracies between cosmological parameters and the EFTofLSS nuisance parameters. As we explicitly showed in Sec. IV B, the strong correlation between σ_8 and the counterterm coefficient c_s^2 at the power spectrum level is largely broken by the inclusion of the cross-spectrum and the marked power spectrum, thanks to their distinct k -dependence. This is consistent with the findings of [24,26], where marked statistics were shown to alleviate degeneracies among cosmological and EFT parameters, a role traditionally played by the bispectrum.

Finally, a distinctive feature of our method is the necessity of imposing perturbativity priors on the mark coefficients. As discussed in Sec. II C, mass conservation is not a symmetry of the marked field, so higher-loop contributions can enter at large scales already at one-loop rather than being suppressed by powers of k . Requiring that these contributions remain subdominant places nontrivial constraints on $\{c_1, c_2, c_3\}$, restricting the optimization to a theoretically controlled region of parameter space. While this boundary limits the accessible fitting space compared to a Gaussian process (GP) with no theoretical prior (indeed, the GP-optimal coefficients of [22] lie outside our perturbative region) we nonetheless recover information gains consistent with those of [22] when using the same analysis specifications. The two approaches are therefore complementary: the GP method has more freedom to explore aggressive mark shapes, while ours provides full analytical control and a clear criterion for when the perturbative description is trustworthy.

We conclude with some remarks on caveats and future directions. A conceptually important result emerges from this optimization procedure. As first pointed out in [22], when the full data vector $\{P, C, M\}$ is used, the Fisher matrix is invariant under affine transformations of the mark, $m \rightarrow Am + B$, in the mildly nonlinear regime where the Gaussian likelihood approximation is valid. We apply this property here following [22]; we note that at highly

nonlinear scales, where the likelihood can develop significant non-Gaussianity, this invariance may no longer hold exactly. However, within our perturbative regime, it has a clear and important consequence: void-enhancing and overdensity-enhancing marks are equivalent in information content, and the optimizer has no reason to prefer one over the other. What it consistently selects for, across all parameters, smoothing scales, and redshifts considered, is a strong contrast between under- and overdense regions, this being the only degree of freedom not removed by the invariance under affine symmetry of the Fisher matrix. The information gain from the marked power spectrum therefore does not reside in the special properties of voids *per se*, but in the environmental modulation of clustering captured by the mark. We stress that this conclusion is specific to the entire $\{P, C, M\}$ data vector; any analysis based on higher order functions such as the marked bispectrum (without including all the cross spectra) would not enjoy the same affine invariance, and in that case, the nature of the up-weighted regions would matter.

An important caveat is that the present analysis is restricted to the dark matter field in real space. Extending to biased tracers in redshift space is a necessary step before application to data and is currently a work in progress; preliminary results are encouraging.

Our framework also makes it straightforward to extend the analysis to cosmological parameters beyond vanilla Λ CDM, without requiring a dedicated simulation suite for each new parameter. Two extensions we are actively pursuing are the following. There is already a significant body of work showing that void-enhancing marked power spectra are particularly sensitive to neutrino masses [19,20,26], consistent with the physical expectation that neutrino free-streaming suppresses power preferentially in underdense regions. Within our framework, it will be interesting to derive the optimal $\{P, C, M\}$ combination for neutrino masses analytically, both to cross-check existing results and to ask whether the optimal mark for neutrinos differs qualitatively from the standard void-enhancing ansatz. We previously studied the performance of the marked power spectrum for primordial non-Gaussianities f_{NL} using the standard mark ansatz in [26]. That mark was originally designed with neutrinos and modified gravity in mind (see also [21]), and it is far from obvious that it is optimal for f_{NL} . The shapes of primordial non-Gaussianity couple very different configurations of the bispectrum (squeezed for local f_{NL} , equilateral or folded for other shapes), and the optimal environmental weighting may be qualitatively different. Applying our optimization pipeline to f_{NL} for different primordial shapes is therefore a genuinely open question that we plan to address in forthcoming work.

More broadly, the marked power spectrum can be understood as a practical compression scheme for information from higher order functions: at one-loop, it provides

a weighted integral over bispectrum triangle configurations, where the weight function is determined by the mark. On mildly nonlinear scales, where the bispectrum dominates the non-Gaussian signal, an optimized marked power spectrum can, in principle, perform comparably to a full $P + B$ analysis [25,26], while being considerably cheaper to compute and model. At more nonlinear scales, the mark also begins to compress information from higher-order correlators, while still retaining the simplicity of a two-point function. This is the deeper motivation for our interest in marked statistics as an efficient and practical summary for nonlinear LSS, and it underlies all the extensions described above. At those more nonlinear scales, however, a purely perturbative treatment is no longer sufficient, and one must rely on simulations and/or machine-learning emulators, as it has been done for f_{NL} in [69].

ACKNOWLEDGMENTS

The authors thank Massimo Pietroni and Francesco Spezzati for useful discussions. M. L. and M. M. acknowledge support by the MIUR Progetti di Ricerca di Rilevante Interesse Nazionale (PRIN) Bando 2022—Grant No. 20228RMX4A. F. S. is partly supported by ICSC—Centro Nazionale di Ricerca in High Performance Computing, Big Data and Quantum Computing, funded by European Union—NextGenerationEU.

DATA AVAILABILITY

There are no publicly available research data or software supporting this manuscript. Requests for further information or data should be sent to the authors.

APPENDIX A: VALIDITY OF PERTURBATION THEORY FOR MARKED FIELDS

In this appendix we further investigate the accuracy of the perturbative approach adopted in this work for the description of marked fields. The mark function we use in this work is

$$m(\delta_R) = c_0 + c_1\delta_R + c_2\delta_R^2 + c_3\delta_R^3. \quad (\text{A1})$$

One may wonder if the arbitrariness of the c_i 's coefficient may invalidate the assumption that the new weighted field is perturbative. Intuitively, the answer to this question depends on the underlying theoretical framework. In the perturbative context of the EFTofLSS, the coefficients are required to satisfy constraints dictated by perturbativity and by the convergence of the theory. In particular, we already know that some one-loop terms can give non-negligible contributions on very large scales, as already pointed out in [17,18,24]. This is due to the fact that mass conservation is not a symmetry of the marked field, even for dark matter, and the usual behavior $\sim k^2$ as $k \rightarrow 0$ is not guaranteed anymore, similarly to what happens for standard dark

matter tracers [70–72]. In addition, higher loop terms affecting large scales is a distinct feature of marked correlators: we are bringing information from nonlinear to linear scales, enhancing the effect of short modes on long ones. This is where the information gain resides, as already stated in previous works [17,18,24,26]. Even though this looks appealing, we need to require that the theory is still under perturbative control: since we stop at the one-loop order, we need to impose that the contributions from higher loops is negligible compared to the lower ones. This idea was previously introduced in the context of the EFTofLSS in [34], see also [35]. Its purpose is to restrict the parameter volume to regions of the bias-parameter space where higher-loop contributions remain under theoretical control. In the context of marked correlators, this kind of prior becomes even more important as we will see, since higher loops contribute non-negligibly already at large scales.

1. Perturbativity prior from scaling arguments

It is useful to evaluate the large scale limit of the one-loop contributions and compare them with the linear terms of the marked power spectrum. We know that the marked power spectrum at one loop at large scales is dominated by two pieces

$$M_{22}(k) = 2 \int_{\mathbf{q}} [H_2(\mathbf{k} - \mathbf{q}, \mathbf{q})]^2 P_{\text{lin}}(q) P_{\text{lin}}(|\mathbf{k} - \mathbf{q}|), \quad (\text{A2})$$

and

$$M_{13}(k) = 6C_{\delta_M}(k)P_{\text{lin}}(k) \int_{\mathbf{q}} H_3(\mathbf{k}, \mathbf{q}, -\mathbf{q})P_{\text{lin}}(q). \quad (\text{A3})$$

Let us consider M_{13} first: its ratio with the linear contribution M_{11} is given by

$$\begin{aligned} \frac{M_{13}(k)}{M_{11}(k)} &= 6 \int_{\mathbf{q}} F_3(\mathbf{k}, \mathbf{q}, -\mathbf{q})P_{\text{lin}}(q) \\ &\quad + \frac{8}{C_{\delta_M}(k)} \int_{\mathbf{q}} C_{\delta_M^2}(q, |\mathbf{k} - \mathbf{q}|)F_2(\mathbf{k}, -\mathbf{q})P_{\text{lin}}(q) \\ &\quad + \frac{6}{C_{\delta_M}(k)} \int_{\mathbf{q}} C_{\delta_M^3}(k, q, q)P_{\text{lin}}(q). \end{aligned} \quad (\text{A4})$$

Recall that the weighting functions of order n contain, roughly, n window functions W_R , explicitly $C_{\delta_M^n}(k) \sim W_R^n(k)$. These terms are under perturbative control: the first piece in Eq. (A4) is the same as in the standard power spectrum, and it behaves as $\sim \sqrt{\Delta^2(k)}$, while the other two pieces behave as $\sim \sqrt{\Delta_M^2(k)}$. For the M_{22} piece, the story is different as it is the one responsible for the shot-noiselike terms that can be dominant already at large scales, where perturbation theory is expected to be valid. It reads

$$\begin{aligned} \frac{M_{22}(k)}{M_{11}(k)} &= \frac{2 \int_{\mathbf{q}} [F_2(\mathbf{k} - \mathbf{q}, \mathbf{q})]^2 P_{\text{lin}}(q) P_{\text{lin}}(|\mathbf{k} - \mathbf{q}|)}{P_{\text{lin}}(k)} + \frac{4 \int_{\mathbf{q}} C_{\delta_M^2}(q, |\mathbf{k} - \mathbf{q}|) F_2(\mathbf{k} - \mathbf{q}, \mathbf{q}) P_{\text{lin}}(|\mathbf{k} - \mathbf{q}|) P_{\text{lin}}(q)}{C_{\delta_M}(k) P_{\text{lin}}(k)} \\ &+ \frac{2 \int_{\mathbf{q}} C_{\delta_M^2}^2(q, |\mathbf{k} - \mathbf{q}|) P_{\text{lin}}(q) P_{\text{lin}}(|\mathbf{k} - \mathbf{q}|)}{C_{\delta_M^2}^2(k) P_{\text{lin}}(k)}. \end{aligned} \quad (\text{A5})$$

The first and the second term are safe, as they behave as $\sim \sqrt{\Delta^2(k)}$ and $\sim \sqrt{\Delta_M^2(k)}$ respectively. The last term in the second line of Eq. (A5) is the more problematic: if we take the large scale limit of such term we obtain

$$\lim_{k \rightarrow 0} \frac{2 \int_{\mathbf{q}} C_{\delta_M^2}^2(q, |\mathbf{k} - \mathbf{q}|) P_{\text{lin}}(q) P_{\text{lin}}(|\mathbf{k} - \mathbf{q}|)}{C_{\delta_M^2}^2(k) P_{\text{lin}}(k)} = \frac{2 \int_{\mathbf{q}} C_{\delta_M^2}^2(q, q) P_{\text{lin}}^2(q)}{(c_0 + c_1)^2 P_{\text{lin}}(k)}, \quad (\text{A6})$$

which becomes problematic since the numerator is constant and the denominator goes to zero as $\lim_{k \rightarrow 0} P_{\text{lin}}(k) \sim k^{n_s}$.

Crucially, if we open up the expression for this limit we can see that it depends on the value of c_2

$$\lim_{k \rightarrow 0} \frac{2 \int_{\mathbf{q}} C_{\delta_M^2}^2(q, |\mathbf{k} - \mathbf{q}|) P_{\text{lin}}(q) P_{\text{lin}}(|\mathbf{k} - \mathbf{q}|)}{C_{\delta_M^2}^2(k) P_{\text{lin}}(k)} = \frac{2(c_2^2 \int_{\mathbf{q}} W_R^4(q) P_{\text{lin}}^2(q) + c_1^2 \int_{\mathbf{q}} W_R^2(q) P_{\text{lin}}^2(q) + 2c_2 c_1 \int_{\mathbf{q}} W_R^3(q) P_{\text{lin}}^2(q))}{(c_0 + c_1)^2 P_{\text{lin}}}, \quad (\text{A7})$$

where the integrals evaluate to

$$\begin{aligned} \int_{\mathbf{q}} W_R^2(q) P_{\text{lin}}^2(q) &\simeq 275, & \int_{\mathbf{q}} W_R^3(q) P_{\text{lin}}^2(q) &\simeq 181, \\ \int_{\mathbf{q}} W_R^4(q) P_{\text{lin}}^2(q) &\simeq 132 \end{aligned} \quad (\text{A8})$$

for $R = 30 \text{ Mpc}/h$ and $z = 0$, and

$$\begin{aligned} \int_{\mathbf{q}} W_R^2(q) P_{\text{lin}}^2(q) &\simeq 863, & \int_{\mathbf{q}} W_R^3(q) P_{\text{lin}}^2(q) &\simeq 643, \\ \int_{\mathbf{q}} W_R^4(q) P_{\text{lin}}^2(q) &\simeq 512 \end{aligned} \quad (\text{A9})$$

for $R = 15 \text{ Mpc}/h$ and $z = 0$.

The results of these calculations are shown in the main text in Fig. 1.

2. Higher-loop contributions for marked power spectra

We start from the expression for the marked overdensity field

$$\delta_M(\mathbf{x}) = \frac{m(\delta_R(\mathbf{x}))}{\bar{m}} [1 + \delta(\mathbf{x})] - 1. \quad (\text{A10})$$

At this stage, the expression is kept fully nonperturbative, since it corresponds to the quantity provided to simulations (or observational data) for the computation of the two-point correlation function of the marked field. We first evaluate the marked power spectrum in its full nonperturbative form, as obtained from simulations, and then compare it to the perturbative prediction. In Fourier space this becomes

$$\begin{aligned} \delta_M(\mathbf{k}) &= \frac{1}{\bar{m}} \left\{ c_0 \delta(\mathbf{k}) + c_1 \delta_R(\mathbf{k}) + \mathcal{I}_{\mathbf{k}; \mathbf{q}_1, \mathbf{q}_2} \left[c_2 \delta_R(\mathbf{q}_1) \delta_R(\mathbf{q}_2) \right. \right. \\ &\quad \left. \left. + \frac{c_1}{2} (\delta_R(\mathbf{q}_1) \delta(\mathbf{q}_2) + \delta(\mathbf{q}_1) \delta_R(\mathbf{q}_2)) \right] \right. \\ &\quad \left. + \mathcal{I}_{\mathbf{k}; \mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3} \left[c_3 \delta_R(\mathbf{q}_1) \delta_R(\mathbf{q}_2) \delta_R(\mathbf{q}_3) \right. \right. \\ &\quad \left. \left. + \frac{c_2}{3} (\delta_R(\mathbf{q}_1) \delta_R(\mathbf{q}_2) \delta(\mathbf{q}_3) + 2 \text{ perms.}) \right] \right. \\ &\quad \left. + \mathcal{I}_{\mathbf{k}; \mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3} c_3 \delta_R(\mathbf{q}_1) \delta_R(\mathbf{q}_2) \delta_R(\mathbf{q}_3) \delta(\mathbf{q}_4) \right\}, \end{aligned} \quad (\text{A11})$$

where we have used the notation

$$\mathcal{I}_{\mathbf{k}; \mathbf{q}_1, \dots, \mathbf{q}_n} \equiv \int \frac{d^3 \mathbf{q}_1}{(2\pi)^3} \dots \int \frac{d^3 \mathbf{q}_n}{(2\pi)^3} (2\pi)^3 \delta_D(\mathbf{k} - \mathbf{q}_{1\dots n}). \quad (\text{A12})$$

We emphasize that this expression is precisely the one employed in the estimation of the power spectrum.

$$\bar{m}^2 \langle \delta_M(\mathbf{k}) \delta_M(\mathbf{k}') \rangle = (2\pi)^2 \delta_D(\mathbf{k} + \mathbf{k}') M(k), \quad (\text{A13})$$

and this expression is not perturbative, meaning that the density fields are the nonlinear ones. The expression for the full marked power spectrum presents many different pieces: we start with the one involving only one field [the first line of Eq. (A11)]

$$\bar{m}^2 \langle \delta_M^{[1]}(\mathbf{k}) \delta_M^{[1]}(\mathbf{k}') \rangle' = (c_0 + c_1 W_R(\mathbf{k}))^2 P_{\text{NL}}(k) \quad (\text{A14})$$

where $\delta_M^{[1]}$ indicates the terms with only one field in Eq. (A11), and $\langle \dots \rangle'$ is the correlator where the $(2\pi)^3 \delta_D(\mathbf{k} + \mathbf{k}')$ has been removed. P_{NL} represents the complete nonlinear power spectrum for the dark matter density perturbations. In our perturbative perspective, P_{NL} would simply be the EFTofLSS prediction up to some order, 1-loop in this work. We could already argue that this piece, in a perturbative framework and at the scales considered in this work, should be the dominant one, as it represents the leading order. At large scales this piece behaves as $\sim (c_0 - c_1)^2 P_{\text{NL}}(k)$, with $P_{\text{NL}}(k) \simeq P_{\text{lin}}(k) \rightarrow k^{n_s}$ for $k \rightarrow 0$. Imposing that this piece has to be the dominant one will turn out to be a constrain on the possible values for c_2 and c_3 compared to c_1 as we will see below. Moving on to contributions to the marked power spectrum that involve on and two fields we have

$$\begin{aligned} \bar{m}^2 \langle \delta_M^{[1]}(\mathbf{k}) \delta_M^{[2]}(\mathbf{k}') \rangle \\ = C_{\delta_M}(k) \int \frac{d^3 \mathbf{q}}{(2\pi)^3} C_{\delta_M^2}(q, |\mathbf{k} + \mathbf{q}|) B_{\text{NL}}(\mathbf{k}, \mathbf{q}, -\mathbf{k} - \mathbf{q}), \end{aligned} \quad (\text{A15})$$

and

$$\begin{aligned} \bar{m}^2 \langle \delta_M^{[1]}(\mathbf{k}') \delta_M^{[2]}(\mathbf{k}) \rangle \\ = C_{\delta_M}(k') \int \frac{d^3 \mathbf{q}}{(2\pi)^3} C_{\delta_M^2}(q, |\mathbf{k} - \mathbf{q}|) B_{\text{NL}}(-\mathbf{k}, \mathbf{q}, \mathbf{k} - \mathbf{q}). \end{aligned} \quad (\text{A16})$$

This interesting exercise show explicitly that the two-point correlation function of the marked field contains pieces of higher order correlators, such as the (convolved) three-point correlation function, as shown in the above expressions. Considering the configurations with four fields $([2] - [2])$ and $([1] - [3])$ we obtain

$$\begin{aligned} \bar{m}^2 \langle \delta_M^{[2]}(\mathbf{k}) \delta_M^{[2]}(\mathbf{k}') \rangle &= \mathcal{I}_{\mathbf{k}; \mathbf{q}_1, \mathbf{q}_2} \mathcal{I}_{\mathbf{k}'; \mathbf{p}_1, \mathbf{p}_2} C_{\delta_M^2}(q_1, q_2) C_{\delta_M^2} \\ &\times (p_1, p_2) \langle \delta(\mathbf{q}_1) \delta(\mathbf{q}_2) \delta(\mathbf{p}_1) \delta(\mathbf{p}_2) \rangle, \end{aligned} \quad (\text{A17})$$

where the four-point function has a disconnected part

$$\begin{aligned} \langle \delta(\mathbf{q}_1) \delta(\mathbf{q}_2) \delta(\mathbf{q}_3) \delta(\mathbf{q}_4) \rangle_{\text{dc}} \\ = 2(2\pi)^6 \delta_D(\mathbf{q}_{13}) \delta_D(\mathbf{q}_{24}) P_{\text{NL}}(q_1) P_{\text{NL}}(q_2), \end{aligned} \quad (\text{A18})$$

where we have already neglected the term that gives a disconnected contribution in Eq. (A17), and a connected piece

$$\begin{aligned} \langle \delta(\mathbf{q}_1) \delta(\mathbf{q}_2) \delta(\mathbf{q}_3) \delta(\mathbf{q}_4) \rangle_{\text{c}} \\ = (2\pi)^3 \delta_D(\mathbf{q}_{1234}) T_{\text{NL}}(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3, \mathbf{q}_4). \end{aligned} \quad (\text{A19})$$

These two give

$$\begin{aligned} \bar{m}^2 \langle \delta_M^{[2]}(\mathbf{k}) \delta_M^{[2]}(\mathbf{k}') \rangle &= (2\pi)^3 \delta_D(\mathbf{k} + \mathbf{k}') \left[2 \int \frac{d^3 \mathbf{q}}{(2\pi)^3} C_{\delta_M^2}^2(q, |\mathbf{k} - \mathbf{q}|) P_{\text{NL}}(q) P_{\text{NL}}(|\mathbf{k} - \mathbf{q}|) \right. \\ &\quad \left. + \int \frac{d^3 \mathbf{q}_1}{(2\pi)^3} \frac{d^3 \mathbf{q}_2}{(2\pi)^3} C_{\delta_M^2}(q_1, |\mathbf{k} - \mathbf{q}_1|) C_{\delta_M^2}(q_2, |\mathbf{k} + \mathbf{q}_2|) T_{\text{NL}}(\mathbf{q}_1, \mathbf{k} - \mathbf{q}_1, \mathbf{q}_2, -\mathbf{k} - \mathbf{q}_2) \right]. \end{aligned} \quad (\text{A20})$$

The other piece with four fields is given by

$$\begin{aligned} \bar{m}^2 (\langle \delta_M^{[1]}(\mathbf{k}) \delta_M^{[3]}(\mathbf{k}') \rangle + \langle \delta_M^{[3]}(\mathbf{k}) \delta_M^{[1]}(\mathbf{k}') \rangle) &= (2\pi)^3 \delta_D(\mathbf{k} + \mathbf{k}') \left[6 C_{\delta_M}(k) P_{\text{NL}}(k) \int \frac{d^3 \mathbf{q}}{(2\pi)^3} C_{\delta_M^3}(k, q, q) P_{\text{NL}}(q) \right. \\ &\quad + C_{\delta_M}(k) \int \frac{d^3 \mathbf{q}_1}{(2\pi)^3} \frac{d^3 \mathbf{q}_2}{(2\pi)^3} (C_{\delta_M^3}(q_1, q_2, |\mathbf{k} + \mathbf{q}_{12}|) T_{\text{NL}}(\mathbf{k}, \mathbf{q}_1, \mathbf{q}_2, -\mathbf{k} - \mathbf{q}_{12}) \\ &\quad \left. + C_{\delta_M^3}(q_1, q_2, |\mathbf{k} - \mathbf{q}_{12}|) T_{\text{NL}}(-\mathbf{k}, \mathbf{q}_1, \mathbf{q}_2, \mathbf{k} - \mathbf{q}_{12})) \right]. \end{aligned} \quad (\text{A21})$$

Note that the terms in the first line of Eq. (A20) are those entering our 1-loop expression for the marked power spectrum.

In what follows we focus on higher-loop contributions that can contaminate the large scale limit. The leading terms of this type arise at two-loop order and originate from the two-loop power spectrum, the one-loop bispectrum, and the tree-level trispectrum. These represent the lowest-order higher-point contributions that can propagate to large scales and therefore potentially spoil the expected large scale behavior. Higher-order correlators, or higher perturbative orders of the same correlators, involve an increasing number of factors of the linear power spectrum. In particular, such terms typically contain at least four powers of the linear spectrum. Their contribution is therefore parametrically suppressed once the conditions that we impose below are enforced. These conditions effectively ensure that the perturbative expansion remains well behaved and that higher-order terms do not generate sizeable contributions at large scales. For this reason, it is sufficient for our purposes to restrict the analysis to the terms listed above, which capture the dominant sources of higher-loop contamination within the perturbative regime. Contributions from higher-order correlators or higher-loop corrections to the same correlators are expected to be subleading under the assumptions adopted here.

The first term we compute is the one involving three fields:

$$C_{\delta_M}(k) \int \frac{d^3 \mathbf{q}}{(2\pi)^3} C_{\delta_M^2}(q, |\mathbf{k} - \mathbf{q}|) B_{\text{NL}}(-\mathbf{k}, \mathbf{q}, \mathbf{k} - \mathbf{q}). \quad (\text{A22})$$

In the EFTofLSS, the nonlinear bispectrum up to one-loop is given by

$$B_{\text{NL}}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = B_{\text{tree}}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) + B_{\text{1loop}}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3), \quad (\text{A23})$$

where

$$B_{\text{tree}}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = 2F_2(\mathbf{k}_1, \mathbf{k}_2)P_L(k_1)P_L(k_2) + (2 \text{ perms.}), \quad (\text{A24})$$

and the one-loop part is given by

$$B_{\text{1loop}}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = B_{222}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) + B_{321}^{(I)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) + B_{321}^{(II)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{q}_3) + B_{411}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3), \quad (\text{A25})$$

with

$$\begin{aligned} B_{222}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) &= 8 \int_{\mathbf{q}} F_2(-\mathbf{q}, \mathbf{k}_1 + \mathbf{q}) F_2(\mathbf{k}_1 + \mathbf{q}, \mathbf{k}_2 - \mathbf{q}) F_2(\mathbf{k}_2 - \mathbf{q}, \mathbf{q}) P_L(q) P_L(|\mathbf{k}_1 + \mathbf{q}|) P_L(|\mathbf{k}_2 - \mathbf{q}|), \\ B_{321}^{(I)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) &= 6P_L(k_1) \int_{\mathbf{q}} F_3(-\mathbf{q}, -\mathbf{k}_2 + \mathbf{q}, -\mathbf{k}_1) F_2(\mathbf{q}, \mathbf{k}_2 - \mathbf{q}) P_L(q) P_L(|\mathbf{k}_2 - \mathbf{q}|) + 5 \text{ perms.}, \\ B_{321}^{(II)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) &= 6P_L(k_1) P_L(k_2) F_2(\mathbf{k}_1, \mathbf{k}_2) \int_{\mathbf{q}} F_3(\mathbf{k}_1, \mathbf{q}, -\mathbf{q}) P_L(q) + 5 \text{ perms.}, \\ B_{411}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) &= 12P_L(k_1) P_L(k_2) \int_{\mathbf{q}} F_4(\mathbf{q}, -\mathbf{q}, -\mathbf{k}_1, -\mathbf{k}_2) P_L(q) + 2 \text{ perms.} \end{aligned} \quad (\text{A26})$$

The tree-level contribution is already contained in the one-loop expression for the marked power spectrum and therefore does not represent an additional independent contribution.

The large scale limit ($k \rightarrow 0$) of the one-loop 321(I) term is

$$\begin{aligned} \lim_{k \rightarrow 0} C_{\delta_M}(k) \int_{\mathbf{q}} C_{\delta_M^2}(q, |\mathbf{k} - \mathbf{q}|) B_{321}^{(I)}(-\mathbf{k}, \mathbf{q}, \mathbf{k} - \mathbf{q}) &= 6(c_0 + c_1) \int_{\mathbf{q}} (c_2 W_R(q)^2 + c_1 W_R(q)) P_L^2(q) \\ \int_{\mathbf{p}} F_3(-\mathbf{p}, \mathbf{q} + \mathbf{p}, -\mathbf{q}) F_2(\mathbf{p}, -\mathbf{q} - \mathbf{p}) P_L(|\mathbf{q} + \mathbf{p}|) &\equiv 6(c_0 + c_1)(c_2 \mathcal{K}_1 + c_1 \mathcal{K}_2) = \text{const.} \end{aligned}$$

For the term with $B_{321}^{(II)}$, in the $k \rightarrow 0$ limit, we obtain

$$\begin{aligned} \lim_{k \rightarrow 0} C_{\delta_M}(k) \int_{\mathbf{q}} C_{\delta_M^2}(q, |\mathbf{k} - \mathbf{q}|) B_{321}^{(II)}(-\mathbf{k}, \mathbf{q}, \mathbf{k} - \mathbf{q}) &= 6C_{\delta_M}(k) \int_{\mathbf{q}} C_{\delta_M^2}(q, q) \\ &\times \left\{ P_L(k) P_L(q) F_2(\mathbf{q}, 0) \int_{\mathbf{p}} F_3(\mathbf{q}, -\mathbf{p}, \mathbf{p}) P_L(p) + P_L(k) P_L(q) F_2(\mathbf{q}, 0) \int_{\mathbf{p}} F_3(0, \mathbf{p}, -\mathbf{p}) P_L(p) \right. \\ &\left. + P_L(k) P_L(q) F_2(0, -\mathbf{q}) \int_{\mathbf{p}} F_3(-\mathbf{q}, \mathbf{p}, -\mathbf{p}) P_L(p) + P_L(k) P_L(q) F_2(0, -\mathbf{q}) \int_{\mathbf{p}} F_3(\mathbf{p}, \mathbf{p}, -\mathbf{p}) P_L(p) \right\}. \end{aligned} \quad (\text{A27})$$

This result shows that this piece goes like $\sim C \times P_L(k)$ which safely goes to zero for $k \rightarrow 0$. The B_{222} part gives results similar to $B_{321}^{(I)}$

$$\lim_{k \rightarrow 0} C_{\delta_M}(k) \int_{\mathbf{q}} C_{\delta_M}^2(q, |\mathbf{k} - \mathbf{q}|) B_{222}(-\mathbf{k}, \mathbf{q}, \mathbf{k} - \mathbf{q}) \simeq 8(c_0 + c_1)(c_2 \mathcal{K}_3 + c_1 \mathcal{K}_4). \quad (\text{A28})$$

These results show how higher order terms can give a non-negligible contribution already at large scales, depending on the mark parameters $\{c_0, c_1, c_2\}$.

Let us consider now four fields

$$\bar{m}^2 \langle \delta_M^{[2]}(\mathbf{k}) \delta_M^{[2]}(\mathbf{k}') \rangle' = 2 \int_{\mathbf{q}} C_{\delta_M}^2(q, |\mathbf{k} - \mathbf{q}|) P_{\text{NL}}(q) P_{\text{NL}}(|\mathbf{k} - \mathbf{q}|) \quad (\text{A29})$$

$$+ \int_{\mathbf{q}_1, \mathbf{q}_2} C_{\delta_M}^2(q_1, |\mathbf{k} - \mathbf{q}_1|) C_{\delta_M}^2(q_2, |\mathbf{k} + \mathbf{q}_2|) T_{\text{NL}}(\mathbf{q}_1, \mathbf{k} - \mathbf{q}_1, \mathbf{q}_2, -\mathbf{k} - \mathbf{q}_2). \quad (\text{A30})$$

In the first line, the one involving the nonlinear power spectrum we expect, at the order we are interested in, contributions up to one-loop, meaning $P_L P_L$ and $P_L P_{1\text{loop}}$. The first contribution is simply the one already obtained with the one-loop calculation of the marked power spectrum

$$2 \int_{\mathbf{q}} C_{\delta_M}^2(q, |\mathbf{k} - \mathbf{q}|) P_L(q) P_L(|\mathbf{k} - \mathbf{q}|) \rightarrow 2 \int_{\mathbf{q}} C_{\delta_M}^2(q, q) P_L^2(q) \sim \text{const.} \quad (\text{A31})$$

The $P_L P_{1\text{loop}}$ terms give

$$2 \int_{\mathbf{q}} C_{\delta_M}^2(q, |\mathbf{k} - \mathbf{q}|) \left[P_L(q) \left(2 \int_{\mathbf{p}} F_2^2(\mathbf{k} - \mathbf{q} - \mathbf{p}, \mathbf{p}) P_L(p) P_L(|\mathbf{k} - \mathbf{q} - \mathbf{p}|) + 6 \int_{\mathbf{p}} F_3(\mathbf{k} - \mathbf{q}, \mathbf{p}, -\mathbf{p}) P_L(p) P_L(|\mathbf{k} - \mathbf{q}|) \right) \right. \quad (\text{A32})$$

$$\left. + P_L(|\mathbf{k} - \mathbf{q}|) \left(2 \int_{\mathbf{p}} F_2^2(\mathbf{q} - \mathbf{p}, \mathbf{p}) P_L(p) P_L(|\mathbf{q} - \mathbf{p}|) + 6 \int_{\mathbf{p}} F_3(\mathbf{q}, \mathbf{p}, -\mathbf{p}) P_L(p) P_L(q) \right) \right], \quad (\text{A33})$$

which becomes, in the limit $k \rightarrow 0$

$$8 \int_{\mathbf{q}} C_{\delta_M}^2(q, q) \left[P_L(q) \left(\int_{\mathbf{p}} F_2^2(\mathbf{p}, -\mathbf{p} - \mathbf{q}) P_L(p) P_L(|\mathbf{q} + \mathbf{p}|) + 3 \int_{\mathbf{p}} F_3(-\mathbf{q}, -\mathbf{p}, \mathbf{p}) P_L(p) P_L(q) \right) \right] = \quad (\text{A34})$$

$$= c_2^2 \mathcal{K}_{n_1} + 2c_1 c_2 \mathcal{K}_{n_2} + c_1^2 \mathcal{K}_{n_3}. \quad (\text{A35})$$

For the second piece in Eq. (A29), the one involving the nonlinear trispectrum, we just need to stop at the tree-level, as it already involves three power spectra

$$\begin{aligned} T_{\text{NL}}(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3, \mathbf{q}_4) &= T_{\text{tree}}(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3, \mathbf{q}_4) = 4P_L(q_1)P_L(q_2)(F_2(\mathbf{q}_1, -\mathbf{q}_1 - \mathbf{q}_3)F_2(\mathbf{q}_2, \mathbf{q}_1 + \mathbf{q}_3)P_L(|\mathbf{q}_1 + \mathbf{q}_3|) \\ &\quad + F_2(\mathbf{q}_1, -\mathbf{q}_1 - \mathbf{q}_4)F_2(\mathbf{q}_2, \mathbf{q}_1 + \mathbf{q}_4)P_L(|\mathbf{q}_1 + \mathbf{q}_4|)) + 5 \text{ perm.} \\ &\quad + 6F_3(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3)P_L(q_1)P_L(q_2)P_L(q_3) + 3 \text{ perm.}, \end{aligned} \quad (\text{A36})$$

The piece with F_3 gives

$$\begin{aligned} 6 \int_{\mathbf{q}_1} \int_{\mathbf{q}_2} C_{\delta_M}^2(q_1, |\mathbf{k} - \mathbf{q}_1|) C_{\delta_M}^2(q_2, |\mathbf{k} + \mathbf{q}_2|) [F_3(\mathbf{q}_1, \mathbf{k} - \mathbf{q}_1, \mathbf{q}_2) P_L(q_1) P_L(q_2) P_L(|\mathbf{k} - \mathbf{q}_1|) \\ + F_3(\mathbf{q}_1, \mathbf{k} - \mathbf{q}_1, -\mathbf{k} - \mathbf{q}_2) P_L(q_1) P_L(|\mathbf{k} - \mathbf{q}_1|) P_L(|\mathbf{k} + \mathbf{q}_2|) \\ + F_3(\mathbf{q}_1, \mathbf{q}_2, -\mathbf{k} - \mathbf{q}_2) P_L(q_1) P_L(q_2) P_L(|\mathbf{k} + \mathbf{q}_2|) \\ + F_3(\mathbf{k} - \mathbf{q}_1, \mathbf{q}_2, -\mathbf{k} - \mathbf{q}_2) P_L(q_2) P_L(|\mathbf{k} - \mathbf{q}_1|) P_L(|\mathbf{k} + \mathbf{q}_2|)], \end{aligned} \quad (\text{A37})$$

which, in the limit $k \rightarrow 0$ becomes

$$12 \int_{\mathbf{q}_1} \int_{\mathbf{q}_2} C_{\delta_M^2}(q_1, q_1) C_{\delta_M^2}(q_2, q_2) [F_3(\mathbf{q}_1, -\mathbf{q}_1, \mathbf{q}_2) P_L(q_1)^2 P_L(q_2) + F_3(\mathbf{q}_1, \mathbf{q}_2, -\mathbf{q}_2) P_L(q_1) P_L(q_2)^2] \\ = c_2^2 \mathcal{K}_x + c_2 c_1 \mathcal{K}_y + c_1^2 \mathcal{K}_z. \quad (\text{A38})$$

Finally, let us consider the first term involving the mark coefficient c_3 , $[3] - [3]$, see Eq. (A11)

$$\bar{m}^2 \langle \delta_M^{[3]}(\mathbf{k}) \delta_M^{[3]}(\mathbf{k}') \rangle = \mathcal{I}_{\mathbf{k}; \mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3} \mathcal{I}_{\mathbf{k}'; \mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3} H_3(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) H_3(\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3) \langle \delta_{\mathbf{q}_1} \delta_{\mathbf{q}_2} \delta_{\mathbf{q}_3} \delta_{\mathbf{p}_1} \delta_{\mathbf{p}_2} \delta_{\mathbf{p}_3} \rangle. \quad (\text{A39})$$

The correlator inside the integral gives a different kind of contribution, $\sim PPP, BB, PT, 6\text{pt}$: here we focus on the contribution $\sim PPP$ which is nothing but a term coming from the 2loop power spectrum. It reads

$$\bar{m}^2 \langle \delta_M^{[3]}(\mathbf{k}) \delta_M^{[3]}(\mathbf{k}') \rangle'_{PPP} = 9 P_L(k) \int_{\mathbf{q}, \mathbf{p}} H_3(\mathbf{k}, \mathbf{q}, -\mathbf{q}) H_3(-\mathbf{k}, \mathbf{p}, -\mathbf{p}) P_L(q) P_L(p) \\ + 6 \int_{\mathbf{q}, \mathbf{p}} [H_3(\mathbf{q}, \mathbf{p}, \mathbf{k} - \mathbf{q}, -\mathbf{p})]^2 P_L(q) P_L(p) P_L(|\mathbf{k} - \mathbf{q} - \mathbf{p}|). \quad (\text{A40})$$

The first line behaves as $\sim P_L(k)$ at large scales, while the second line gives constant contributions $\propto (c_3^2, c_2 c_3, c_2^2)$. As a consequence of this, both c_2 and c_3 have to undergo the perturbativity criteria delineated in Section II C.

APPENDIX B: AFFINE INVARIANCE OF THE FISHER MATRIX FOR HIGHER-ORDER MARKED CORRELATORS

In the main text, we stated that the affine invariance of the Fisher matrix extends beyond the marked power spectrum to any complete set of n -point correlators of the density field δ and the marked field Δ . Here we provide the explicit proof.

1. General argument

Consider a data vector $\mathbf{d}$ comprising all auto- and cross-spectra of δ and Δ up to order n . The Fisher matrix is

$$F_{\alpha\beta} = (\partial_\alpha \mathbf{d})^T C^{-1} \partial_\beta \mathbf{d}, \quad (\text{B1})$$

where C is the covariance matrix, which we leave completely general. Under the affine transformation of the mark function $m' = A m + B$, the marked field transforms linearly as

$$\Delta' = A \Delta + B \delta, \quad (\text{B2})$$

since $\Delta = m(\delta_R)(1 + \delta) - \langle m(\delta_R)(1 + \delta) \rangle$ is linear in m by construction. Because every element of $\mathbf{d}$ is multilinear in the fields, each n -point correlator transforms as a polynomial in A and B of degree at most n . The data

vector therefore transforms linearly as $\mathbf{d}' = \mathbf{T} \mathbf{d}$, where $\mathbf{T}$ is a square matrix whose entries are polynomials in A and B . The covariance and derivatives transform accordingly as

$$C' = \mathbf{T} C \mathbf{T}^T, \quad \partial_\alpha \mathbf{d}' = \mathbf{T} \partial_\alpha \mathbf{d}. \quad (\text{B3})$$

Substituting into the Fisher matrix,

$$F'_{\alpha\beta} = (\partial_\alpha \mathbf{d}')^T \mathbf{T}^T (\mathbf{T} C \mathbf{T}^T)^{-1} \mathbf{T} \partial_\beta \mathbf{d} = (\partial_\alpha \mathbf{d})^T C^{-1} \partial_\beta \mathbf{d} = F_{\alpha\beta}, \quad (\text{B4})$$

where we used $\mathbf{T}^T (\mathbf{T}^T)^{-1} = \mathbf{I}$ and $\mathbf{T}^{-1} \mathbf{T} = \mathbf{I}$.

The proof therefore reduces entirely to showing that $\mathbf{T}$ is invertible. This follows from the fact that $\mathbf{T}$ is lower-triangular, as can be seen by recalling that substituting $\Delta' = A \Delta + B \delta$ into any n -point correlator only generates correlators of equal or lower degree in Δ , never higher. The determinant is therefore simply the product of the diagonal entries, each of which is a power of A , giving $\det \mathbf{T} = A^{n(n+1)/2} \neq 0$ for $A \neq 0$, as explicitly verified in the example below.

2. Explicit example: Marked power spectrum and bispectrum

We illustrate the argument with the data vector comprising the full set of marked power spectra and bispectra,

$$\mathbf{d} = (P_{\delta\delta}, P_{\delta\Delta}, P_{\Delta\Delta}, B_{\delta\delta\delta}, B_{\delta\delta\Delta}, B_{\delta\Delta\Delta}, B_{\Delta\Delta\Delta}). \quad (\text{B5})$$

Using $\Delta' = A \Delta + B \delta$, the transformation rules follow from the multilinearity of each correlator. For the power spectra

$$P'_{\delta\delta} = P_{\delta\delta}, \quad (\text{B6})$$

The transformation matrix $\mathbf{T}$ is therefore lower-triangular,

$$P'_{\delta\Delta} = AP_{\delta\Delta} + BP_{\delta\delta}, \quad (\text{B7})$$

$$P'_{\Delta\Delta} = A^2P_{\Delta\Delta} + 2ABP_{\delta\Delta} + B^2P_{\delta\delta}. \quad (\text{B8})$$

For the bispectra, expanding $\Delta' = A\Delta + B\delta$ in each marked slot:

$$B'_{\delta\delta\delta} = B_{\delta\delta\delta}, \quad (\text{B9})$$

$$B'_{\delta\delta\Delta} = AB_{\delta\delta\Delta} + BB_{\delta\delta\delta}, \quad (\text{B10})$$

$$B'_{\delta\Delta\Delta} = A^2B_{\delta\Delta\Delta} + 2ABB_{\delta\delta\Delta} + B^2B_{\delta\delta\delta}, \quad (\text{B11})$$

$$B'_{\Delta\Delta\Delta} = A^3B_{\Delta\Delta\Delta} + 3A^2BB_{\delta\Delta\Delta} + 3AB^2B_{\delta\delta\Delta} + B^3B_{\delta\delta\delta}. \quad (\text{B12})$$

$$\mathbf{T} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ B & A & 0 & 0 & 0 & 0 & 0 \\ B^2 & 2AB & A^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & B & A & 0 & 0 \\ 0 & 0 & 0 & B^2 & 2AB & A^2 & 0 \\ 0 & 0 & 0 & B^3 & 3AB^2 & 3A^2B & A^3 \end{pmatrix}, \quad (\text{B13})$$

with determinant $\det \mathbf{T} = A^9$, which is nonvanishing for any $A \neq 0$. The matrix $\mathbf{T}$ is therefore invertible, and the Fisher matrix is invariant under the affine transformation $F'_{\alpha\beta} = F_{\alpha\beta}$ for any covariance structure C .

APPENDIX C: ADDITIONAL RESULTS

We report here further results of our analysis. In Table III we report the mark coefficient obtained through the optimization procedure. The resulting mark functions are plotted in 2.

TABLE III. Values for the polynomial mark for smoothing radii $R = 15, 30 \text{ Mpc}/h$ at $z = 0$ and $z = 0.5$.

$R = 10 \text{ Mpc}h^{-1}, z = 0$				$R = 15 \text{ Mpc}h^{-1}, z = 0$				$R = 30 \text{ Mpc}h^{-1}, z = 0$			
Parameter	c_1	c_2	c_3	Parameter	c_1	c_2	c_3	Parameter	c_1	c_2	c_3
FoM	-0.3	0.02	0.0036	FoM	-0.4	-0.15	0.068	FoM	3.6	4.8	0.4
Ω_m	-0.3	0.02	0.0036	Ω_m	-0.4	-0.15	0.068	Ω_m	-3.6	4.0	4.8
σ_8	-0.3	0.02	0.0036	σ_8	-0.4	-0.15	0.068	σ_8	-2.4	2.8	2.8
h	-0.3	0.02	0.0036	h	-0.4	-0.15	0.068	h	-3.6	4.0	4.8
$R = 10 \text{ Mpc}h^{-1}, z = 0.5$				$R = 15 \text{ Mpc}h^{-1}, z = 0.5$				$R = 30 \text{ Mpc}h^{-1}, z = 0.5$			
Parameter	c_1	c_2	c_3	Parameter	c_1	c_2	c_3	Parameter	c_1	c_2	c_3
FoM	-0.4	0.03	0.0085	FoM	-3.3	2.14	1.39	FoM	4.4	8.4	3.6
Ω_m	-0.4	0.03	0.0085	Ω_m	2.1	1.35	0.23	Ω_m	0.4	0.8	1.2
σ_8	-0.4	0.03	0.0085	σ_8	2.5	1.07	-0.70	σ_8	4.4	8.4	3.6
h	-0.4	0.03	0.0085	h	0.8	0.07	-0.046	h	2.4	4.8	3.2
$R = 10 \text{ Mpc}h^{-1}, z = 1$				$R = 15 \text{ Mpc}h^{-1}, z = 1$				$R = 30 \text{ Mpc}h^{-1}, z = 1$			
Parameter	c_1	c_2	c_3	Parameter	c_1	c_2	c_3	Parameter	c_1	c_2	c_3
FoM	-0.6	0.048	0.021	FoM	-2.5	1.3	0.64	FoM	-2.8	6.0	8.8
Ω_m	-0.4	0.14	0.062	Ω_m	4.1	3.0	2.55	Ω_m	1.6	4.8	4.0
σ_8	-0.6	0.048	0.021	σ_8	-2.8	2.5	1.18	σ_8	2.4	7.2	-9.9
h	-0.4	0.14	0.062	h	1.7	1.5	-1.36	h	3.6	8.4	-3.6

In Table IV we report the complete results of our analysis using both Gaussian and simulation-based covariances.

We also show in Fig. 6 a simple example used to estimate the value for f_{cond} . We vary only c_1 and compare the error

bars on the different cosmological parameters obtained using only P and the combination $P + C + M$, with c_s^2 fixed to make it easier to visualize.

TABLE IV. Improvement on the parameter error bars for Gaussian and simulation-based covariances.

Optimization of FoM				Optimization of single parameters			
$R = 10 \text{ Mpch}^{-1} \ z = 0$				$R = 10 \text{ Mpch}^{-1} \ z = 0$			
Covariance	$r(\Omega_m)$	$r(\sigma_8)$	$r(h)$	Covariance	$r(\Omega_m)$	$r(\sigma_8)$	$r(h)$
Gaussian	1.54	14.69	1.60	Gaussian	1.54	14.169	1.60
Simulations	2.34	8.32	2.47	Simulations	2.34	8.32	2.47
$R = 10 \text{ Mpch}^{-1} \ z = 0.5$				$R = 10 \text{ Mpch}^{-1} \ z = 0.5$			
Covariance	$r(\Omega_m)$	$r(\sigma_8)$	$r(h)$	Covariance	$r(\Omega_m)$	$r(\sigma_8)$	$r(h)$
Gaussian	1.20	2.75	1.23	Gaussian	1.21	2.75	1.23
Simulations	1.53	3.11	1.50	Simulations	1.53	3.11	1.50
$R = 10 \text{ Mpch}^{-1} \ z = 1$				$R = 10 \text{ Mpch}^{-1} \ z = 1$			
Covariance	$r(\Omega_m)$	$r(\sigma_8)$	$r(h)$	Covariance	$r(\Omega_m)$	$r(\sigma_8)$	$r(h)$
Gaussian	1.12	1.86	1.13	Gaussian	1.12	1.86	1.13
Simulations	1.43	2.06	1.38	Simulations	1.38	2.06	1.27
$R = 15 \text{ Mpch}^{-1} \ z = 0$				$R = 15 \text{ Mpch}^{-1} \ z = 0$			
Covariance	$r(\Omega_m)$	$r(\sigma_8)$	$r(h)$	Covariance	$r(\Omega_m)$	$r(\sigma_8)$	$r(h)$
Gaussian	2.17	12.05	2.30	Gaussian	2.17	12.05	2.30
Simulations	1.48	5.55	1.57	Simulations	1.48	5.55	1.57
$R = 15 \text{ Mpch}^{-1} \ z = 0.5$				$R = 15 \text{ Mpch}^{-1} \ z = 0.5$			
Covariance	$r(\Omega_m)$	$r(\sigma_8)$	$r(h)$	Covariance	$r(\Omega_m)$	$r(\sigma_8)$	$r(h)$
Gaussian	1.75	3.97	1.81	Gaussian	1.81	5.10	1.65
Simulations	1.89	2.79	1.83	Simulations	1.20	1.19	1.15
$R = 15 \text{ Mpch}^{-1} \ z = 1$				$R = 15 \text{ Mpch}^{-1} \ z = 1$			
Covariance	$r(\Omega_m)$	$r(\sigma_8)$	$r(h)$	Covariance	$r(\Omega_m)$	$r(\sigma_8)$	$r(h)$
Gaussian	1.42	2.50	1.44	Gaussian	1.58	2.61	1.98
Simulations	1.43	2.06	1.38	Simulations	1.09	2.21	1.23
$R = 30 \text{ Mpch}^{-1} \ z = 0$				$R = 30 \text{ Mpch}^{-1} \ z = 0$			
Covariance	$r(\Omega_m)$	$r(\sigma_8)$	$r(h)$	Covariance	$r(\Omega_m)$	$r(\sigma_8)$	$r(h)$
Gaussian	2.01	11.17	2.04	Gaussian	2.19	12.88	2.62
Simulations	1.28	1.83	1.25	Simulations	1.72	2.50	1.69
$R = 30 \text{ Mpch}^{-1} \ z = 0.5$				$R = 30 \text{ Mpch}^{-1} \ z = 0.5$			
Covariance	$r(\Omega_m)$	$r(\sigma_8)$	$r(h)$	Covariance	$r(\Omega_m)$	$r(\sigma_8)$	$r(h)$
Gaussian	1.97	4.12	2.00	Gaussian	2.24	4.12	2.17
Simulations	1.23	1.25	1.16	Simulations	1.29	1.25	1.19
$R = 30 \text{ Mpch}^{-1} \ z = 1$				$R = 30 \text{ Mpch}^{-1} \ z = 1$			
Covariance	$r(\Omega_m)$	$r(\sigma_8)$	$r(h)$	Covariance	$r(\Omega_m)$	$r(\sigma_8)$	$r(h)$
Gaussian	1.35	2.21	1.42	Gaussian	2.07	2.77	2.13
Simulations	1.48	1.72	1.32	Simulations	1.36	1.67	1.26

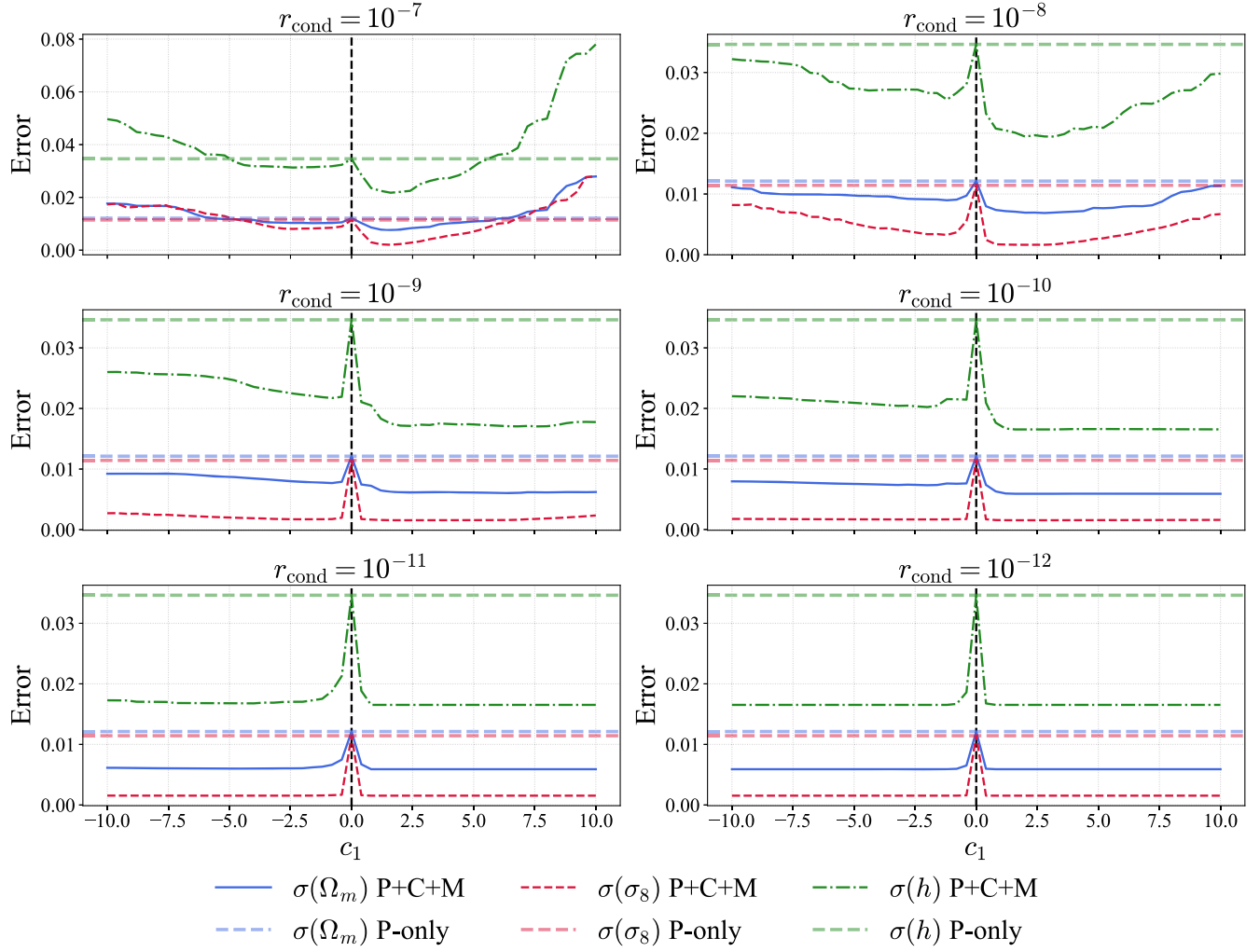


FIG. 6. Dependence of the forecasted errors on the conditioning number.

-
- [1] R. Laureijs *et al.* (EUCLID Collaboration), Euclid Definition Study Report, ESA Report, [arXiv:1110.3193](#).
- [2] Y. Mellier *et al.* (Euclid Collaboration), Euclid. I. Overview of the Euclid mission, *Astron. Astrophys.* **697**, A1 (2025).
- [3] A. Aghamousa *et al.* (DESI Collaboration), The DESI experiment Part I: Science, targeting, and survey design, [arXiv:1611.00036](#).
- [4] O. Doré *et al.* (SPHEREx Collaboration), Cosmology with the SPHEREx all-sky spectral survey, [arXiv:1412.4872](#).
- [5] A. Abate *et al.* (LSST Dark Energy Science Collaboration), Large synoptic survey telescope: Dark energy science collaboration 10.2172/1156445, [arXiv:1211.0310](#).
- [6] Y. Wang *et al.*, The high latitude spectroscopic survey on the nancy grace Roman space telescope, *Astrophys. J.* **928**, 1 (2022).
- [7] R. Scoccimarro, H. M. P. Couchman, and J. A. Frieman, The bispectrum as a signature of gravitational instability in redshift-space, *Astrophys. J.* **517**, 531 (1999).
- [8] R. Scoccimarro, The bispectrum: From theory to observations, *Astrophys. J.* **544**, 597 (2000).
- [9] F. Bernardeau, S. Colombi, E. Gaztanaga, and R. Scoccimarro, Large scale structure of the universe and cosmological perturbation theory, *Phys. Rep.* **367**, 1 (2002).
- [10] E. Sefusatti, M. Crocce, S. Pueblas, and R. Scoccimarro, Cosmology and the bispectrum, *Phys. Rev. D* **74**, 023522 (2006).
- [11] 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).

- [12] 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).
- [13] E. Krause *et al.* (Beyond-2pt Collaboration), A parameter-masked mock data challenge for beyond-two-point galaxy clustering statistics, *Astrophys. J.* **990**, 99 (2025).
- [14] R. K. Sheth, A. J. Connolly, and R. Skibba, Marked correlations in galaxy formation models, *arXiv:astro-ph/0511773*.
- [15] M. White, A marked correlation function for constraining modified gravity models, *J. Cosmol. Astropart. Phys.* **11** (2016) 057.
- [16] G. Valogiannis and R. Bean, Beyond δ : Tailoring marked statistics to reveal modified gravity, *Phys. Rev. D* **97**, 023535 (2018).
- [17] O. H. E. Philcox, E. Massara, and D. N. Spergel, What does the marked power spectrum measure? Insights from perturbation theory, *Phys. Rev. D* **102**, 043516 (2020).
- [18] O. H. E. Philcox, A. Aviles, and E. Massara, Modeling the marked spectrum of matter and biased tracers in real- and redshift-space, *J. Cosmol. Astropart. Phys.* **03** (2021) 038.
- [19] E. Massara, F. Villaescusa-Navarro, S. Ho, N. Dalal, and D. N. Spergel, Using the marked power spectrum to detect the signature of neutrinos in large-scale structure, *Phys. Rev. Lett.* **126**, 011301 (2021).
- [20] E. Massara, F. Villaescusa-Navarro, C. Hahn, M. M. Abidi, M. Eickenberg, S. Ho, P. Lemos, A. Moradinezhad Dizgah, and B. R.-S. Blacard, Cosmological information in the marked power spectrum of the galaxy field, *Astrophys. J.* **951**, 70 (2023).
- [21] M. Kärcher, J. Bel, and S. de la Torre, Towards an optimal marked correlation function analysis for the detection of modified gravity, *Astron. Astrophys.* **694**, A253 (2025).
- [22] J. A. Cowell, D. Alonso, and J. Liu, Optimizing marked power spectra for cosmology, *Mon. Not. R. Astron. Soc.* **535**, 3129 (2024).
- [23] J. A. Cowell, J. Armijo, L. Thiele, G. A. Marques, C. P. Novaes, D. Grandón, S. Cheng, M. Shirasaki, D. Alonso, and J. Liu, First constraints from marked angular power spectra with Subaru Hyper Suprime-Cam Survey First-Year Data, *Mon. Not. R. Astron. Soc.* **546**, stag033 (2026).
- [24] H. Ebina and M. White, An analytically tractable marked power spectrum, *J. Cosmol. Astropart. Phys.* **01** (2025) 150.
- [25] H. Ebina, M. White, and E. Chaussidon, The marked power spectrum as a practical bispectrum measure for galaxy redshift surveys, *J. Cosmol. Astropart. Phys.* **08** (2026) 003.
- [26] M. Marinucci *et al.*, The constraining power of the marked power spectrum: an analytical study, *J. Cosmol. Astropart. Phys.* **09** (2025) 036.
- [27] E. Massara, C. Hahn, M. Eickenberg, S. Ho, J. Hou, P. Lemos, C. Modi, A. Moradinezhad Dizgah, L. Parker, and B. R.-S. Blacard, Cosmological constraints using simulation-based inference of galaxy clustering with marked power spectra, *Phys. Rev. D* **112**, 083507 (2025).
- [28] D. Bianchi, Augmented correlation functions for spectroscopic galaxy surveys, *arXiv:2605.30305*.
- [29] F. Semenzato, B. D. Wandelt, M. Liguori, and A. Raccanelli, Interpretable neural marked statistics for cosmological inference, *arXiv:2606.11295*.
- [30] 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.
- [31] D. Baumann, A. Nicolis, L. Senatore, and M. Zaldarriaga, Cosmological non-linearities as an effective fluid, *J. Cosmol. Astropart. Phys.* **07** (2012) 051.
- [32] L. Senatore and M. Zaldarriaga, The IR-resummed effective field theory of large scale structures, *J. Cosmol. Astropart. Phys.* **02** (2015) 013.
- [33] L. Senatore and M. Zaldarriaga, Redshift space distortions in the effective field theory of large scale structures, *arXiv:1409.1225*.
- [34] 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.
- [35] F. Spezzati, M. Marinucci, and M. Simonović, Equivalence of the field-level inference and conventional analyses on large scales, *Phys. Rev. D* **113**, 063546 (2026).
- [36] F. Villaescusa-Navarro *et al.*, The Quijote simulations, *Astrophys. J. Suppl. Ser.* **250**, 2 (2020).
- [37] G. D'Amico, Y. Donath, M. Lewandowski, L. Senatore, and P. Zhang, The one-loop bispectrum of galaxies in redshift space from the Effective Field Theory of Large-Scale Structure, *J. Cosmol. Astropart. Phys.* **07** (2024) 041.
- [38] T. Bakx, M. M. Ivanov, O. H. E. Philcox, and Z. Vlah, One-loop galaxy bispectrum: Consistent theory, efficient analysis with COBRA, and implications for cosmological parameters, *arXiv:2507.22110*.
- [39] T. Bakx, M. Garny, H. Rubira, and Z. Vlah, Two-loop renormalization and running of galaxy bias, *J. Cosmol. Astropart. Phys.* **03** (2026) 056.
- [40] T. Bakx, H. Rubira, N. E. Chisari, and Z. Vlah, Rapid cosmological inference with the two-loop matter power spectrum, *arXiv:2508.00611*.
- [41] C. Anastasiou, A. Favorito, M. Lewandowski, L. Senatore, and H. Zheng, Efficient evaluation of the dark-matter two-loop power spectrum in the EFT of LSS, *arXiv:2509.05187*.
- [42] T. Steele and T. Baldauf, Precise calibration of the one-loop trispectrum in the effective field theory of large scale structure, *Phys. Rev. D* **103**, 103518 (2021).
- [43] 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.
- [44] D. Gualdi, H. é. Gil-Marín, and L. Verde, Joint analysis of anisotropic power spectrum, bispectrum and trispectrum: Application to N-body simulations, *J. Cosmol. Astropart. Phys.* **07** (2021) 008.
- [45] N. Aghanim *et al.* (Planck Collaboration), Planck 2018 results. VI. Cosmological parameters, *Astron. Astrophys.* **641**, A6 (2020); **652**, C4(E) (2021).
- [46] G. Cabass, M. Simonović, and M. Zaldarriaga, Cosmological information in perturbative forward modeling, *Phys. Rev. D* **109**, 043526 (2024).
- [47] 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.

- [48] G. D’Amico, A. Refregier, L. Senatore, and P. Zhang, The cosmological analysis of DES 3×2pt data from the effective field theory of large-scale structure, *J. Cosmol. Astropart. Phys.* **03** (2026) 062.
- [49] M. Tegmark, Measuring cosmological parameters with galaxy surveys, *Phys. Rev. Lett.* **79**, 3806 (1997).
- [50] G. D’Amico, J. Gleyzes, N. Kokron, K. Markovic, L. Senatore, P. Zhang, F. Beutler, and H. Gil-Marín, The cosmological analysis of the SDSS/BOSS data from the effective field theory of large-scale structure, *J. Cosmol. Astropart. Phys.* **05** (2020) 005.
- [51] M. M. Ivanov, M. Simonović, and M. Zaldarriaga, Cosmological parameters from the BOSS galaxy power spectrum, *J. Cosmol. Astropart. Phys.* **05** (2020) 042.
- [52] A. G. Adame *et al.* (DESI Collaboration), DESI 2024 VII: Cosmological constraints from the full-shape modeling of clustering measurements, *J. Cosmol. Astropart. Phys.* **07** (2025) 028.
- [53] V. Yankelevich and C. Porciani, Cosmological information in the redshift-space bispectrum, *Mon. Not. R. Astron. Soc.* **483**, 2078 (2019).
- [54] N. Agarwal, V. Desjacques, D. Jeong, and F. Schmidt, Information content in the redshift-space galaxy power spectrum and bispectrum, *J. Cosmol. Astropart. Phys.* **03** (2021) 021.
- [55] C. F. Park, E. Allys, F. Villaescusa-Navarro, and D. Finkbeiner, Quantification of high-dimensional non-Gaussianities and its implication to fisher analysis in cosmology, *Astrophys. J.* **946**, 107 (2023).
- [56] J. Carron, Information escaping the correlation hierarchy of the convergence field in the study of cosmological parameters, *Phys. Rev. Lett.* **108**, 071301 (2012).
- [57] M. Tegmark and P. J. E. Peebles, The Time evolution of bias, *Astrophys. J. Lett.* **500**, L79 (1998).
- [58] L. Piga, M. Marinucci, G. D’Amico, M. Pietroni, F. Vernizzi, and B. S. Wright, Constraints on modified gravity from the BOSS galaxy survey, *J. Cosmol. Astropart. Phys.* **04** (2023) 038.
- [59] P. Carrilho, C. Moretti, and A. Pourtsidou, Cosmology with the EFTofLSS and BOSS: Dark energy constraints and a note on priors, *J. Cosmol. Astropart. Phys.* **01** (2023) 028.
- [60] P. Taule, M. Marinucci, G. Biselli, M. Pietroni, and F. Vernizzi, Constraints on dark energy and modified gravity from the BOSS Full-Shape and DESI BAO data, *J. Cosmol. Astropart. Phys.* **03** (2025) 036.
- [61] M. Tsedrik, P. Carrilho, and C. Moretti, The simple way to measure evolving dark energy without prior-volume effects, *J. Cosmol. Astropart. Phys.* **04** (2026) 030.
- [62] A. Reeves, P. Zhang, and H. Zheng, Debiasing inference in large-scale structure with non-flat volume measures, [arXiv:2507.20991](https://arxiv.org/abs/2507.20991).
- [63] S. Alam *et al.* (BOSS Collaboration), The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: Cosmological analysis of the DR12 galaxy sample, *Mon. Not. R. Astron. Soc.* **470**, 2617 (2017).
- [64] T. Bonnaire, N. Aghanim, J. Kuruvilla, and A. Decelle, Cosmology with cosmic web environments—I. Real-space power spectra, *Astron. Astrophys.* **661**, A146 (2022).
- [65] T. Bonnaire, J. Kuruvilla, N. Aghanim, and A. Decelle, Cosmology with cosmic web environments—II. Redshift-space auto and cross-power spectra, *Astron. Astrophys.* **674**, A150 (2023).
- [66] F. Villaescusa-Navarro, Pylans: Python libraries for the analysis of numerical simulations, Astrophysics Source Code Library, record ascl:1811.008 (2018).
- [67] N. Hand, Y. Feng, F. Beutler, Y. Li, C. Modi, U. Seljak, and Z. Slepian, nboddykit: An open-source, massively parallel toolkit for large-scale structure, *Astron. J.* **156**, 160 (2018).
- [68] J. Hartlap, P. Simon, and P. Schneider, Why your model parameter confidences might be too optimistic: Unbiased estimation of the inverse covariance matrix, *Astron. Astrophys.* **464**, 399 (2007).
- [69] G. Jung, A. Ravenni, M. Liguori, M. Baldi, W. R. Coulton, F. Villaescusa-Navarro, and B. D. Wandelt, Quijote-PNG: Optimizing the summary statistics to measure primordial non-Gaussianity, *Astrophys. J.* **976**, 109 (2024).
- [70] 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.
- [71] M. Marinucci, K. Pardede, and M. Pietroni, Bootstrapping Lagrangian perturbation theory for the large scale structure, *J. Cosmol. Astropart. Phys.* **10** (2024) 051.
- [72] G. Biselli, M. Marinucci, G. D’Amico, and M. Pietroni, Probing nonlinear structure formation beyond Λ CDM with the LSS bootstrap: A joint power spectrum and bispectrum analysis, [arXiv:2605.13298](https://arxiv.org/abs/2605.13298).