UC Berkeley
UC Berkeley Previously Published Works

Title
The marked power spectrum as a practical bispectrum measure for galaxy redshift surveys

Permalink
https://escholarship.org/uc/item/48b1z9tw

Journal
Journal of Cosmology and Astroparticle Physics, 2026(08)

ISSN
1475-7516

Authors
Ebina, Haruki
White, Martin
Chaussidon, Edmond

Publication Date
2026-08-01

DOI
10.1088/1475-7516/2026/08/003

Copyright Information
This work is made available under the terms of a Creative Commons Attribution License, available at https://creativecommons.org/licenses/by/4.0/

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Journal of Cosmology
and Astroparticle
Physics

SISSA

PAPER • OPEN ACCESS

The marked power spectrum as a practical bispectrum measure for galaxy redshift surveys

To cite this article: Haruki Ebina *et al* JCAP08(2026)003

View the article online for updates and enhancements.

You may also like

- An analytically tractable marked power spectrum
Haruki Ebina and Martin White
- Galaxy redshift-space bispectrum: the importance of being anisotropic
D. Gualdi and L. Verde
- Joint analysis of anisotropic power spectrum, bispectrum and trispectrum: application to N-body simulations
Davide Gualdi, Héctor Gil-Marín and Licia Verde

This content was downloaded from IP address 108.75.79.166 on 10/11/2026 at 19:47

Journal of Cosmology and Astroparticle Physics
An IOP and SISSA journal

Received: *March 9, 2026*
Revised: *May 30, 2026*
Accepted: *June 10, 2026*
Published: *August 3, 2026*

The marked power spectrum as a practical bispectrum measure for galaxy redshift surveys

Haruki Ebina[a,b,c] **Martin White**[a,b,c] **and Edmond Chaussidon**[c]

[a] *Department of Physics, University of California, Berkeley, CA 94720, U.S.A.*
[b] *Berkeley Center for Cosmological Physics, UC Berkeley, CA 94720, U.S.A.*
[c] *Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, U.S.A.*

E-mail: ebina@berkeley.edu, mwhite@berkeley.edu, echaussidon@lbl.gov

Abstract: Modern datasets have the precision necessary to uncover new information by including higher-order, non-Gaussian information into cosmological inference. The marked power spectrum offers access to such information while preserving the structure of two-point correlators. This approach to higher-order statistics has the advantage that many modeling questions can directly benefit from progress already made in standard cosmological analyses using the power spectrum and correlation function, while increasing the data vector size negligibly and retaining much of the degeneracy-breaking power of the bispectrum through a particular data compression. In this work, we first restructure the marked power spectrum to isolate its higher-order information and demonstrate its ability to break parameter degeneracies. We then investigate the effect of survey geometry on the marked power spectrum and find that a treatment similar to that of the power spectrum is sufficient. Additionally, we investigate the perturbative modeling and covariance structure of the marked power spectrum, shedding light on its degeneracy breaking power and cross-covariance with the power spectrum. Finally, we demonstrate that the cosmology dependence of the marked power spectrum is smooth, indicating that cosmological inference is possible by modeling the cosmology dependence through interpolation rather than analytical modeling.

Keywords: cosmological parameters from LSS, power spectrum, redshift surveys

ArXiv ePrint: 2603.03581

© 2026 The Author(s). Published by IOP Publishing Ltd on behalf of Sissa Medialab. Original content from this work may be used under the terms of the Creative Commons Attribution 4.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

https://doi.org/10.1088/1475-7516/2026/08/003

JCAP08(2026)003

Contents

1 Introduction **1**

2 Marked power spectrum **3**
2.1 Definition and modeling 3
2.2 Perturbative modeling 6
2.3 Nuisance parameters 7
2.4 Degeneracy breaking 9

3 Covariance matrix **12**
3.1 Covariance of M 13
3.2 Covariance of P-M 14
3.3 Scaling and relation to the bispectrum 15

4 The effect of survey geometry **17**
4.1 Power spectrum 18
4.2 Marked power spectrum 19
4.3 Implementation 20

5 Validation on mocks **21**
5.1 Periodic box mocks 21
5.2 Stochasticity with low number density 22
5.3 Cutsky mocks 23

6 Alcock-Paczynski effect **25**

7 Conclusions **26**

A The beyond-2pt information in $\mathcal{M}$ **28**

B Choice of marks **28**

C Scale-dependence of M_{13} **30**

D Covariance matrix **31**
D.1 Covariance of M 31
D.2 Covariance of P-M 38

1 Introduction

Large-scale structure offers a unique probe into the evolution of the universe by tracing the gravitational growth of the density field, enabling local observers to learn about cosmology, fundamental physics, and structure formation [1–3]. For redshift surveys to date, including the ongoing Dark Energy Spectroscopic Instrument (DESI) [4], the majority of the information

JCAP08(2026)003

has come from the two-point function, either in Fourier space (the power spectrum) or configuration space (the correlation function) [5, 6]. The near-Gaussianity of the density field on large scales makes this an efficient compression of information. However, this compression is not entirely lossless and with rapidly increasing observational precision, attention has turned to extracting the small amount of additional information present in higher-order statistics on quasi-linear scales. There are many means to access this additional information. Higher-order n-point functions, such as the bispectrum [7, 8] and trispectrum [9, 10], are the most straightforward generalizations of the two-point function, while a variety of alternative statistics exist, including skew spectra [11, 12], wavelet scattering transforms [13–16], density-split statistics [17–20] and the filter-squared bispectrum [21, 22]. Many of these methods have been compared in the recent "Beyond 2-pt Challenge" [23].

While straightforward in principle, higher-order n-point functions pose several practical challenges: computational complexity, large data vectors, requirements on covariances, window functions, sensitivity to systematics or artifacts from the survey operations [24–27]. Since much of the new information provided by higher order statistics comes from their ability to break degeneracies between "nuisance parameters" [28–30], that are themselves degenerate with cosmological parameters, there is reason to believe that alternatives that are practically easier to handle may achieve similar performance.

In particular, marked spectra can offer a practical means of breaking degeneracies among parameters in our theoretical models while being theoretically well controlled, having small data sizes and making the most use of the existing survey infrastructure [31]. The perturbative modeling of the marked power spectrum is on the same footing as the existing power spectrum models, with the marked spectrum having a straightforward relation to the bispectrum as well. Extensive infrastructure exists to handle survey practicalities and systematics in two-point functions and marked spectra can make use of much of this. Density field estimation, an essential component of marked spectra observation, is also well developed, since it is used as a key part of analyses of baryon acoustic oscillations. Further, marks that are low-order in the smoothed overdensity field, with smoothing radii that are larger than the non-linear scale, are especially amenable to an analytic treatment. For all of these reasons, properly constructed marked power spectra have been suggested as a means of enhancing cosmological inference [32, 33].

Previously, marked spectra have been proposed as a means of testing cosmology by up-weighting underdense regions with a particular inverse-weighting mark, which enhances sensitivity to modified gravity effects or massive neutrinos [31, 34, 35]. More recently, perturbative models for marked spectra have been developed [32, 36, 37] and the possibility of using different forms of the mark has been investigated [32, 38]. Ref. [32] in particular found that modeling is easier for marks that were low-order polynomials in the smoothed overdensity and have demonstrated that marked spectra is capable of improving cosmological inference by breaking degeneracies among parameters in the theoretical models. Following these works, we develop a method to incorporate survey window functions into the marked power spectrum, which is essential for applying marked spectra to real survey data. Furthermore, we restructure the marked power spectrum to isolate the higher-order information and demonstrate its ability to break parameter degeneracies.

JCAP08(2026)003

This paper is organized as follows. We begin in section 2 with a review of the marked power spectrum and its perturbative modeling, and restructure the spectrum to isolate the higher-order information. In section 3 we describe the covariance structure of marked power spectra and quantify the amount of new information as compared to the power spectrum. In section 4 we describe how to measure the marked spectrum on survey data and incorporate survey window functions in a similar manner to the power spectrum. In section 5.1 and section 5.3 we validate the modeling on periodic box and cutsky simulations, respectively. In section 6 we show how to accurately and cheaply incorporate the fiducial cosmology dependence in cosmology inference. Finally, we present our conclusions in section 7.

2 Marked power spectrum

The marked power spectrum (MPS) offers a principled, controlled method to include higher-order information beyond the power spectrum. MPS was initially introduced in cosmology[1] to look for modified gravity signals by up-weighting void (underdense) regions [31]. This particular mark (weighting) has been adopted by subsequent studies that predicted significant improvement in cosmological constraints, in particular neutrino mass [34, 35]. Recent work [32, 36, 37] has shown that a perturbative modeling of the MPS is possible without introducing UV uncertainties beyond those already in the power spectrum. Refs. [32, 38] investigate the possibility of using different forms of the mark beyond that was initially proposed, where ref. [32] in particular found that modeling is easier for marks that were low-order polynomials in the smoothed overdensity. In the following text we introduce the theoretical set up of the MPS, focusing on the linear mark investigated in ref. [32]. While the inverse-weighting mark [31] can be described in an identical, perturbative formalism to what we employ in this work [37], the infinite Taylor expansion of the mark creates unpredictable higher-loop contributions, making the pertubative modeling practically challenging. To one-loop power spectrum order ($\mathcal{O}(\delta^4)$), we believe that the linear formalism captures all new information introduced by any mark, as we will describe below.

As we will frequently transition between configuration and Fourier space, it will be useful to define the following notation

$$\langle f_1(\mathbf{x}_1) f_2(\mathbf{x}_2)\rangle_{\rm FT}(\mathbf{k}) = \int d\mathbf{x}\; e^{-i\mathbf{k}\,\cdot\,\mathbf{x}}\,\langle f_1(\mathbf{x}_1) f_2(\mathbf{x}_2)\rangle|_{\mathbf{x}=\mathbf{x}_1-\mathbf{x}_2} \tag{2.1}$$

where f_1 and f_2 are fields defined in configuration-space and we assume that the two-point correlator depends only on the separation vector between the two fields $\mathbf{x} = \mathbf{x}_1 - \mathbf{x}_2$. To simplify later expressions we shall also adopt the notation

$$\int_{\mathbf{k}_1\cdots\mathbf{k}_n} = \int \prod_{i=1}^{n} \frac{d^3k_i}{(2\pi)^3}\,. \tag{2.2}$$

2.1 Definition and modeling

The marked density field is defined by weighting the galaxy density field ρ_g by a mark field, m,

$$\rho_M(\mathbf{x}) = m\left[\delta_{g,R}(\mathbf{x})\right]\rho_g(\mathbf{x}) \tag{2.3}$$

[1]Preceding work to this was mostly in the context of astronomy, where marks were defined with observables, such as luminosity, rather than density [39, 40].

JCAP08(2026)003

taken to be a functional of the smoothed galaxy overdensity field, $\delta_{g,R}$. We define $\delta_{g,R}(k) = W_R(k)\delta_g(k)$ with a Gaussian kernel $W_R(k) = \exp\{-k^2R^2/2\}$.[2] Expanding the mark field in powers of $\delta_{g,R}$

$$m(\boldsymbol{x}) = \sum_n \widetilde{C}_n \delta^n_{g,R}(\boldsymbol{x}) \tag{2.4}$$

one finds that only terms with $n \leq 3$ appear when perturbatively modeling the power spectrum to one-loop order [32, 37]. For a controlled modeling of correlators of ρ_M, it is useful to consider marks that are low order in $\delta_{g,R}$ [32]. In this work, we follow ref. [32] and focus on marks that are linear in $\delta_{g,R}$

$$\frac{m(\boldsymbol{x})}{\bar{m}} = C_0 + C_1\delta_{g,R}(\boldsymbol{x}) \tag{2.5}$$

making the marked overdensity field

$$\delta_M = (m/\bar{m})\,(\delta_g + 1) - 1 \tag{2.6}$$
$$= (C_0 + C_1\delta_{g,R})(\delta_g + 1) - 1 \tag{2.7}$$

where $\bar{m} = \langle\rho_M\rangle\,/\,\langle\rho_g\rangle$ is an overall normalization that can be measured from simulations or data and we have defined $\widetilde{C}_n = \bar{m}\,C_n$. We shall be exclusively concerned with the Fourier-space two-point function of δ_M, which is known as the marked power spectrum (MPS). Being a two-point function, it can be efficiently computed and manipulated using the same infrastructure as is commonly used to measure the power spectrum in galaxy surveys. A key quality of the MPS is that it introduces no additional UV-divergences than the power spectrum, as the smoothing on all 'additional' fields prevents UV-divergent contact terms (zero-lag contractions, e.g. $\delta^2_g(\mathbf{x}) \supset \int P(p)dp$). This places the MPS on an equal footing with the power spectrum from the perspective of modeling.

Previous studies on the MPS [31, 32, 37] have focused directly on the two-point correlator of the marked overdensity field $\mathcal{M} = \langle\delta^2_M\rangle$ which, due to the constant term in each δ_g, inevitably includes two-point information already available from the power spectrum [32, 37]. Namely for two fields with marks m^a and m^b,

$$\mathcal{M} = \left\langle \delta^a_M \delta^b_M \right\rangle := \left\langle C^a_0 C^b_0 \delta^2_g + (C^a_0 C^b_1 + C^b_0 C^a_1)\delta_g\delta_{g,R} + C^a_1 C^b_1 \delta^2_{g,R} \right\rangle + M \tag{2.8}$$

where the first set of terms is merely a collection of power spectrum information, with $\left\langle\delta^2_g\right\rangle = P(k,\mu)$, $\langle\delta_g\delta_{g,R}\rangle = W_R(k)P(k,\mu)$ and $\left\langle\delta^2_{g,R}\right\rangle = W^2_R(k)P(k,\mu)$, and we have (re-)defined M to be the collection of terms involving higher-point information. From here on, we will refer to M as the *marked power spectrum*.[3] We refer the reader to appendix A for discussion on the two-point and beyond two-point information content in $\mathcal{M}$.

[2] The possibility of using a non-Gaussian kernel has been explored in past work [34, 36] and more recently in ref. [41], but will be outside the scope of this study.

[3] One can equally use a convention where one weighs the overdensity δ_g instead of ρ_g, making the marked overdensity $\bar{m}\delta_M = m\delta_g$, as in ref. [38]. Both conventions produce the spectra studied in this work, but are not in general equal.

JCAP08(2026)003

For any mark, the MPS M can be decomposed into three- and four-point correlators [32, 37]. For a linear mark, this decomposition is

$$M = \left[C_0^a C_1^b + C_0^b C_1^a\right] \langle \delta_{g,R}(\mathbf{x}_1)\, \delta_g(\mathbf{x}_1)\, \delta_g(\mathbf{x}_2)\rangle + 2C_1^a C_1^b\, \langle \delta_{g,R}(\mathbf{x}_1)\, \delta_g(\mathbf{x}_1)\, \delta_{g,R}(\mathbf{x}_2)\rangle + C_1^a C_1^b\, \langle \delta_{g,R}(\mathbf{x}_1)\, \delta_g(\mathbf{x}_1)\, \delta_{g,R}(\mathbf{x}_2)\, \delta_g(\mathbf{x}_2)\rangle \quad (2.9)$$

where the freedom in the coefficients shows our ability to access different terms by choice of mark(s). To access the first three-point correlator, we can choose one mark to be a constant, e.g. a cross-spectrum between the unmarked field ($m = 1$, so $\rho_M = \rho_g$) and $m = 1 + \delta_{g,R}$. In fact, any cross-spectrum between the unmarked field and a linear mark isolates the first term. To access the second and third terms, one must choose the spectrum between two linear marks, e.g. the auto-spectrum of $m = 1+\delta_{g,R}$, but to isolate these terms, one will need to null the first term, by e.g. the cross-spectrum between $1 + \delta_{g,R}$ and $1 - \delta_{g,R}$. To the extent that, at one-loop power spectra order ($\mathcal{O}(\delta^4)$), the four-point correlators merely contribute the disconnected trispectrum (convolutions of the power spectrum), the three-point correlation terms of MPS contain all of the new (beyond two point) information introduced by the MPS.

Thus, in this work, we mainly focus on the cross-spectrum between the unmarked field δ_g and the mark $m = 1 + \delta_{g,R}$ (further discussion of alternative marks can be found in appendix section B). This leaves us with

$$M = \frac{1}{\bar{m}} \langle \delta_{g,R}(\mathbf{x}_1)\, \delta_g(\mathbf{x}_1)\, \delta_g(\mathbf{x}_2)\rangle \quad (2.10)$$

in configuration space. In Fourier space

$$M(\mathbf{k}) = \frac{1}{\bar{m}} \langle [\delta_{g,R}\delta_g](\mathbf{k})\delta_g(-\mathbf{k})\rangle = \frac{1}{\bar{m}} \int \frac{d^3p}{(2\pi)^3} W_R(p) B(\mathbf{p}, \mathbf{k}-\mathbf{p}, -\mathbf{k}) \quad (2.11)$$

where B is the galaxy bispectrum. The MPS is an integral of the bispectrum with the integration limited to $p \sim R^{-1}$ by (one power of) the smoothing kernel.

This form of the MPS makes apparent the similarity of MPS with the skew spectrum, a wide class of composite-field two-point correlators [11, 42–44]. One of the first and simplest skew spectra, $\left\langle [\delta_{g,R}^2](\mathbf{k})\delta_g(-\mathbf{k})\right\rangle$, merely differs from the MPS expression above by an extra smoothed field in the convolution. This indicates that the information obtained between the MPS and skew spectra are similar, but likely greater from the MPS due to the lack of smoothing. On the other hand, the extra smoothing of the skew spectra likely aids the modeling and control of non-linearities. There are also practical differences in the measurement and window treatment. We will elaborate on this further in section 4. In general, an investigation into the differences between various composite-field two-point correlators would be interesting, but is beyond the scope of this work.

For the remainder of this work, we will not explicitly write the factor of $\bar{m}$ as it is only an overall factor of ≈ 1. This corresponds to setting the mark to $m = 1 + \bar{m}\delta_{g,R}$ and is well-defined, however one cannot know $\bar{m}$ without selecting a dataset and hence this mark would not be fit for a general theoretical discussion.

JCAP08(2026)003

2.2 Perturbative modeling

Similarly to the power spectrum, the MPS can be modeled perturbatively in powers of the linear overdensity field δ_L [32, 37]. In Eulerian Perturbation Theory (EPT; [1, 45]), the equations of motion are constructed based on the assumption that cold dark matter and baryons behave as perfect, pressure-less fluids. These equations of motions can then be solved perturbatively, predicting the nonlinear matter overdensity field $\delta = \rho/\bar{\rho} - 1$ and velocity divergence $\theta = \nabla \cdot \mathbf{v}$ order-by-order. In the Einstein-de Sitter limit, writing $\delta = \sum_{n=1}^{\infty} D^n \delta^{(n)}$ and $\theta = \sum_{n=1}^{\infty} D^n \theta^{(n)}$, with D the linear growth factor, one finds [1, 45]

$$\delta^{(n)}(\mathbf{k}) = \int_{\mathbf{k}_1 \ldots \mathbf{k}_n} (2\pi)^3 \delta_D \left(\sum_i \mathbf{k}_i - \mathbf{k} \right) \delta_L(\mathbf{k}_1) \cdots \delta(\mathbf{k}_n) F_n(\mathbf{k}_1, \ldots, \mathbf{k}_n, \mathbf{k}) \tag{2.12}$$

where F_n is a function of the wavevectors that can be computed via recurrence, and for which we give explicit expressions for $n = 1$ and 2 below. A similar expression holds for $\theta^{(n)}$ with the substitution of G_n for F_n. These predictions for the matter field are then translated into observable, redshift-space galaxy field by applying redshift-space distortions and using the bias expansion

$$\delta_g = b_1 \delta_m + b_2 (\delta_m^2 - \langle \delta_m^2 \rangle) + b_s (s^2 - \langle s^2 \rangle) + b_3 \mathcal{O}_3 + \cdots \tag{2.13}$$

where we expand in the underlying matter and shear field, and the expansion terms are limited by symmetry and perturbative order. The final perturbative solution to the redshift-space galaxy overdensity field then becomes [1, 45]

$$\delta_g^{(n)}(\mathbf{k}) = \int_{\mathbf{k}_1 \ldots \mathbf{k}_n} (2\pi)^3 \delta_D \left(\sum_i \mathbf{k}_i - \mathbf{k} \right) \delta_L(\mathbf{k}_1) \cdots \delta(\mathbf{k}_n) Z_n(\mathbf{k}_1, \ldots, \mathbf{k}_n, \mathbf{k}) \tag{2.14}$$

where the kernels Z_n are defined by a combination of density kernels, F_n, velocity kernels, G_n, and bias terms. The first few kernels relevant for this work are [1]

$$Z_1(\mathbf{k}_1) = b_1 + f\mu_1^2 \tag{2.15}$$

$$\begin{aligned} Z_2(\mathbf{k}_1, \mathbf{k}_2) = {} & b_1 F_2(\mathbf{k}_1, \mathbf{k}_2) + f\mu_k^2 G_2(\mathbf{k}_1, \mathbf{k}_2) + \frac{b_2}{2} + b_s \left(\frac{(\mathbf{k}_1 \cdot \mathbf{k}_2)^2}{k_1^2 k_2^2} - \frac{1}{3} \right) \\ & + \frac{fk\mu_k}{2} \left[\frac{\mu_1}{k_1} (b_1 + f\mu_2^2) + \frac{\mu_2}{k_2} (b_1 + f\mu_1^2) \right] \end{aligned} \tag{2.16}$$

with $\mathbf{k} = \mathbf{k}_1 + \mathbf{k}_2$, $\mu_k = \hat{k} \cdot \hat{n}$, $\mu_n = \hat{k}_n \cdot \hat{n}$, and the second-order density and velocity kernels are

$$F_2(\mathbf{k}_1, \mathbf{k}_2) = \frac{5}{7} + \frac{2}{7} \left(\hat{k}_1 \cdot \hat{k}_2 \right)^2 + \frac{1}{2} \left(\hat{k}_1 \cdot \hat{k}_2 \right) \left(\frac{k_1}{k_2} + \frac{k_2}{k_1} \right) \tag{2.17}$$

$$G_2(\mathbf{k}_1, \mathbf{k}_2) = \frac{3}{7} + \frac{4}{7} \left(\hat{k}_1 \cdot \hat{k}_2 \right)^2 + \frac{1}{2} \left(\hat{k}_1 \cdot \hat{k}_2 \right) \left(\frac{k_1}{k_2} + \frac{k_2}{k_1} \right). \tag{2.18}$$

It is then clear that the contributions to the (unmarked) power spectrum up to one-loop ($\mathcal{O}(\delta_L^4)$) order only arise from three distinct contributions

$$P(\mathbf{k}) = P_{11}(\mathbf{k}) + 2P_{13}(\mathbf{k}) + P_{22}(\mathbf{k}) \tag{2.19}$$

JCAP08(2026)003

with

$$P_{11} = \left\langle \left(\delta_g^{(1)}\right)^2 \right\rangle, \qquad P_{13} = \left\langle \delta_g^{(1)} \delta_g^{(3)} \right\rangle, \qquad P_{22} = \left\langle \left(\delta_g^{(2)}\right)^2 \right\rangle \tag{2.20}$$

where the first term is the tree-level power spectrum [46] and the latter two are one-loop contributions [47]. Note that the modeling here is limited to large scales by the validity of the original equations of motion set up for δ and θ. To go beyond this we take an effective field theory (EFT) approach, where one integrates over the small-scales that are not described by the large-scale theory. We will discuss this, along with the treatment of field stochasticity later in section 2.3. We refer the reader to ref. [1] for a detailed discussion of perturbative modeling and ref. [45] for a recent review of EFT in this context.

To extend the perturbative modeling to the MPS, it is useful to note that the smoothing of the overdensity field $\delta_{g,R}(\mathbf{k}) = W_R(k)\delta_g(\mathbf{k})$ simply extends to a smoothing on the perturbative solutions, i.e. $\delta_{g,R}(\mathbf{k}) = \sum_{n=1}^{\infty} W_R(k)\delta_g^{(n)}(\mathbf{k})$. This alone is sufficient to model the MPS, modulo stochastic and small-scale dynamics (EFT) terms that we describe in section 2.3. Recalling the expression for M (eq. (2.9)) the lowest order (one-loop in power spectrum, $\mathcal{O}(\delta_L^4)$) contributions then become

$$\langle \delta_{g,R,1}\, \delta_{g,1}\, \delta_{g,2} \rangle = \left\langle \delta_{g,R,1}^{(1)}\, \delta_{g,1}^{(2)}\, \delta_{g,2}^{(1)} \right\rangle + \left\langle \delta_{g,R,1}^{(2)}\, \delta_{g,1}^{(1)}\, \delta_{g,2}^{(1)} \right\rangle + \left\langle \delta_{g,R,1}^{(1)}\, \delta_{g,1}^{(1)}\, \delta_{g,2}^{(2)} \right\rangle, \tag{2.21}$$

$$\langle \delta_{g,R,1}\, \delta_{g,1}\, \delta_{g,R,2} \rangle = \left\langle \delta_{g,R,1}^{(1)}\, \delta_{g,1}^{(2)}\, \delta_{g,R,2}^{(1)} \right\rangle + \left\langle \delta_{g,R,1}^{(2)}\, \delta_{g,1}^{(1)}\, \delta_{g,R,2}^{(1)} \right\rangle + \left\langle \delta_{g,R,1}^{(1)}\, \delta_{g,1}^{(1)}\, \delta_{g,R,2}^{(2)} \right\rangle, \tag{2.22}$$

$$\langle \delta_{g,R,1}\, \delta_{g,1}\, \delta_{g,R,2}\, \delta_{g,2} \rangle = \left\langle \delta_{g,R,1}^{(1)}\, \delta_{g,1}^{(1)}\, \delta_{g,R,2}^{(1)}\, \delta_{g,2}^{(1)} \right\rangle \tag{2.23}$$

where the subscripts on the fields denote their spatial coordinates. The terms with three fields are all different compressions of the tree-level bispectrum $B_{112} = \left\langle \delta_g^{(1)} \delta_g^{(1)} \delta_g^{(2)} \right\rangle$ and the terms with four fields are convolutions of the tree-level power spectrum P_{11}. These terms can also be reorganized by the order of $\delta_M^{(n)}$ they belong to, i.e.

$$M = 2M_{13} + M_{22} \tag{2.24}$$

$$\begin{aligned} 2M_{13} = &\left[C_0^a C_1^b + C_0^b C_1^a\right] \left(\left\langle \delta_{g,R,1}^{(1)}\, \delta_{g,1}^{(2)}\, \delta_{g,2}^{(1)} \right\rangle + \left\langle \delta_{g,R,1}^{(2)}\, \delta_{g,1}^{(1)}\, \delta_{g,2}^{(1)} \right\rangle \right) \\ &+ 2C_1^a C_1^b \left(\left\langle \delta_{g,R,1}^{(1)}\, \delta_{g,1}^{(2)}\, \delta_{g,R,2}^{(1)} \right\rangle + \left\langle \delta_{g,R,1}^{(2)}\, \delta_{g,1}^{(1)}\, \delta_{g,R,2}^{(1)} \right\rangle \right) \end{aligned} \tag{2.25}$$

$$\begin{aligned} M_{22} = &\left[C_0^a C_1^b + C_0^b C_1^a\right] \left\langle \delta_{g,R,1}^{(1)}\, \delta_{g,1}^{(1)}\, \delta_{g,2}^{(2)} \right\rangle \\ &+ C_1^a C_1^b \left(2 \left\langle \delta_{g,R,1}^{(1)}\, \delta_{g,1}^{(1)}\, \delta_{g,R,2}^{(2)} \right\rangle + \left\langle \delta_{g,R,1}^{(1)}\, \delta_{g,1}^{(1)}\, \delta_{g,R,2}^{(1)}\, \delta_{g,2}^{(1)} \right\rangle \right). \end{aligned} \tag{2.26}$$

Note again that for this work, we will isolate the contribution $\left[C_0^a C_1^b + C_0^b C_1^a\right] \langle \delta_{g,R,1} \delta_{g,1} \delta_{g,2} \rangle$ by choosing to observe the cross-spectrum between the unmarked field and marked field with mark $m = 1 + \delta_{g,R}$, as it contains the least amount of smoothing while accessing information in the bispectrum directly. The contractions that contribute to this term can be diagrammatically described as shown in figure 1. We refer the reader to refs. [32, 37] for detailed discussions of the subtleties in the perturbative modeling of the MPS ($\mathcal{M}$).

2.3 Nuisance parameters

The equations of motion central to the perturbative modeling above hold only for large-scale dynamics where the ideal assumptions about matter are approximately correct. At

JCAP08(2026)003

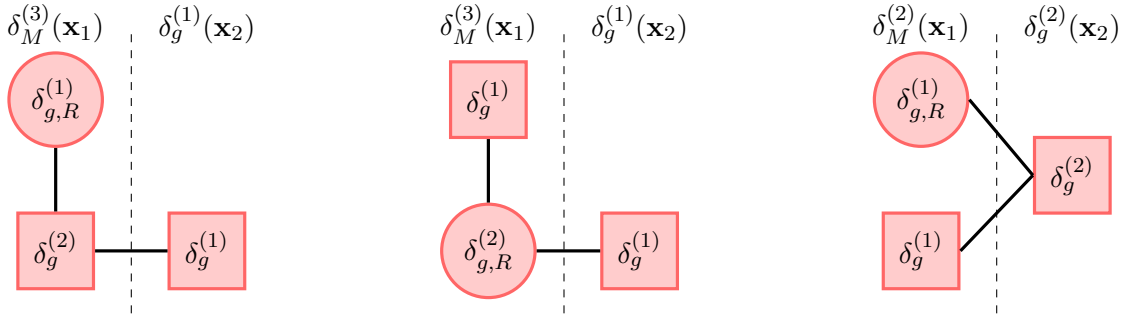


Figure 1. The three diagrams contributing to higher-order information in the MPS for a cross-spectrum between marked field δ_M at $\mathbf{x}_1$ and unmarked field δ_g at $\mathbf{x}_2$. The red squares and circles represent the raw and smoothed overdensity operators δ_g and $\delta_{g,R}$, respectively. The operators on the left of each diagram belong to the marked field, δ_M, and those on the right belong to the unmarked field, δ_g. The left two diagrams contribute to M_{13}, whereas the rightmost diagram contributes to M_{22}. Notice that all diagrams include a $\delta^{(2)} \supset Z_2$, giving rise to a "leading-order" dependence on b_2 and b_s that will be important later.

small-scales, gravitational collapse and structure formation induce effects beyond the model. These effects are incorporated into the large-scale theory by the effective field theory (EFT) approach, where one integrates over the small-scales and introduce the maximal degrees of freedom allowed by symmetry. In addition, one needs to account for the stochasticity ϵ of the density field. These effects together introduce the counterterm corrections

$$P(k,\mu) \supset \sum_{n=0}^{2} \alpha_{2n} \mu^{2n} k^2 P_L(k) \tag{2.27}$$

and stochastic terms

$$P(k,\mu) \supset \sum_{n=0}^{2} N_{2n} (k\mu)^{2n} \tag{2.28}$$

where $N_0 = N$ is the shotnoise and N_2, N_4 include the fingers-of-god [47].

While the MPS can largely be described using nuisance parameters already present in the modeling of the power spectrum, the new density contractions in the MPS implies that there is a set of new nuisance parameters introduced [32]. The only parameters that are included at one-loop power spectrum order (tree-level bispectrum order; $\mathcal{O}(\delta_L^4)$) are the stochastic terms $B_{\rm shot}$ and $A_{\rm shot}$, which stem from the tree-level bispectrum stochastic terms [48–50]

$$B \supset \left[\left(\frac{B_{\rm shot}}{\bar{n}} b_1 + N f \mu^2 \right) Z_1(\mathbf{k_1}) P_L(k_1) + \text{cycl.} \right] + \frac{A_{\rm shot}}{\bar{n}^2} \tag{2.29}$$

The contributions proportional to $B_{\rm shot}$, N, and $A_{\rm shot}$ arise from the products $\langle(\delta)(\epsilon)(\epsilon\delta)\rangle = \langle\epsilon^2\rangle\langle\delta^2\rangle$, $\langle(\delta)(\epsilon)(\epsilon f\mu^2\theta)\rangle = f\mu^2\langle\epsilon^2\rangle\langle\delta^2\rangle$ and $\langle(\epsilon)^3\rangle$ sourced by short wavelength contributions ϵ, $\epsilon\delta$ and $\epsilon f\mu^2\theta$ with ϵ a stochastic field uncorrelated with δ [32, 48]. As each of these quantities has an order of magnitude set by Poisson statistics, one expects $B_{\rm shot}$, $A_{\rm shot} = \mathcal{O}(1)$ and $P_{\rm shot} = \bar{n}N - 1 = \mathcal{O}(1)$ [49–51].

JCAP08(2026)003

In addition, we include a contribution that is formally higher-loop order following ref. [32]. In typical EFT modeling, contributions from higher-loop orders are introduced only at small scales and hence by limiting the analysis to large scales, one can terminate modeling at finite order. For this work, we limit M to (quasi-)linear scales ($k < 0.12\,h^{-1}\,\mathrm{Mpc}$) and thus will not require typical EFT correction terms (counterterms).[4] However, for the MPS there is a large-scale, perturbative,[5] higher-loop correction that arise from contact terms (fields at the same point in configuration space) [32, 37]. For this work, we only consider the MPS monopole and hence require only one free parameter

$$M_0 \supset A_0 P_L \tag{2.31}$$

where, due to the perturbative nature of the correction, we estimate $A_0 = \mathcal{O}(\sigma_R^4, \sigma_R^2\bar{\sigma}_R^2, \bar{\sigma}_R^4) \lesssim \sigma_R^2, \bar{\sigma}_R^2$ for the expansion coefficients defined as

$$\sigma_R^2 = \int_{\boldsymbol{p}} W_R(p) P_L(p), \qquad \bar{\sigma}_R^2 = \int_{\boldsymbol{p}} (pR)^2 W_R(p) P_L(p), \tag{2.32}$$

which, for the smoothing scales that we consider in this work, are $\lesssim 0.2$, 0.11, 0.07, for $R = 10$, 15, and $20\,h^{-1}\,\mathrm{Mpc}$ respectively, at $z = 0.7$,[6] [32]. In practice, we find that we do not need to vary A_0 for any of our theory fits to mock catalogs in this work. However, we will continue to include A_0 in the discussion of the text following theoretical predictions. Thus in what follows we will consider only the following new nuisance parameters

$$\{B_{\rm shot}, A_{\rm shot}, A_0\} \tag{2.33}$$

in addition to those from the one-loop power spectrum

$$\{b_1, b_2, b_s, \alpha_0, \alpha_2, N, N_2\} \tag{2.34}$$

and show that this provides the level of accuracy expected of next-generation surveys to jointly fit P_0, P_2, and M_0.[7] Note that choose to focus on the MPS monopole in this work due to modeling difficulties [32] the quadrupole and higher multipoles also have the potential to assist degeneracy breaking. This investigation will be left for future work, but detecting the value of b_s, which is a subdominant contribution to M_0 as described below, may be a potential use case.

2.4 Degeneracy breaking

The MPS depends upon bias parameters differently than the power spectrum, offering the possibility of breaking degeneracies. This is most apparent for the quadratic biases, b_2 and

[4] If one aims to model smaller scales, one will require counterterm corrections, such as this ansatz used in ref. [32]

$$M(k,\mu) \supset C_{\delta_M}(k) k^2 \left(N_{m,2}^{(0)} + \mu^2 N_{m,2}^{(2)}\right) \tag{2.30}$$

where $C_{\delta_M}(k) = C_0 + C_1 W_R(k)$.

[5] These corrections are similar to that in the Lyman-α forest two-point function, although with the critical distinction that the ones in MPS are UV-safe [32].

[6] The maximum correction (at $R = 10\,h^{-1}\,\mathrm{Mpc}$) grows from $\lesssim 0.15$ at $z = 1.1$ [32] to $\lesssim 0.2$ at $z = 0.7$.

[7] We neglect the third order bias b_3 following standard practice [6] for power spectrum analyses, as M does not introduce new b_3 dependencies.

JCAP08(2026)003

b_s, where the parameter dependence comes at leading order for M, in contrast to P where they arise at one-loop. The b_2-degeneracy breaking power of the bispectrum has long been known to the community [28–30, 52] and has been confirmed for the marked spectrum in ref. [32]. Here, we will develop this insight further.

The leading order contribution to (the bispectrum term in) the MPS comes in three contractions

$$\left\langle \delta^{(1)}_{g,R}(\mathbf{x_1})\delta^{(2)}_{g}(\mathbf{x_1})\delta^{(1)}_{g}(\mathbf{x_2})\right\rangle, \quad \left\langle \delta^{(2)}_{g,R}(\mathbf{x_1})\delta^{(1)}_{g}(\mathbf{x_1})\delta^{(1)}_{g}(\mathbf{x_2})\right\rangle, \quad \left\langle \delta^{(1)}_{g,R}(\mathbf{x_1})\delta^{(1)}_{g}(\mathbf{x_1})\delta^{(2)}_{g}(\mathbf{x_2})\right\rangle \tag{2.35}$$

which are different integrations over B_{112}, as shown in section 2.2.

Let us start with the first term which, following eq. (2.14) for $\delta^{(n)}$, can be expanded as

$$\left\langle \delta^{(1)}_{g,R,1}\delta^{(2)}_{g,1}\delta^{(1)}_{g,2}\right\rangle_{\rm FT} = \int_{\mathbf{p}} Z_2(\mathbf{k},\mathbf{p})Z_1(\mathbf{k})Z_1(\mathbf{p})W_R(p)P_L(k)P_L(p) \tag{2.36}$$

where we have abbreviated the coordinates to subscripts on the fields. The only b_2 dependence in this term is that sourced by $Z_2 \supset b_2/2$, yielding

$$\left\langle \delta^{(1)}_{g,R,1}\delta^{(2)}_{g,1}\delta^{(1)}_{g,2}\right\rangle_{\rm FT} \supset Z_1(\mathbf{k})P_L(k)\,\frac{b_2}{2}\int_{\mathbf{p}} Z_1(\mathbf{p})W_R(p)P_L(p) \tag{2.37}$$

$$= \frac{b_2}{2}\left(b_1+\frac{f}{3}\right)Z_1(\mathbf{k})P_L(k)\int\frac{p^2\,dp}{2\pi^2}W_R(p)P_L(p) \tag{2.38}$$

which again highlights the importance of the smoothing, W_R, as the expression would otherwise be UV-sensitive.

The second contribution $\left\langle \delta^{(2)}_{g,R,1}\delta^{(1)}_{g,1}\delta^{(1)}_{g,2}\right\rangle$ merely differs from the derivation above by the argument of the smoothing, which will now be $W_R(|\mathbf{k}+\mathbf{p}|)$. The third contribution and its b_2 dependence can be expanded as

$$\left\langle \delta^{(1)}_{g,R,1}\delta^{(1)}_{g,1}\delta^{(2)}_{g,2}\right\rangle_{\rm FT} = \int_{\mathbf{p}} Z_2(\mathbf{k}-\mathbf{p},\mathbf{p})Z_1(\mathbf{k}-\mathbf{p})Z_1(\mathbf{p})W_R(p)P_L(|\mathbf{k}-\mathbf{p}|)P_L(p) \tag{2.39}$$

$$\supset \frac{b_2}{2}\int_{\mathbf{p}} Z_1(\mathbf{k}-\mathbf{p})Z_1(\mathbf{p})W_R(p)P_L(|\mathbf{k}-\mathbf{p}|)P_L(p) \tag{2.40}$$

which in the low-k limit ($k\to 0$) contains

$$\frac{b_2}{2}\left(b_1^2+\frac{2b_1 f}{3}+\frac{f^2}{5}\right)\int\frac{p^2\,dp}{2\pi^2}W_R(p)P_L^2(p) \tag{2.41}$$

having contributions $\propto b_2$, b_1b_2, and $b_1^2b_2$. These contributions are distinct from that in the power spectrum

$$P_{22} = \left\langle\left(\delta^{(2)}_g\right)^2\right\rangle = 2\int_{\mathbf{p}} Z_2^2(\mathbf{p},\mathbf{k}-\mathbf{p})P_L(k)P_L(|\mathbf{k}-\mathbf{p}|) \tag{2.42}$$

$$P_{13} = \left\langle\delta^{(1)}_g\delta^{(3)}_g\right\rangle = 3\int_{\mathbf{p}} Z_3(\mathbf{k},\mathbf{p},-\mathbf{p})Z_1(\mathbf{k})P_L(k)P_L(p) \tag{2.43}$$

which have a different set of kernels with different momenta as arguments. In particular, the P_{22} term produces a b_2^2 contribution that cannot appear in the marked spectrum. Additionally,

JCAP08(2026)003

the b_2 dependence of P, which only enters at one-loop, is subdominant to linear theory terms, whereas in M there is a b_2 dependence in every leading-order contraction.

Similarly, b_s contributes to Z_2 at leading order and the b_s dependence can be captured by substituting the b_2 term in Z_2 above by the b_s term ($b_s[(\hat{k}_1 \cdot \hat{k}_2)^2 - 1/3]$). For example, the b_s dependence of the monopole of $\left\langle \delta^{(1)}_{g,R,1}\delta^{(2)}_{g,1}\delta^{(1)}_{g,2}\right\rangle$ is

$$\left\langle \delta^{(1)}_{g,R,1}\delta^{(2)}_{g,1}\delta^{(1)}_{g,2}\right\rangle_0 \supset \frac{8f^2}{675} b_s P_L(k) \int \frac{p^2\,dp}{2\pi^2} W_R(p) P_L(p) \tag{2.44}$$

For $b_1 = 2$ and $f \sim 0.9$, the monopole prefactor here is weaker than that of b_2 in eq. (2.38) by a factor of ~ 550, which indicates that using M_0 to detect b_s may be difficult, as implied later in figure 2. The expression also indicates that this b_s dependence is independent of the other nuisance parameters, but this is not a general result. For the second contraction $\left\langle \delta^{(2)}_{g,R,1}\delta^{(1)}_{g,1}\delta^{(1)}_{g,2}\right\rangle$ the 'extra' exponential introduced by the change in smoothing argument ($W_R(p) \to W_R(|\mathbf{k}+\mathbf{p}|)$) breaks the result above, although it can be recovered in the low-k limit. Finally the b_s dependence of the third contraction is

$$\left\langle \delta^{(1)}_{g,R,1}\delta^{(1)}_{g,1}\delta^{(2)}_{g,2}\right\rangle_{\rm FT} \supset \int_{\mathbf{p}} b_s \left(\frac{((\mathbf{k}-\mathbf{p})\cdot\mathbf{p})^2}{|\mathbf{k}-\mathbf{p}|^2 p^2} - \frac{1}{3}\right) Z_1(\mathbf{k}-\mathbf{p}) Z_1(\mathbf{p}) W_R(p) P_L(|\mathbf{k}-\mathbf{p}|) P_L(p) \tag{2.45}$$

which, as $k \to 0$, reduces to

$$b_s \left(b_1^2 + \frac{2b_1 f}{3} + \frac{f^2}{5}\right) \int \frac{p^2\,dp}{2\pi^2} W_R(p) P_L^2(p) \tag{2.46}$$

taking a similar form to the dependency of b_2 in the same term (eq. (2.41)).

Figure 2 shows the potential of M_0 for breaking degeneracy in practice for a tracer with LRG-like biases ($b_1 = 2.07$, $b_2 = 0.43$, $b_s = 1.18$; fit to DESI DR1 cutsky mocks section 5.3) at $z = 0.7$. The left panel shows the effect of a $\sim 1\sigma$ deviation of b_2 and b_s in current and near-future surveys ($\Delta b_2 = 2$, $\Delta b_s = 1$; see e.g. table VII of ref. [50]). We first observe that, as expected from preliminary results in the calculation above, the b_2 dependence is considerably stronger than that of b_s, indicating that b_2 will likely be the nuisance parameter of interest when employing the MPS, although the inclusion of the quadrupole M_2 may assist detecting b_s deviations as well as other angular bispectrum signals. The figure further shows that this 'signal' dominates against both observational uncertainty ($\sigma_{\rm sim}$) and model uncertainty from higher-order corrections ($\pm A_0 P_L$) both by over an order of magnitude, showing certain potential for employment. The latter, in particular, is crucial, as the scale-dependence of the signal is largely degenerate with $P_L(k)$ and would be challenging to exploit without control over A_0. While here we find that the scale-dependence of ΔM_0 resembles that of P_L, we also find that M_{13} has a similar scale-dependence. This may be of interest in the future when making theoretical insights and has been explored in appendix C. Now for the right panel let us consider two sets of biases: the LRG-like bias above ($\theta_{\rm LRG}$) and another ($\theta_{\rm fit}$) that returns an indistinguishable power spectrum but with a b_2 lower by 2. While $\Delta b_2 = 2$ here as in the left panel, it is a priori unclear whether the strong b_2 dependency seen in the left panel will fully manifest itself under this circumstance as we have altered other nuisance

JCAP08(2026)003

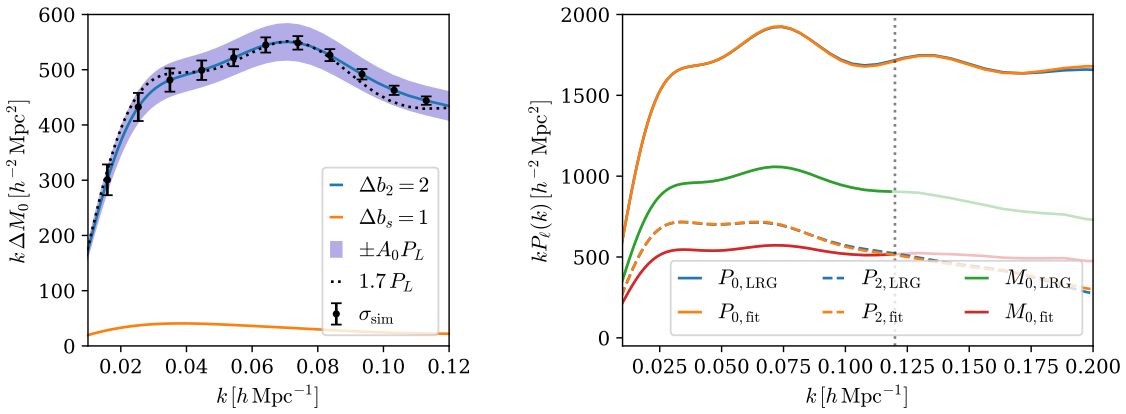


Figure 2. *Left*: the change in M_0 (with $R = 15h^{-1}\,\mathrm{Mpc}$ at $z = 0.7$) from shifts in second-order bias b_2 (blue line; $\Delta b_2 = 2$) and shear bias b_s (orange; $\Delta b_s = 1$) of size equal to the $1\,\sigma$ uncertainty from current generation surveys [50]. The b_2 dependence dominates despite both biases being quadratic in δ. Overlaid on the b_2 curve is the error observed from $8\,h^{-3}\,\mathrm{Gpc}^3$ simulations (black error bars; section 5.1) and the anticipated maximum effect of higher-loop contributions (blue shade; section 2.3) with $A_0 \lesssim 0.11$, both of which are subdominant to the expected b_2 change. The dotted black line shows 1.7 times the linear power spectrum P_L demonstrating both that the shape of ΔM_0 is degenerate with P_L (appendix C) and that the higher-loop correction $A_0 P_L$ is subdominant by over an order of magnitude to the b_2 change. *Right*: the power spectrum and $R = 15h^{-1}\,\mathrm{Mpc}$ marked power spectrum of two nuisance parameter sets $\{\theta_{\rm LRG}\}$ and $\{\theta_{\rm fit}\}$ at $z = 0.7$. $\theta_{\rm LRG}$ is the LRG-like bias parameters that fit the DESI DR1 cutsky mock at $z = 0.7$ (section 5.3) with $b_2 \approx 0.5$, while $\theta_{\rm fit}$ is another nuisance set with nearly identical power spectrum predictions but with $b_2 = -1.5$. The similarity in $P_\ell(k)$ captures the $\Delta b_2 \sim 2$ fully expected in near-term redshift surveys, while $M_0(k)$ differs by a factor of 2, showing its potential to assist standard two-point analyses. The dotted vertical line indicates the $k^M_{\rm max} = 0.12\,h^{-1}\,\mathrm{Mpc}$ used for this work.

parameters as well. The figure, however, demonstrates that this is not of significant concern as the amplitude of M_0 deviates by a factor of 2, while P_0 and P_2 are indistinguishable.

The two panels also offer some insight into what the binning of M in data should be, as long as the b_2 signal is the principle aim. Both panels show that the b_2 dependence is smooth and weakly scale-dependent. This indicates that the precise values of $k_{\rm min}$ and $k_{\rm max}$ are not important, as this would only marginally change the statistical constraining power and would not reveal new information. This also suggests that a coarse binning will suffice, which is beneficial in terms of the covariance matrix measurement.

3 Covariance matrix

To combine the power spectrum multipoles and the MPS we require not only an estimate of the covariance of each set of spectra, but also their cross-covariance. As the MPS has a significantly smaller data vector size than the bispectrum, we expect it will be possible to estimate the total covariance matrix between P_ℓ and M_0 using simulations, as was done for the recent power spectrum analysis of DESI DR1 full-shape data [6]. Regardless, it will be useful to gain analytic insight into the covariance structure. To this end, here we will

JCAP08(2026)003

discuss the covariance structure in real-space and neglect the effects of survey geometry for brevity, although extending to redshift-space will be straightforward in principle as described briefly in appendix D.

We start by defining the N-point correlator

$$\langle \delta(\mathbf{k}_1) \dots \delta(\mathbf{k}_N) \rangle = (2\pi)^3 \delta_D \left(\sum_i \mathbf{k}_i \right) P_N(\mathbf{k}_1, \dots, \mathbf{k}_N) \tag{3.1}$$

where δ_D is the Dirac delta function. As the following discussion will be centered around discrete k-bins with width Δk, it will also be useful to define the mode-counting factor, i.e. the volume of each (thin) shell

$$V_k = \int_{V_k} d^3q \simeq 4\pi \int_{k-\Delta k/2}^{k+\Delta k/2} q^2 dq \approx 4\pi k^2 \Delta k \tag{3.2}$$

and the translation between continuous and discrete δ-functions

$$(2\pi)^3 \delta_D(\mathbf{k}_1 - \mathbf{k}_2) \leftrightarrow V \delta^K_{k_1,k_2} \tag{3.3}$$

where V is the survey volume and δ^K is the Kronecker delta function. Using these definitions, the estimators for P and M become

$$\widehat{P}_{\mathbf{k}} = V^{-1} \int_{V_k} \frac{d^3q}{V_k} \delta(-\mathbf{q}) \delta(\mathbf{q}), \qquad \widehat{M}_{\mathbf{k}} = V^{-1} \int_{V_k} \frac{d^3q}{V_k} \delta(-\mathbf{q}) \tilde{\delta}_M(\mathbf{q}) \tag{3.4}$$

where we have defined $\tilde{\delta}_M(\mathbf{q}) = [\delta_{g,R} \star \delta_g](\mathbf{q})$ as a shorthand for the relevant contribution to the Fourier transform of δ_M. Throughout this section we will also utilize the abbreviation $f(\mathbf{k}) = f_{\mathbf{k}}$ to simplify the expressions.

3.1 Covariance of M

The covariance of a bin-averaged marked spectrum can be calculated as

$$\mathrm{Cov}\left(\widehat{M}_k, \widehat{M}_{k'}\right) = \left\langle \widehat{M}_k \widehat{M}_{k'} \right\rangle - \left\langle \widehat{M}_k \right\rangle \left\langle \widehat{M}_{k'} \right\rangle. \tag{3.5}$$

The general form of this auto-covariance is largely similar to that for the filtered-square bispectrum [21], although the smoothing kernel and the number of smoothed fields are different. The latter will appear as an increase in the number of terms due to a loss of symmetry between the two fields used to construct the composite field (as $\tilde{\delta}_M = \delta_{g,R}\delta_g$ and $\delta_{g,R} \neq \delta_g$ in this work). We will calculate the cross-covariance between P and M in section 3.2 after the auto-covariance.

Expanding the expression above we obtain

$$\mathrm{Cov}\left(\widehat{M}_k, \widehat{M}_{k'}\right) = V^{-2} \int_{V_k} \frac{d^3q}{V_k} \int_{V_{k'}} \frac{d^3q'}{V_{k'}} \int_{\mathbf{p}} \int_{\mathbf{p}'} W_R(p) W_R(p') \tag{3.6}$$

$$\left(\left\langle \delta_{-\mathbf{q}} \delta_{\mathbf{p}} \delta_{\mathbf{q}-\mathbf{p}} \delta_{-\mathbf{q}'} \delta_{\mathbf{p}'} \delta_{\mathbf{q}'-\mathbf{p}'} \right\rangle - \left\langle \delta_{-\mathbf{q}} \delta_{\mathbf{p}} \delta_{\mathbf{q}-\mathbf{p}} \right\rangle \left\langle \delta_{-\mathbf{q}'} \delta_{\mathbf{p}'} \delta_{\mathbf{q}'-\mathbf{p}'} \right\rangle \right) \tag{3.7}$$

To calculate the whole covariance, we must consider all possible contractions in the six-field correlator. As each δ is a mean-zero field, contractions with $\langle \delta \rangle$ can be excluded.

JCAP08(2026)003

This leaves us with four possible categories of contractions: $(2,2,2)$ (e.g. $\langle ab\rangle\langle cd\rangle\langle ef\rangle$), $(3,3)$ (e.g. $\langle abc\rangle\langle def\rangle$), $(4,2)$ (e.g. $\langle abcd\rangle\langle ef\rangle$), and (6) (e.g. $\langle abcdef\rangle$). These terms are traditionally referred to as PPP, BB, PT, and P_6, respectively, in bispectrum covariance calculations [53]. We run through the contractions for each category in appendix D.

The calculations reveal that the covariance can be organized into diagonal (D_{MM}) and non-diagonal (N_{MM}) terms in the following fashion

$$\mathrm{Cov}\left(\widehat{M}_k, \widehat{M}_{k'}\right) = (2\pi)^3 \delta^K_{k,k'} V^{-1} V_k^{-1} D_{MM} + V^{-1} N_{MM} \ . \tag{3.8}$$

The inverse factor of $V_k \propto k^2$ accompanying D_{MM} determines that it dominates at large-scales, where we have the best theoretical control. The diagonal term can be further factorized into each contraction type (PPP, BB, PT)

$$D_{MM} = D_{222} + D_{33} + D_{42} \tag{3.9}$$

where

$$D_{222}(\mathbf{k}) = \int_{\mathbf{p}} W_R(p)\left[W_R(p) + W_R(|\mathbf{k}-\mathbf{p}|)\right] P(\mathbf{k}) P(\mathbf{p}) P(\mathbf{k}-\mathbf{p}) \tag{3.10}$$

$$D_{33}(\mathbf{k}) = M^2(\mathbf{k}) \tag{3.11}$$

$$D_{42}(\mathbf{k}) = \int_{\mathbf{p}} \int_{\mathbf{p}'} W_R(p) W_R(p') P(\mathbf{k}) T(\mathbf{p}, \mathbf{k}-\mathbf{p}, \mathbf{p}', -\mathbf{k}-\mathbf{p}') \tag{3.12}$$

This full diagonal contribution can be recovered from the Gaussian (disconnected) covariance, i.e.

$$\mathrm{Cov}\left(\widehat{M}_k, \widehat{M}'_k\right) \supset \mathrm{Cov}_G\left(\widehat{M}_k, \widehat{M}_{k'}\right) = (2\pi)^3 \delta^K_{k,k'} V^{-1} V_k^{-1} \left[D_{22} + D_{33} + D_{42}\right] \tag{3.13}$$

indicating that for large-scale covariances the Gaussian approximation is sufficient. This behavior is anticipated and has been demonstrated previously for the filtered-square bispectrum covariance [21]; we further confirm this agreement along the covariance diagonal using the 25 periodic box mocks employed in this work (section 5.1). We note, however, that for the skew spectrum of type $\langle[\delta_R^2]\delta\rangle$, ref. [54] reports an appreciable SNR deviation of $\lesssim 50\%$ when the covariance is restricted to its diagonal entries. A precision study of the MPS covariance requires a larger simulation suite than is available here and therefore lies beyond the scope of this work, but revisiting this question in the future will be important.

3.2 Covariance of *P*-*M*

Now let us shift our attention to the covariance between P and M,

$$\mathrm{Cov}\left(\widehat{P}_k, \widehat{M}_{k'}\right) = V^{-2} \int_{V_k} \frac{d^3k}{V_k} \int_{V_k} \frac{d^3k'}{V_k} \int_{\mathbf{p}'} W_R(p') \tag{3.14}$$

$$\left(\left\langle \delta_{-\mathbf{q}} \delta_{\mathbf{q}} \delta_{-\mathbf{q}'} \delta_{\mathbf{p}'} \delta_{\mathbf{q}'-\mathbf{p}'} \right\rangle - \left\langle \delta_{-\mathbf{q}} \delta_{\mathbf{q}} \right\rangle \left\langle \delta_{-\mathbf{q}'} \delta_{\mathbf{p}'} \delta_{\mathbf{q}'-\mathbf{p}'} \right\rangle\right) \tag{3.15}$$

With five density field operators, the possible contractions are of the form $(3,2)$ or (5). While we refer the reader to appendix D for a detailed calculation, we once again find a factorization of covariance into diagonal (D_{PM}) and non-diagonal (N_{PN}) terms

$$\mathrm{Cov}\left(\widehat{P}_k, \widehat{M}_{k'}\right) = (2\pi)^3 \delta^K_{k,k'} V^{-1} V_k^{-1} D_{PM} + V^{-1} N_{PM} \ . \tag{3.16}$$

JCAP08(2026)003

The inverse mode-counting factor, $V_k^{-1} \propto k^{-2}$, again indicates that D_{PM} dominates at large scales. This diagonal term only consists of contributions from the (3,2) contraction

$$D_{PM}(\mathbf{k}) = D_{23}(\mathbf{k}) = 2P(\mathbf{k})M(\mathbf{k}) \ . \tag{3.17}$$

Similar to the covariance, we find that the Gaussian approximation recovers the full diagonal contribution, i.e.

$$\mathrm{Cov}_G\left(\widehat{P}_k, \widehat{M}_{k'}\right) = (2\pi)^3 \delta^K_{k,k'} V^{-1} V_k^{-1} D_{23} \tag{3.18}$$

Due to the similarity in structure, this result carries over to other two-field composite fields (e.g. the filtered square bispectrum) as well by a substitution of smoothing kernels.

3.3 Scaling and relation to the bispectrum

For the bispectrum, the cross-covariance is known to be subdominant to the auto-covariances for general triangles with sides of similar length ($k_1 \sim k_2 \sim k_3$), becoming important only for particular configurations such as the squeezed bispectrum ($k_1 \ll k_2 \simeq k_3$) [53, 55]. However, recent work shows that the correlations between the monopoles of the power spectrum and bispectrum are large ($\gtrsim 0.5$) at quasi-linear scales ($k \sim 0.1\,h\,\mathrm{Mpc}^{-1}$) suggesting that the cross-covariance is not negligible for all compressions of the bispectrum [56]. If the cross-covariance is subdominant, it is both a convenient and scientifically motivating feature, as a smaller cross-covariance decreases the uncertainty requirement on measurement and covariance matrix, and qualitatively indicates that there is more 'new' information in the probe. We will investigate how this situation translates to the MPS.

The small cross-covariance of the bispectrum can be traced back to the difference in the mode-counting factor normalizing the estimators. For bins of width $\Delta k \ll k_i$, the volume of triangles included in the bispectrum is

$$V_{\mathrm{tr}} \simeq \int_{\mathbf{q}_1 \in k_1} d^3q_1 \int_{\mathbf{q}_2 \in k_2} d^3q_2 \int_{\mathbf{q}_3 \in k_3} d^3q_3 \ \delta_D(\mathbf{q}_{123}) \simeq 8\pi^2 k_1 k_2 k_3 \Delta k^3 \tag{3.19}$$

where $\int_{\mathbf{q}\in k} d^3q = \int d\Omega_q \int_{k-\Delta k/2}^{k+\Delta k/2} q^2 dq$ and $\mathbf{q}_{123} = \mathbf{q}_1 + \mathbf{q}_2 + \mathbf{q}_3$ [53, 57]. Similar to the power spectrum, where $\widehat{P} \propto V_k^{-1}$, the bispectrum estimator scales as $\widehat{B} \propto V_{\mathrm{tr}}^{-1}$.

This mode-counting factor enters the (auto-)covariance as

$$\mathrm{Cov}\left[\widehat{B}(t), \widehat{B}(t')\right] \simeq \mathrm{Cov}\left[\widehat{B}(t), \widehat{B}(t')\right]_{222} \propto V_{\mathrm{tr}}^{-2} V_{\mathrm{tr}} P(k_1)P(k_2)P(k_3) \tag{3.20}$$

where V_{tr}^{-2} comes from the two powers of $\widehat{B}$ and the V_{tr} from the integrals present after performing momentum-matching using δ_D's [53, 55]. This is distinct from the cross-covariance

$$\mathrm{Cov}\left[\widehat{P}(k), \widehat{B}(t')\right] \simeq \mathrm{Cov}\left[\widehat{P}(k), \widehat{B}(t')\right]_{32} \propto V_{\mathrm{tr}}^{-1} V_k^{-1} V_{\mathrm{tr}} P(k) B(k_1, k_2, k_3) \tag{3.21}$$

where the estimators $\widehat{P}$ and $\widehat{B}$ contribute V_k^{-1} and V_{tr}^{-1}, respectively, and the integrals after momentum-matching contribute V_{tr} [55]. This results in an overall volume factor of V_k^{-1},

JCAP08(2026)003

which is distinct from the auto-covariance scaling as $V_{\rm tr}^{-1}$. This is the key reason for the cross-covariance suppression of the bispectrum, as

$$\frac{\mathrm{Cov}\left[\widehat{P},\widehat{B}\right]}{\sqrt{\mathrm{Cov}\left[\widehat{P},\widehat{P}\right]\mathrm{Cov}\left[\widehat{B},\widehat{B}\right]}} \sim \frac{V_k^{-1}PB}{\sqrt{V_k^{-1}V_{\rm tr}^{-1}P^5}} \sim \frac{\Delta k}{k}\sqrt{k^3P} \tag{3.22}$$

where we assume $k_1 \sim k_2 \sim k_3 \sim k$ and approximate $B(k,k,k) \sim P^2(k)$, and k^3P is dimensionless [53]. Hence, when $\Delta k \ll k$, the cross-covariance is suppressed by $\Delta k/k$. Since this suppression is a consequence of the difference in the mode-counting factor, the cross-covariance is 'unsuppressed' as one considers larger k-bins (e.g. $\Delta k \sim k$) and allow more triangles configurations to contribute, indicating that it may be misguided to use the suppression of cross-covariance per bin as a heuristic of new scientific information. In what follows, we will see an analogous situation for the MPS.

To trace the mode-counting factor of the MPS, let us define

$$V_M = \int_{\mathbf{q}\in k}\int_{\mathbf{p}_1}\int_{\mathbf{p}_2} W_R(p_1)\delta_D(\mathbf{q}+\mathbf{p}_1+\mathbf{p}_2) \tag{3.23}$$

For the purpose of this exercise, it will also be useful to define a narrow smoothing kernel[8]

$$W_{\mathrm{hat},p_*}(k) = \begin{cases} 1 & p_* - \frac{\Delta p}{2} < k < p_* + \frac{\Delta p}{2} \\ 0 & \text{otherwise} \end{cases} \tag{3.24}$$

where the mode-counting factor becomes $V_M \approx 16\pi^2 k^2 p_*^2 \Delta k \Delta p$. In what follows we will use this narrow kernel when explicitly invoking the kernel to make the mode-counting factor more transparent. In terms of mode-counting, the broad Gaussian kernel used for this work can roughly be recovered when $\Delta p \sim p_*$.

Let us now consider the dominant diagonal contribution to the cross-covariance D_{23} (eq. (D.105))

$$\begin{aligned}(2\pi)^3\delta^K_{k,k'}V^{-1}V_k^{-1}D_{23} &= (2\pi)^3V^{-1}V_k^{-1}\delta^K_{k,k'}P(\mathbf{k})M(\mathbf{k}) &(3.25)\\ &= (2\pi)^3V^{-1}V_k^{-1}\delta^K_{k,k'}P(\mathbf{k})\int_{\mathbf{p}} W_R(p)B(\mathbf{p},\mathbf{k}-\mathbf{p},-\mathbf{k}) &(3.26)\\ &= V^{-1}V_k^{-2}V_M\delta^K_{k,k'}P(\mathbf{k})\langle B(\mathbf{p}_*,\mathbf{k}-\mathbf{p}_*,-\mathbf{k})\rangle_\varphi &(3.27)\end{aligned}$$

where at the final step we apply the thin-shell approximation $B(\mathbf{p}',\mathbf{q}-\mathbf{p}',-\mathbf{q}) = B(\mathbf{p}_*,\mathbf{k}-\mathbf{p}_*,-\mathbf{k})$. Similarly for the auto-covariance, the (2,2,2) diagonal term D_{222} (eq. (D.65)) scales as

$$\begin{aligned}&(2\pi)^3\delta^K_{k,k'}V^{-1}V_k^{-1}D_{222}\\ &= (2\pi)^3\delta^K_{k,k'}V^{-1}V_k^{-1}\int_{\mathbf{p}} W_R(p)\left[W_R(p)+W_R(|\mathbf{k}-\mathbf{p}|)\right]P(\mathbf{k})P(\mathbf{p})P(\mathbf{k}-\mathbf{p}) &(3.28)\\ &\approx V^{-1}V_k^{-2}V_M\delta^K_{k,k'}P(\mathbf{k})P(\mathbf{p}_*)\langle P(\mathbf{k}-\mathbf{p}_*)\rangle_\varphi &(3.29)\end{aligned}$$

[8] For example, one could imagine constructing this by using as a mark the difference between two densities with similar but unequal smoothing lengths.

JCAP08(2026)003

Note that we have not shown that the (2,2,2) term dominates out of the three diagonal terms, but we consider it to draw parallels with the bispectrum case.

Thus, if we adopt the approximation $p_* \sim k \sim |\mathbf{k} - \mathbf{p}_*|$ and (crudely) approximate a constant contribution over the angle average $\langle f \rangle_\varphi$, the ratio between the cross- and auto-covariances scale as

$$\frac{\mathrm{Cov}\left[\widehat{P}, \widehat{M}\right]}{\sqrt{\mathrm{Cov}\left[\widehat{P}, \widehat{P}\right] \mathrm{Cov}\left[\widehat{M}, \widehat{M}\right]}} \sim \frac{V_k^{-2} V_M B}{\sqrt{V_k^{-3} V_M P^5}} \sim \frac{V_M^{1/2} P^{1/2}}{V_k^{1/2}} = \sqrt{\frac{k^2 p_*^2 \Delta k \Delta p\, P}{k^2 \Delta k}} = \frac{p_*}{k} \sqrt{\frac{\Delta p}{k}} \sqrt{k^3 P} \tag{3.30}$$

showing a suppression of the cross-covariance by $\sqrt{\Delta p/k}$ as opposed to the $\Delta k/k$ of the bispectrum. By extension, we see that the MPS considered in this work, with a Gaussian smoothing kernel W_R, corresponds to the case $\Delta p \sim k$, with little suppression. This is a direct consequence of the change in mode-counting factors. Simply put, the (wide) integration over the inner momentum $\mathbf{p}$ in M means that there are far more triangles that contribute, whereas for B only specific triangle configurations give matches. This is analogous to the suppression in $P - B$ cross-covariance becoming weaker when considering wider k-bins and allowing more triangles to contribute. Note that a wide window breaks the thin-shell approximation applied above and thus is merely an analogy valid for order-of-magnitude estimates. This result is in qualitative agreement with past work investigating the cross-covariance of skew spectra with the power spectrum, which found weaker but significant correlations [54, 58].

A stronger suppression with a narrow window indicates that there is more "new" information when using a narrow window and will be less affected from data uncertainties when computing the inverse covariance. Of course, this is in trade-off with other factors influenced by the window choice, such as the SNR, degeneracy breaking in the whole parameter space, and any numerical uncertainties and systematics dependence when using sharp windows with significant ringing in real-space. In particular, it is expected that the SNR will be impacted, as one is decreasing the number of data points that contribute.

A full investigation of the results above is outside of the scope of this work, as it will require a large suite of simulations considering specific survey footprints. At the present stage, we still conduct a preliminary validation of the results using 25 mock catalogs in $2\,h^{-1}\,\mathrm{Gpc}$ periodic boxes, which are later introduced in section 5.1. We indeed find that the Gaussian contribution dominates and the cross-covariance is not significantly suppressed relative to the auto-covariances, indicating that we are in a qualitatively different situation than when considering the joint analyses of P and B.

4 The effect of survey geometry

Previous calculations and applications of the MPS were limited to those on periodic boxes.[9] However, real galaxy surveys observe only a fraction of the sky, requiring an estimator that accounts for the (often complex) survey footprint. This technique is established for power spectrum calculations and we demonstrate how this can be carried over directly to the MPS.

[9] Ref. [33] has applied the marked angular power spectrum to weak lensing data, but did not model the window effect as it used an emulator over simulations as the theoretical prediction.

JCAP08(2026)003

4.1 Power spectrum

The modeling of the power spectrum over a part of the sky is well-established [59]. Observationally, the galaxy distribution is captured by the weighted galaxy density, n_g, and it is compared to a weighted random density, n_r, which captures the survey geometry and mean density variation without any clustering. The randoms thus encode information about the survey footprint, which can be defined as a survey geometry selection function $W(\mathbf{x})$ through ensemble averages over the two densities

$$W(\mathbf{x}) = \langle n_g(\mathbf{x}) \rangle = \alpha \langle n_r(\mathbf{x}) \rangle \quad \text{with} \quad \alpha = \frac{\int d\mathbf{x}\; n_g(\mathbf{x})}{\int d\mathbf{x}\; n_r(\mathbf{x})} \,. \tag{4.1}$$

Using these ingredients, the power spectrum multipoles over the survey footprint can be estimated by [60]

$$\widehat{P}_\ell(k_\mu) = \frac{2\ell+1}{AV_{k_\mu}} \int_{V_{k_\mu}} \mathrm{d}\mathbf{k} \int \mathrm{d}\mathbf{x}_1 \int \mathrm{d}\mathbf{x}_2\, e^{i\mathbf{k}\,\cdot\,(\mathbf{x}_2-\mathbf{x}_1)} \mathcal{F}(\mathbf{x}_1)\mathcal{F}(\mathbf{x}_2)\mathcal{L}_\ell(\widehat{\mathbf{k}}\,\cdot\,\widehat{\mathbf{x}}_1) - \mathcal{N}_\ell \tag{4.2}$$

where $\mathcal{F}(\mathbf{x}) = n_g(\mathbf{x}) - \alpha n_r(\mathbf{x})$ is the FKP field [59], V_{k_μ} is the volume of the shell of $\mathbf{k}$ being integrated over, $\mathcal{L}_\ell$ is the Legendre polynomial of order ℓ and $\mathcal{N}_\ell$ is the estimated shot-noise contribution. The normalization, A, is conventionally taken to be[10]

$$A = \alpha \int d\mathbf{x}\; n_g(\mathbf{x}) n_r(\mathbf{x}) \approx \int d\mathbf{x}\; \langle n_g(\mathbf{x}) \rangle^2 \tag{4.3}$$

where $\langle n_g(\mathbf{x}) \rangle$ corresponds to the mean density.

For Poisson shot-noise $\langle \mathcal{F}(\mathbf{x})\mathcal{F}(\mathbf{x}') \rangle = W(\mathbf{x})W(\mathbf{x}')\xi(\mathbf{x},\mathbf{x}') + W(\mathbf{x})\delta_D(\mathbf{x}-\mathbf{x}')$ [59]. The δ_D term is subtracted by $\mathcal{N}_\ell$ above and so for a thin bin in k the expectation of $\widehat{P}_\ell$ becomes [61, 62]

$$\left\langle \widehat{P}_\ell(k) \right\rangle = \frac{2\ell+1}{A} \int \frac{d\Omega_k}{4\pi} \int d\mathbf{x}_1 \int d\mathbf{x}_2\, e^{i\mathbf{k}\,\cdot\,(\mathbf{x}_2-\mathbf{x}_1)}\, W(\mathbf{x}_1)\, W(\mathbf{x}_2)\, \mathcal{L}_\ell\Big(\hat{\mathbf{k}}\,\cdot\,\hat{\mathbf{x}}_1\Big)\; \xi(\mathbf{x}_1,\mathbf{x}_2) \tag{4.4}$$

$$= \frac{2\ell+1}{A} \sum_p \int \frac{d\Omega_k}{4\pi} \int d\mathbf{x} \int d\mathbf{s}\, e^{-i\mathbf{k}\,\cdot\,\mathbf{s}} W(\mathbf{x})W(\mathbf{x}-\mathbf{s}) \mathcal{L}_\ell\Big(\hat{\mathbf{k}}\,\cdot\,\hat{\mathbf{x}}\Big)\, \mathcal{L}_p(\hat{\mathbf{x}}\,\cdot\,\hat{\mathbf{s}})\, \xi_p(s) \tag{4.5}$$

$$= 4\pi(-i)^\ell(2\ell+1) \sum_{\ell_1,\ell_2} \begin{pmatrix} \ell_1 & \ell_2 & \ell \\ 0 & 0 & 0 \end{pmatrix}^2 \int s^2\, ds\; \xi_{\ell_1}(s)\, \mathcal{W}_{\ell_2}(s)\, j_\ell(ks) \tag{4.6}$$

where we define $\mathbf{s} = \mathbf{x}_1 - \mathbf{x}_2$ and the real space window matrix $\mathcal{W}$ is defined as

$$\mathcal{W}_\ell(s) = \frac{2\ell+1}{4\pi A} \int d\Omega_s \int d\mathbf{x}\; W(\mathbf{x})W(\mathbf{x}-\mathbf{s}) \mathcal{L}_\ell(\hat{\mathbf{x}}\,\cdot\,\hat{\mathbf{s}}) \tag{4.7}$$

While this does not include wide-angle corrections to the plane-parallel approximation, these can be incorporating by expanding the correlation function ξ in powers of s/d, where $\mathbf{s}$ is the pair separation and $\mathbf{d}$ is the line-of-sight distance [61, 62]. The window matrix at each order then becomes

$$\mathcal{W}^{(n)}_\ell = \frac{2\ell+1}{4\pi A} \int d\Omega_s \int d\mathbf{x}\; x^{-n} W(\mathbf{x})W(\mathbf{x}-\mathbf{s}) \mathcal{L}_\ell(\hat{\mathbf{x}}\,\cdot\,\hat{\mathbf{s}}) \tag{4.8}$$

[10] The product of n_g and n_r is used instead of the product of two n_r's to avoid a shot-noise bias. See https://pypower.readthedocs.io/en/latest/api/api.html#pypower.fft_power.normalization for a numerical implementation.

JCAP08(2026)003

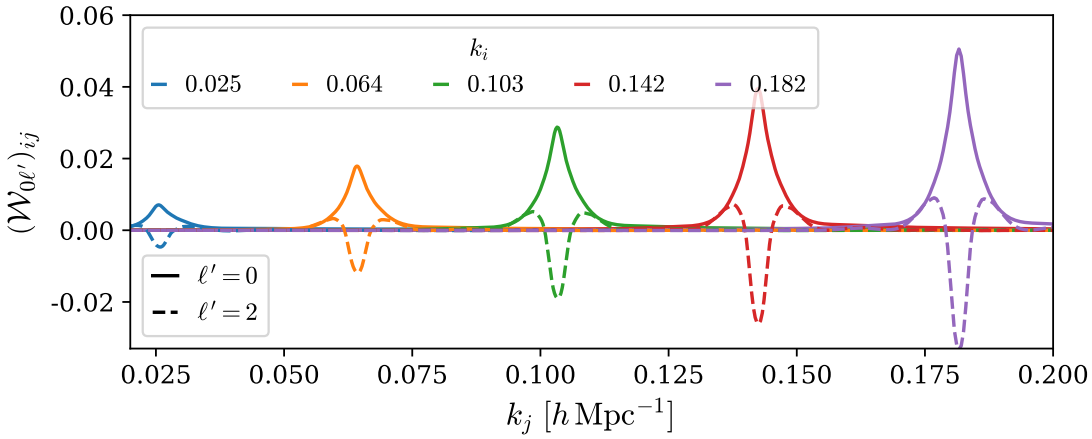


Figure 3. Some pieces of the Fourier-space DR1 cutsky window matrix $(\mathcal{W}_{\ell\ell'})_{ij}$ (eqs. (4.9) and (4.13)). We choose $\ell = 0$ and select k_i, for bins of width $\Delta k \sim 0.01$. The real-space window matrix is shown in figure 6 of ref. [63].

Following ref. [63] we will include the first order correction ($n = 1$). The convolved power spectrum over a finite k-bin, k_i, can then be modeled as a matrix multiplication of the window matrix and the theory model

$$(\widehat{P}_\ell^{\rm obs})_i = (\mathcal{W}_{\ell\ell'})_{ij}(P_{\ell'})_j \ . \tag{4.9}$$

We show in figure 3 some slices of the window matrix, $\mathcal{W}$, for the DESI DR1 cutsky simulations considered in section 5.3.

4.2 Marked power spectrum

The estimation of the marked power spectrum, including the effects of the survey geometry, closely follows the power spectrum calculation. Using the expression for the overdensity field $\delta_g = (n_g - \alpha n_r)/\alpha n_r$ one can describe the marked field as

$$\bar{m}\delta_M = m(\delta_g + 1) - \bar{m} = \frac{m n_g}{\alpha n_r} - \bar{m} = \frac{m n_g - \bar{m}\alpha n_r}{\alpha n_r}. \tag{4.10}$$

In the power spectrum calculation, the FKP field $\mathcal{F}$ is related to the overdensity field by a factor of the mean density

$$\mathcal{F}(\mathbf{x}) = n_g(\mathbf{x}) - \alpha n_r(\mathbf{x}) \quad \text{cf.} \quad \delta_g = \frac{n_g - \alpha n_r}{\alpha\, n_r} \tag{4.11}$$

In parallel to this, we define

$$\mathcal{F}_M(\mathbf{x}) = m(\mathbf{x}) n_g(\mathbf{x}) - \bar{m}\alpha\, n_r(\mathbf{x}) \tag{4.12}$$

which substitutes $\mathcal{F}$ in the power spectrum estimator, while holding the remaining components the same, including the normalization A. Since M is also a 2-point function and the geometrical effect is fully captured by the randoms designed for the power spectrum, the same window

JCAP08(2026)003

matrix ($\mathcal{W}$) established to model the power spectrum over a fraction of the sky can be reused, making the overall equation simply

$$(\widehat{M}_\ell^{\rm obs})_i = (\mathcal{W}_{\ell\ell'})_{ij}(M_{\ell'})_j \tag{4.13}$$

This is a major advantage, as there is no need to rebuild a new window formalism incorporating the same corrections and it bypasses the need for many validations.

The above derivation neglects the effect of the survey geometry on the mark, i.e. it assumes that $W(x)\delta_M(x) \approx m(x)W(x)\delta_g(x)$. As long as the survey volume is much larger than R^3 and not full of holes this approximation is well-justified. Phrased another way, the mark is a local function of the density field and the window effect is most significant at large scales. Furthermore, our use of this approximation is the same as that employed in the standard post-reconstruction BAO analysis (with smoothing $R = 15\,h^{-1}\,$Mpc), which itself applies a (likely) less local operation on δ_g and yet yields results consistent with the configuration-space analysis that fully incorporates the window in the 2PCF.

This approach to survey geometry differs from those in the past literature. Ref. [64] approximates the window convolution of the skew spectrum by decomposing the tree-level bispectrum into products of power spectra and convolving each leg with the power spectrum window matrix. This approximation is fairly successful with few-percent errors typically, but can reach up to 18% deviations for select shapes in the presence of primordial non-Gaussianity. Recently, ref. [65] has developed the exact window treatment for the Lyman-α forest skew spectrum $\langle\delta_F^2\,\delta_F\rangle$, where the absence of derivative operators in δ^2 allows a clean factorization of the window from the three-point correlator, yielding a window kernel involving $W^2(\mathbf{x}_1)W(\mathbf{x}_2)$ that can be computed via pair counting. For the MPS or galaxy skew spectra involving (transversely) smoothed fields, this factorization is complicated by the smoothing kernel, which entangles the window with the density field on scale R, although the MPS minimizes this entanglement by involving only the single smoothing kernel required to regulate UV divergences in the quadratic field. For galaxy surveys, the dense three-dimensional sampling furthermore permits a robust estimate of the smoothed density field in the survey interior using random catalogs, in the same manner as BAO reconstruction, such that the mark is approximately independent of the survey geometry. Under this approximation the window treatment of the MPS reduces to that of a standard two-point function. We validate this on cutsky mocks in section 5.3, finding sub-σ agreement across all modes. The situation is fundamentally different for the Lyman-α forest, where the sparse transverse sampling of quasar sightlines makes a stable local estimate of the smoothed selection function impractical, necessitating the explicit forward modeling of the window.

4.3 Implementation

We implement the algorithm above largely inheriting the code structure of `pypower`[11] [66], which includes both the methods necessary for two-point correlator measurements and window matrix calculations. As mentioned above, the latter does not require modifications, all of the changes will be in the former; in particular, we will aim to replace the FKP field $\mathcal{F}$ with $\mathcal{F}_M = mn_g - \bar{m}\alpha n_r$.

[11] https://github.com/cosmodesi/pypower.

JCAP08(2026)003

To this end we require a robust way to measure the mark m, which is dependent on the measurement of the smoothed field $\delta_{g,R}$ over an inhomogeneous footprint. This is a well-known technique to the field, as $\delta_{g,R}$ is a key component of BAO reconstruction. For our implementation we will follow that of `pyrecon`[12]

$$\delta_{g,R}(\mathbf{x}) = \frac{n_{g,R}(\mathbf{x}) - \alpha n_{r,R}(\mathbf{x})}{\alpha n_{r,R}(\mathbf{x})} \tag{4.14}$$

where $n_{g,R}(\mathbf{k}) = W_R(k) n_g(\mathbf{k})$ and $n_{r,R}(\mathbf{k}) = W_R(k) n_r(\mathbf{k})$. For stability we set $\delta = 0$ for mesh points with less than 1% of the average random weight. Note that the density field calculation increases the importance of the density of randoms and mesh grid size compared to a standard power spectrum calculation. This will be revisited at the end of this section. Given this stable density calculation, measuring $\mathcal{M}$ (eq. (2.8)) from data is straightforward.[13]

From a practical point of view of a $P + M$ fit, it is beneficial to fit the higher-point information M directly, instead of $\mathcal{M}$, which (doubly) includes two-point information from P, making the modeling of off-diagonal covariance more important. In order to measure M there is one additional procedure necessary to account for the smoothing of two-point information. Since the smoothing $W_R(k)$ is not commutative with the effect of the survey geometry,

$$M_{\rm obs} \neq \mathcal{M}_{\rm obs} - (C_0^a + C_1^a W_R(k))(C_0^b + C_1^b W_R(k)) P_{\rm obs} \tag{4.15}$$

Rather, we must consider the smoothed power spectra individually and subtract their contributions

$$M_{\rm obs} = \mathcal{M}_{\rm obs} - (C_0^a C_0^b P_{\rm obs} + (C_0^a C_1^b + C_0^b C_1^a) P_{R,\rm obs} + C_1^a C_1^b P_{RR,\rm obs}) \tag{4.16}$$

where $P_R = \langle \delta_{g,R} \delta_g \rangle$ and $P_{RR} = \left\langle \delta_{g,R}^2 \right\rangle$ can be computed analogously to P itself.

Although the necessity to measure the density field numerically increases the importance of the density of randoms and meshgrid size compared to a standard power spectrum measurement, we have verified that the algorithm performance converges within reasonable requirements. We validate the convergence of the code using both periodic box and 'cutsky' simulations, as introduced later in section 5. In this work we adopt algorithm settings of meshsize of $2\,h^{-1}\,$Mpc for the periodic box and $3.5\,h^{-1}\,$Mpc for cutsky, and a random density of $3.5 \times 10^{-2}\,h^3\,\mathrm{Mpc}^{-3}$ (all 18 random catalogs associated) with cutsky. With these settings, we achieve performance that converges to $\lesssim 1\%$ in MPS amplitude, which is sufficient for this work. We do not require randoms for the periodic box since we can assume uniform random density.

5 Validation on mocks

5.1 Periodic box mocks

Here we perform validations of the theory against mock catalogs generated on the periodic boxes from the AbacusSummit N-body simulation suite [67], produced with the Abacus

[12] https://github.com/cosmodesi/pyrecon.

[13] The software used for this study will be incorporated into the DESI pipeline in the future.

JCAP08(2026)003

N-body code [68, 69]. We adopt the 25 'base' ($2\,h^{-1}\,$Gpc) boxes at $z = 0.8$, with the DESI LRG2 tracer in mind ($0.6 < z < 0.8$), and populate the N-body simulation with halos generated using the standard, five-parameter Halo Occupation Distribution (HOD) model [70] implemented in AbacusUtils[14] software [71]. The standard HOD model defines probability distribution of central and satellite galaxies depending on host halo mass. The central and satellite galaxies are populated based on a binomial and Poisson-based distribution, respectively. This can be described by

$$\langle N_{\rm cen}(M_h)\rangle = \frac{1}{2}\mathrm{erfc}\left(\frac{\ln M_{\rm cut}/M_h}{\sqrt{2}\sigma}\right) \tag{5.1}$$

$$\langle N_{\rm sat}(M_h)\rangle = \langle N_{\rm cen}(M_h)\rangle \left(\frac{M_h - \kappa M_{\rm cut}}{M_1}\right)^{\alpha} \quad \text{for } M_h > \kappa M_{\rm cut} \tag{5.2}$$

where M_h is the halo mass and $\{M_{\rm cut},\, M_1,\, \sigma,\, \kappa,\, \alpha\}$ are model parameters.

Using the best-fit parameters for the LRG2 sample in the DESI one-percent survey $\{\log M_{\rm cut} = 12.78,\, \log M_1 = 13.94,\, \sigma = 0.17,\, \alpha = 1.07,\, \kappa = 0.55\}$ [72], we generate mock catalogs. We omit the incompleteness fraction, $f_{\rm ic}$, from the HOD fit and simply randomly downsample the mock galaxies to produce mocks at two number densities, $\bar{n} = 10^{-3}$ and $3\times10^{-4}\,h^{-3}\,\mathrm{Mpc}^3$, which roughly spans the range of DESI galaxy densities [5, 73]. The mock galaxies have linear bias $b_1 \approx 2.11$ [72] and power spectrum monopoles and quadrupoles as shown in the blue and orange lines in figure 4.

We will use the mean and standard deviation between the boxes to validate the MPS theory for near-future data. The volume of each box corresponds to the volume of the LRG2 z-bin ($0.6 < z < 0.8$) but with a sky area of $\approx 22000\,\mathrm{deg}^2$. This is $\approx 30\%$ larger than the footprint at the end of the DESI survey ($17000\,\mathrm{deg}^2$), providing a sufficient theory validation for the data in the near future. This is in similar spirit to the effort in ref. [32], but differs crucially in the (reduced) number of nuisance parameters, as discussed in section 2.3, and the question of stochasticity of the marked field as discussed below.

5.2 Stochasticity with low number density

It has recently been pointed out [74] that the MPS involves a new stochasticity due to estimating the marked field m (and hence density field $\delta_{g,R}$) from a finite number of objects. This problem can be restructured as an introduction of a new stochastic field ϵ' for the MPS that is distinct from the power spectrum. To the extent that this stochasticity is scale-independent, this would introduce a new nuisance parameter N' where the power spectrum shot noise enters the MPS. However, as evident from eq. (2.29), the contribution is largely degenerate with $B_{\rm shot}$ when considering only the monopole M_0. By fitting to mock catalogs with varying number densities with the same non-stochastic nuisance parameters, we demonstrate that this is indeed the case practically.

In figure 4 we show the fit against P_0, P_2, and M_0 of periodic boxes, with $k^P_{\rm max} = 0.2\,h\,\mathrm{Mpc}^{-1}$ and $k^M_{\rm max} = 0.12\,h\,\mathrm{Mpc}^{-1}$. Through random downsampling, we fit the catalog at two different number densities $\bar{n} = 10^{-3}$ and $3\times10^{-4}\,h^{-3}\,\mathrm{Mpc}^3$, with three different smoothing radii $R = 10,\, 15,\, 20\ h^{-1}\,\mathrm{Mpc}$. We will employ the `scipy` minimizer module in

[14] https://abacusutils.readthedocs.io/en/latest/.

JCAP08(2026)003

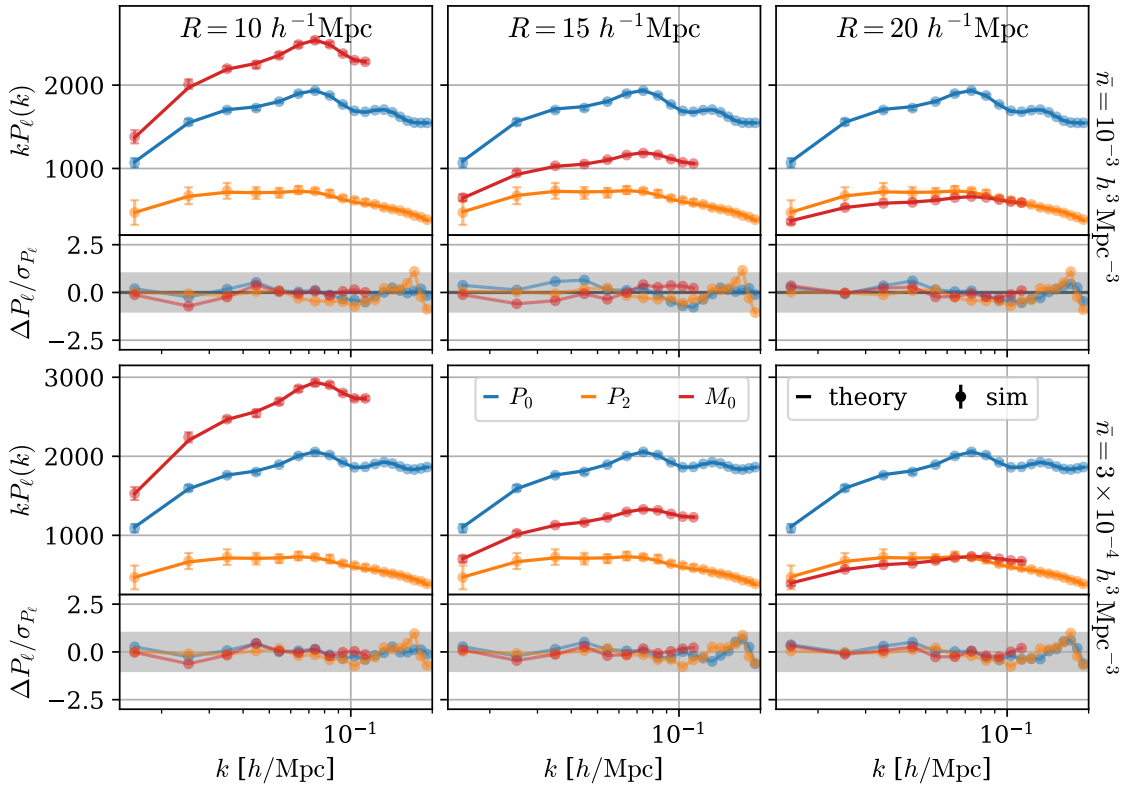


Figure 4. P_ℓ and M_0 joint-fits to the average of 25 simulations in $8\,h^{-3}\,\mathrm{Gpc}^3$ periodic boxes, with each column corresponding to smoothing scales $R = 10,\ 15,\ 20\,h^{-1}\,\mathrm{Mpc}$, and each row corresponding to number densities $\bar{n} = 10^{-3}$ and $3\times10^{-4}\,h^3\,\mathrm{Mpc}^{-3}$. In the top panels, the theory fits are in solid lines and simulation data points in circular markers. The bottom panels show the residuals of the fits using the standard deviation of simulations as errors and we add gray bands to demonstrate the 1σ range. For all panels, blue, orange, and red colors represent P_0, P_2, and M_0, respectively, and the fits adopt $k^P_{\mathrm{max}} = 0.2\,h\,\mathrm{Mpc}^{-1}$ and $k^M_{\mathrm{max}} = 0.12\,h\,\mathrm{Mpc}^{-1}$. The successful $< 1\sigma$ fit for all cases demonstrate that the theoretical precision is enough to be adopted for near-term datasets. Furthermore the fits at different number densities indicate that potential stochastic effects are degenerate with other parameters and can be neglected (section 5.2).

`Cobaya` [75–77] for the fit. Based on the errors calculated from the 25 boxes, we are able to fit all boxes to $\sim 1\sigma$ with the same non-stochastic nuisance parameters (biases b_i and counterterms α_{2n}). The volume of each box ($8\,h^{-3}\,\mathrm{Gpc}^3$) is larger than the effective volume of all DESI DR2 redshift bins [5], indicating that the model accuracy is sufficient for analysis on upcoming data releases, although systematics modeling, such as fiber assignment effects, still must be addressed before data application.

5.3 Cutsky mocks

We use the public, DESI DR1 cutsky mocks [78] for validation of the MPS with survey geometry, focusing on the LRG2 redshift bin, spanning $0.6 < z < 0.8$. The LRG2 mocks are generated based on 25 periodic AbacusSummit simulations [67–69] in cubic boxes of $2\,h^{-1}\,\mathrm{Gpc}$. The mocks are based on the $z = 0.8$ output with galaxies included using the

JCAP08(2026)003

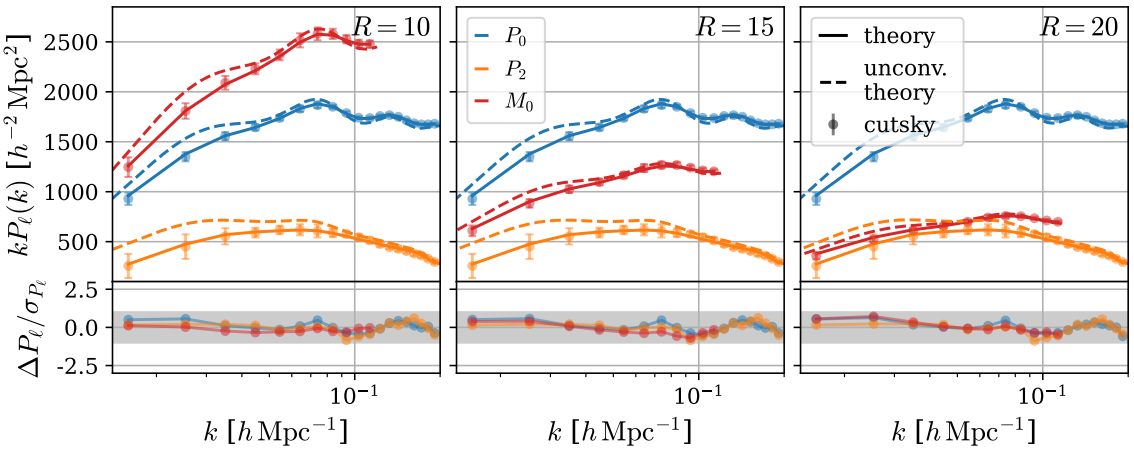


Figure 5. The P-M joint fits to DESI DR1 cutsky mocks in the LRG2 ($0.6 < z < 0.8$) bin with varying smoothing scales $R = 10, 15, 20\,h^{-1}\,\mathrm{Mpc}$. The top panels show the fits directly, with the window-convolved and unconvolved theory shown in solid and dashed lines against the mock measurements in circular markers. The bottom panels show the residual of the fits using simulation errors, with the 1σ range highlighted by the gray band. The fits are $< 1\sigma$ for all cases, demonstrating that the survey geometry is propagated to the spectrum at precision satisfactory for any near-future data.

extended HOD model that incorporates velocity bias through two parameters: α_c and α_s [79]. The first adds a velocity dispersion for central galaxies of α_c times the halo velocity dispersion. The second scales the satellite-halo relative velocity by α_s. The standard five-parameter HOD model (eq. (5.1) and (5.2)) [70] can be recovered in the limit $\alpha_c = 0$ and $\alpha_s = 1$. The HOD parameters used are $\{\log M_{\rm cut} = 12.64, \log M_1 = 13.71, \sigma = 0.09, \alpha = 1.18, \kappa = 0.6, \alpha_c = 0.19, \alpha_s = 0.95\}$ which are best-fits to the 3D correlation function and number density in the DESI One-Percent Survey [72], modulo the incompleteness fraction. The periodic box results are then matched to the DESI NGC and SGC footprints of DR1 by applying coordinate transforms and including the line-of-sight velocities for RSD. This transforms the mocks into a total footprint of $\sim 5800\,\mathrm{deg}^2$ with the window matrix as shown in figure 3. For this work, we will focus on the mocks without fiber assignment ("complete" mocks) to capture the survey geometry effects independent of other observational artifacts. The effects of fiber assignment will be investigated in the future when applying the methodology to observational data. We refer the reader to ref. [78] for further details about these simulations.

We measure both the power spectrum multipoles and MPS monopole from each of the 25 cutsky mocks with smoothing radii of $R = 10$, 15, and $20\,h^{-1}$Mpc, and measure their standard deviation. Using these errors we once again use the `scipy` minimizer module in `Cobaya` [75–77] to inform our joint fits of P and M. The fit results are shown in figure 5, showing that one can model the MPS with a realistic survey window jointly with the power spectrum to within $1\,\sigma$ of observational error bars. Note that since these are in DR1 mocks with a smaller and less homogeneous footprint than future data releases, this agreement indicates that we can model such effects for future DESI results.

JCAP08(2026)003

6 Alcock-Paczynski effect

In practice, galaxy surveys detect the 3D position of LSS using angles and redshift, which require translation to positions using a fiducial cosmology. As it is unlikely that the fiducial cosmology is identical to the "true" cosmology, the choice of cosmology results in a coordinate distortion and one must account for the cosmology mis-specification. This is referred to as the Alcock-Paczynski (A-P) effect [80].

Using the angle on the sky $\Delta\theta$ and redshift difference Δz, one can compute their perpendicular and los separations using

$$r_\perp \approx D_A(z)\Delta\theta \ , \qquad r_\parallel = \Delta\chi \approx \frac{c\Delta z}{H(z)} \tag{6.1}$$

where the comoving distance χ and angular diameter distance D_A are

$$\chi = \int \frac{c\,dz'}{H(z')} \ , \qquad D_A(z) = \frac{\chi(z)}{(1+z)} \tag{6.2}$$

This implies that the fiducial and true wavevectors are related by

$$\mathbf{k}_{\rm fid} = \mathbf{k}_\parallel \alpha_\parallel(z) + \mathbf{k}_\perp \alpha_\perp(z) \tag{6.3}$$

for $\alpha_\parallel = H(z)_{\rm fid}/H(z)_{\rm true}$ and $\alpha_\perp = D_A(z)_{\rm true}/D_A(z)_{\rm fid}$. Accounting for the units of volume in the power spectrum ($P_{\rm obs}(\mathbf{k}_{\rm fid}, z)d^3k_{\rm fid} = P_{\rm true}(\mathbf{k}_{\rm true}, z)d^3k_{\rm true}$) we get [80, 81]

$$P_{\rm obs}(\mathbf{k}_{\rm fid}, z) = \alpha_\parallel^{-1}(z)\alpha_\perp^{-2}(z) P_{\rm true}(\mathbf{k}_{\rm true}, z) \tag{6.4}$$

The inclusion of A-P effects in the power spectrum has been standardized, such that the effect can be calculated at every cosmology analytically (to lowest order in $\delta\alpha$) without significant loss in the MCMC inference pipeline.

For the MPS the fiducial cosmology dependence is more complex than for the power spectrum, due to both the involvement of an additional overdensity field and the smoothing $W_R(k)$. An analytical solution can be best captured by recalling the expression of M as an integral over B (eq. (2.11)). The cosmology dependence of the (tree-level) bispectrum is routinely included and can be calculated extremely rapidly; for example with software such as `FOLPS-D`[15] [82] the evaluation of B takes less than a millisecond. In practice, however, incorporating this effect without loss in computational time is non-trivial, as the integration over B is the computationally expensive procedure in evaluating M. While some of this cost can be handled through the use of emulators, we also wanted to explore an alternative route for incorporating the A-P effect. Below we shall demonstrate that the change of MPS over the fiducial cosmology choice is smooth enough that we can interpolate between cosmologies to match the observed cosmology to the evaluated cosmology at every step of the MCMC chain [83]. This way there will be no position distortion, as the cosmologies used for MPS measurement and model evaluation are identical (alternatively we can use the inverse scaling to introduce the distortion into the theory). Parameter estimation runtime is not compromised significantly, as the measurement pipeline is run once over a sufficiently large cosmology parameter space and then the results interpolated during the MCMC steps.

[15] https://github.com/alejandroaviles/folpsD.

JCAP08(2026)003

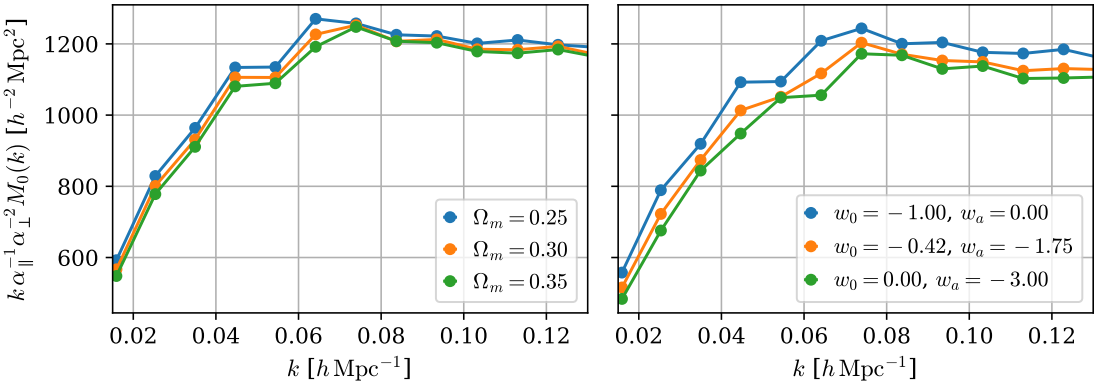


Figure 6. The residual A-P effect of M in ΛCDM (left) and w_0w_aCDM (right), using a smoothing radius $R = 15\,h^{-1}$ Mpc and one instance of DESI DR1 cutsky mocks. For ΛCDM Ω_m is varied between 0.25 and 0.35. The range for w_0w_aCDM correspond to the -2σ to $+2\sigma$ extent in the recent DESI+CMB constraints [5]. Over these cosmologies the change in $\alpha_\parallel^{-1}\alpha_\perp^{-2}\,M_0$ is small and smooth.

The MPS shares the same units of volume as the power spectrum, so the volume factor $\alpha_\parallel^{-1}\alpha_\perp^{-2}$ can be factored out in an identical manner. We therefore focus exclusively on the residual transformation arising from the cosmological dependence of $\mathbf{k}$. In a ΛCDM cosmology, the only cosmological parameter that A-P depends on is Ω_m. For extended cosmology models, there can be more parameter dependencies. For instance, in a w_0w_aCDM, which has been found to be preferred by DESI+CMB constraints [5], there are additional dependencies on the dark energy equation of state w_0, w_a. These also induce very smooth distortions.

Figure 6 shows the residual A-P variation in both ΛCDM and w_0w_aCDM, varying Ω_m between 0.25 and 0.35 and w_0w_a within recent -2σ to $+2\sigma$ constraints with DESI+CMB [5]. The variation of $\alpha_\parallel^{-1}\alpha_\perp^{-2}\,M_0$ over these wide range of cosmologies is small and smooth, implying that it can be interpolated easily.

7 Conclusions

Higher-order statistics of large-scale structure offer insight into non-Gaussian information that is not captured in traditional two-point correlators. As cosmological surveys make advance in precision, accessing these non-Gaussian information becomes important to break parameter degeneracies and improve cosmological constraints [50, 84]. The marked power spectrum (MPS) [31, 32, 37] probes particular compressions of higher-order correlations while retaining the structure of two-point correlators, allowing much of the existing two-point analysis infrastructure to be reused and enabling perturbative modeling without introducing new theoretical uncertainties. In this work we address both the perturbative and survey modeling of the MPS, in order to prepare for application in near-future datasets.

Based on the analytical model, we redefine the MPS in order to isolate the higher-order information and decrease overlap with the power spectrum. Through this redefinition we find strong potential for degeneracy breaking, with reasonably expected changes in the secondary bias b_2 yielding MPS that are different in amplitude by a factor of 2, which is in agreement with, but strengthens, the results of ref. [32]. This signal exceeds observational errors and higher-order corrections by more than an order of magnitude.

JCAP08(2026)003

Recent work raised the possibility that the MPS may have additional stochasticity from constructing the density field from a finite number of objects [74]. We show this source of noise is already encompassed in the model's free parameters and use 25 mock catalogs in $2\,h^{-1}$Gpc periodic boxes to show that this is not a concern. We also demonstrate that small higher-loop corrections to the MPS are not needed for the fits in this work, though they may be required in future analyses [32, 37].

The effect of survey geometry on the MPS can be modeled similarly to the power spectrum, since the mark is a local, smooth function. This is a major advantage in terms of data application, as this simplifies the modeling and infrastructure to implement these effects already exists and is well-known in two-point analyses [5, 6]. We show how to modify standard 2-point function codes to compute the MPS and its window matrix, demonstrating that converged results can be obtained with modest computational requirements. We test the modeling of the survey geometry by jointly fitting the power spectrum and MPS against DESI DR1 cutsky catalogs. The results show consistency between model and simulation within 1σ in all cases, up to $k^P_{\rm max} = 0.2\,h\,{\rm Mpc}^{-1}$ for the power spectrum multipoles and $k^M_{\rm max} = 0.12\,h\,{\rm Mpc}^{-1}$ for the MPS monopole. As the DESI DR1 data are more non-uniform than future DESI data releases, we expect that this agreement will hold for all near-future datasets.

Another major advantage of the MPS is the (relatively) small number of data points, as it has one free momentum vector constrained to low k. This will especially help in the construction of covariance matrices, as the number of simulations necessary to construct the covariance is likely similar to that of the power spectrum. While a full investigation requires a large simulation suite and is left for future work, we develop insight through several avenues. Our redefinition of the MPS decreases the covariance between P and M, lowering the required precision of their cross-covariance. Direct calculation of the covariance reveals that the Gaussian approximation captures all diagonal contributions, in agreement with ref. [21], but that the cross-covariance is not strongly suppressed, unlike the bispectrum case [53, 55] due to difference in mode-counting factors. We qualitatively confirm these findings with our 25 periodic box simulations.

We also address fiducial cosmology dependence. While evaluating the Alcock-Paczynski effect [80] for the MPS is technically possible, implementing it without significant computational cost is non-trivial due to the additional density field and smoothing kernel $W_R(k)$. This can be circumvented by re-evaluating the redshift-distance relation at every likelihood stage. We demonstrate that the cosmology dependence of the MPS is marginal and smooth, allowing simple interpolation over cosmologies to perform this at high precision.

Further work is still required before application to data. The effects of fiber assignment must be modeled, as redshift surveys like DESI do not obtain redshifts for all galaxies and fiber completeness varies with environmental variables such as local density. We must also validate that the MPS returns unbiased cosmological constraints from mock catalogs, which requires faster analytical techniques such as FFTlog [85] or cosmology-based emulators [6], or both. The covariance question must likewise be addressed with a large simulation suite providing precision necessary for modern datasets.

Data availability. The software used for the analytical calculations in this work are publicly available at https://github.com/HarukiEbina/markedPS.

JCAP08(2026)003

Acknowledgments

HE and MW were supported by the DOE. This research was supported in part by grant NSF PHY-2309135 to the Kavli Institute for Theoretical Physics (KITP). This work made use of the Cobaya analysis code [75, 76]. This research used resources of the National Energy Research Scientific Computing Center (NERSC), a Department of Energy User Facility.

A The beyond-2pt information in $\mathcal{M}$

As described in eq. (2.8), the original marked power spectrum $\mathcal{M}$ is a combination of two-point and beyond two-point information. For brevity, we will summarize the two-point information as $\mathcal{M}_{\rm 2pt}$

$$\mathcal{M} = \mathcal{M}_{\rm 2pt} + M \tag{A.1}$$

where M is the marked power spectrum in the main text, extracting the beyond two-point information explicitly. Here, we will explore the respective fraction of these two components in order to gain better perspective of statements made in past work referencing $\mathcal{M}$ [31, 32, 37]. To do this, we measure $\mathcal{M}_{\rm 2pt}$ and M for a high-density ($\bar{n} = 10^{-3}\,h^3\,{\rm Mpc}^{-3}$) DESI LRG-like mock catalogs over 25 periodic simulation boxes of volume $8\,h^{-3}\,{\rm Gpc}^3$ at $z = 0.8$ (see section 5.1).

Of course, as the MPS does not break the perturbative scaling, for an unspecified mark the linear term in $\mathcal{M}_{\rm 2pt}$ will dominate. However, taking advantage of the overall scaling of $\mathcal{M}_{\rm 2pt}$ [32, 37]

$$\mathcal{M}_{{\rm 2pt},\ell}(k) = C^a_{\delta_M}(k) C^b_{\delta_M}(k) P^{\rm one-loop}_\ell(k) \tag{A.2}$$

$$\text{for} \quad C^a_{\delta_M}(k) = C^a_0 + C^a_1 W_R(k) \tag{A.3}$$

one can set one or both of the marks to $m = 1 - \delta_{g,R}$ to suppress two-point information at theoretically well-described large-scales (low-k). This was the approach adopted in ref. [32]. Using the cross-correlation of this mark with the unmarked density field for $R = 10$, 15, and 20 $h^{-1}\,{\rm Mpc}$,[16] we find the results in figure 7. As intended, the mark choice suppresses the $\mathcal{M}_{\rm 2pt}$ sufficiently such that M dominates at low-k. This decomposition, however, also shows that the two-point information provides a non-negligible contribution to even linear scales, which can introduce significant covariance between P and $\mathcal{M}$ and in turn make the requirements on the covariance more stringent. Thus, it is better motivated to directly access the beyond two-point term M by subtracting $\mathcal{M}_{\rm 2pt}$ at the level of the estimator.

B Choice of marks

The calculation in eq. (2.9) indicates that the cross-spectrum of the unmarked field with either $m = 1 \pm \delta_{g,R}$ or $m = \delta_{g,R}$ agree, up to an overall sign. Ref. [32] found that there are practical differences, with the spectra of $m = 1 + \delta_{g,R}$ more stable than that of $m = 1 - \delta_{g,R}$, especially for small smoothing scales (e.g. $R = 10h^{-1}{\rm Mpc}$) where the long tail of $\delta_{g,R}$ can lead to zero-crossings of the mark. By contrast, we find that the cross-spectrum of $m = 1 \pm \delta_{g,R}$ and

[16] As mentioned in section 2, this yields the same M as using $m = 1 + \delta_{g,R}$, as done in the main text.

JCAP08(2026)003

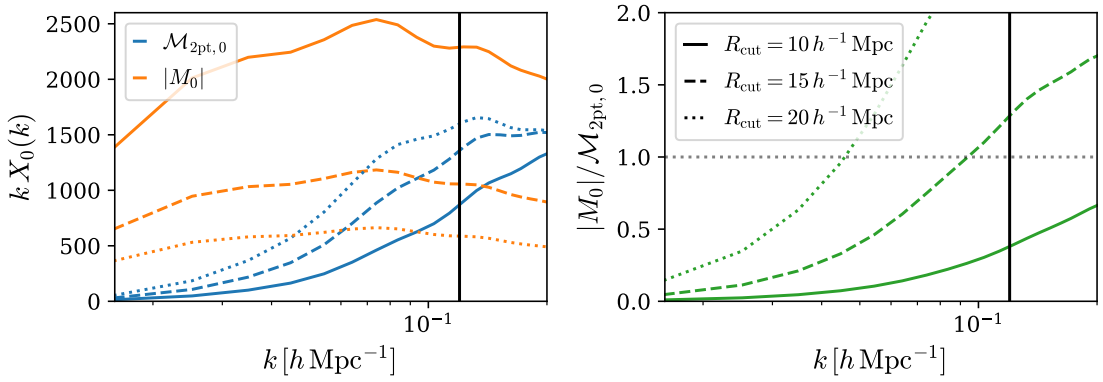


Figure 7. The two-point ($\mathcal{M}_{2\mathrm{pt}}$) and beyond-two-point ($M$) power in $\mathcal{M}_0$. The left panel directly shows the $\mathcal{M}_{2\mathrm{pt},0}$ (blue) and $|M_0|$ (orange) as measured from mock catalogs in periodic boxes, with the smoothing scales $R = 10$ (solid), 15 (dashed) and $20h^{-1}$ Mpc (dotted). The right panel shows the ratio of the two components $|M_0|/\mathcal{M}_{2\mathrm{pt},0}$ (green), with the same line styles corresponding to smoothing radii as the left panel. As anticipated the choice of mark suppresses the two-point component at low-k, with the rate of increase in two-point power strongly dependent on the smoothing radius.

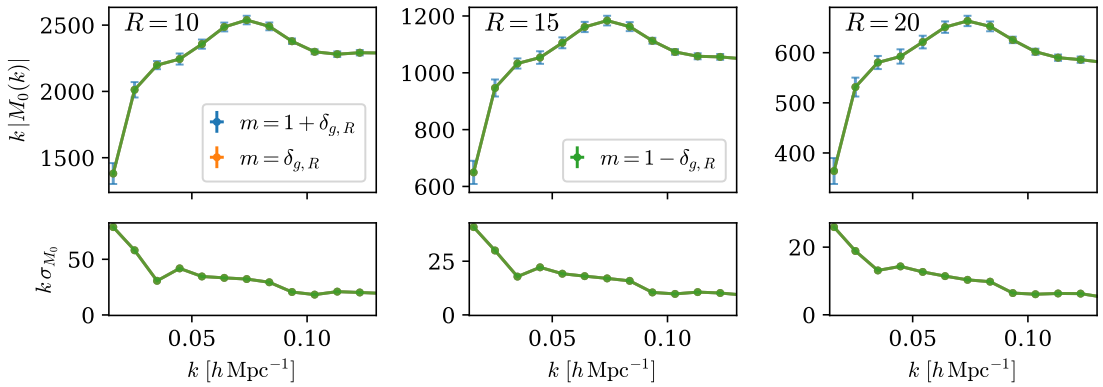


Figure 8. $|M_0|$ and its standard deviation σ_{M_0} for the cross-spectrum of $m = 1 \pm \delta_{g,R}$ and $m = \delta_{g,R}$ for $R = 10$, 15, and $20h^{-1}$Mpc, measured over 25 $8\,h^{-3}\,\mathrm{Gpc}^3$ boxes. The three marks give identical means (up to a sign) and dispersion as expected from eq. (2.9). In the figure the blue, orange and green points and lines lie on top of each other.

$m = \delta_{g,R}$ are in agreement with each other, as shown in figure 8. This may be a result of a more stable observation code, with the density estimation following that of the reconstruction code `pyrecon`. The situation may be different for scenarios involving systematics and systematics correction weights, such as fiber assignment effects. The correction schemes have been tested for subsets of the data, but under significant reweighting it is unclear whether the corrections continue to work as anticipated. In these scenarios, it is conceivable that additional constants in the mark can help counter systematics due to less relative reweighting. Furthermore, one can optimize the mark in a higher polynomial order to tighten specific constraints [38]. We will leave such investigation to future work.

JCAP08(2026)003

C Scale-dependence of M_{13}

Here we discuss the similarity between the scale-dependence of M_{13} and P_L, as seen in figure 2. We have discussed in section 2 that the M_{13} contribution consists of $\left\langle \delta^{(1)}_{g,R,1}\delta^{(2)}_{g,1}\delta^{(1)}_{g,2}\right\rangle$ and $\left\langle \delta^{(2)}_{g,R,1}\delta^{(1)}_{g,1}\delta^{(1)}_{g,2}\right\rangle$ that merely differ by the argument of the smoothing kernel. Focusing on the first term, we have

$$\left\langle \delta^{(1)}_{g,R,1}\delta^{(2)}_{g,1}\delta^{(1)}_{g,2}\right\rangle = \int_{\mathbf{p}} Z_2(\mathbf{k},\mathbf{p})Z_1(\mathbf{k})Z_1(\mathbf{p})W_R(p)P_L(k)P_L(p) \tag{C.1}$$

in the absence of stochastic terms. In section 2 and figure 2 we have already demonstrated that the b_2 contributions are approximately $\propto P_L(k)$, with little additional shape (k) dependence, and the b_s contribution is subdominant to that of b_2 by a factor of 5 or more. Now we inspect the other terms. As shown in eqs. (2.17) and (2.18), the first two terms of the kernels F_2 and G_2 merely contribute angular dependencies that will only matter when considering the marked spectrum quadrupole. The final terms are the only contributors to spectrum shape for both kernels. For the F_2 contribution, we find, however, that the final term vanishes due to symmetry, i.e.

$$\left\langle \delta^{(1)}_{g,R,1}\delta^{(2)}_{g,1}\delta^{(1)}_{g,2}\right\rangle \supset \int_{\mathbf{p}} b_1 F_2(\mathbf{k},\mathbf{p})Z_1(\mathbf{k})Z_1(\mathbf{p})W_R(p)P_L(k)P_L(p) \tag{C.2}$$

$$= b_1 Z_1(\mathbf{k})P_L(k)\int W_R(p)P_L(p)p^2dp \tag{C.3}$$

$$\times\left[\frac{5}{7}\left(\frac{4\pi}{3}(3b_1+f)\right)+\frac{2}{7}\left(\frac{4\pi}{15}(5b_1+f+2f\mu^2)\right)+\frac{1}{2}\times 0\right] \tag{C.4}$$

adding no additional scale-dependence. The G_2 contribution offers a small correction to the shape, with the monopole contribution being

$$\left\langle \delta^{(1)}_{g,R,1}\delta^{(2)}_{g,1}\delta^{(1)}_{g,2}\right\rangle_0 \supset \int_{\mathbf{p}} f\mu^2_{\mathbf{k}+\mathbf{p}}G_2(\mathbf{k},\mathbf{p})Z_1(\mathbf{k})Z_1(\mathbf{p})W_R(p)P_L(k)P_L(p) \tag{C.5}$$

$$= P_L(k)\int\frac{p^2\,dp}{(2\pi)^2}\left(c_0+c_2\frac{k^2}{p^2}+c_4\frac{k^4}{p^4}+\mathcal{O}(k^6/p^6)\right)W_R(p)P_L(p) \tag{C.6}$$

with

$$c_0 = \frac{2f}{11025}\left(2275b_1^2+2394b_1f+603f^2\right)\approx 2.269 \tag{C.7}$$

$$c_2 = \frac{-128f^2}{25725}(7b_1+3f)\approx -0.067 \tag{C.8}$$

$$c_4 = \frac{256f^2}{77125}(7b_1+3f)\approx 0.045 \tag{C.9}$$

where the numerical approximations are for $b_1 = 2$ and $f = 0.9$. Finally the last contribution to Z_2 is

$$\left\langle \delta^{(1)}_{g,R,1}\delta^{(2)}_{g,1}\delta^{(1)}_{g,2}\right\rangle \supset \int_{\mathbf{p}}\frac{fk\mu_{\mathbf{k}+\mathbf{p}}}{2}\left[\frac{\mu}{k}(b_1+f\mu_p^2)+\frac{\mu_p}{k_p}(b_1+f\mu^2)\right]Z_1(\mathbf{k})Z_1(\mathbf{p})W_R(p)P_L(k)P_L(p) \tag{C.10}$$

$$= Z_1(\mathbf{k})P_L(k)\left(\frac{b_1f}{15}(5b_1+3f)+\left(b_1^2f+b_1f^2+\frac{2f^3}{5}\right)\mu^2\right)\int\frac{p^2\,dp}{(2\pi)^2}W_R(p)P_L(p) \tag{C.11}$$

again indicating that the scale-dependence of M resembles that of P_L.

JCAP08(2026)003

Although the triangle configurations are different, since $k \lesssim 0.08\,h\,\mathrm{Mpc}^{-1}$ and $p \lesssim 1/R < 0.1\,h\,\mathrm{Mpc}^{-1}$, the situation here is mathematically similar to the squeezed bispectrum, which also display similar scale-dependence to the power spectrum. For the squeezed bispectrum, one operates with one of the triangle sides q 'squeezed' to zero in momentum space

$$\lim_{q\to 0} B(\mathbf{k},-\mathbf{k},\mathbf{q}) = Z_1^2(\mathbf{k})Z_2(\mathbf{k},-\mathbf{k})P_L^2(k) + Z_1(\mathbf{k})P_L(k)Z_1(\mathbf{q})P_L(q)\left[Z_2(\mathbf{k},\mathbf{q}) + Z_2(-\mathbf{k},\mathbf{q})\right] \tag{C.12}$$

where we use $Z_1(\mathbf{k}) = Z_1(-\mathbf{k})$ to simplify the expression. Since $Z_2(\mathbf{k},-\mathbf{k}) = b_2/2$, the first term is simply $(b_2/2)\,Z_1^2(\mathbf{k})P_L^2(k)$. The sum of Z_2's in the second term cancels any term that is odd[17] in $\hat{k}$, i.e.

$$Z_2(\mathbf{k},\mathbf{q})+Z_2(-\mathbf{k},\mathbf{q}) = 2\left[\frac{5}{7} + \frac{2}{7}\left(\hat{k}\cdot\hat{q}\right)^2 + \mu^2\left(\frac{3}{7} + \frac{4}{7}\left(\hat{k}\cdot\hat{q}\right)^2\right) + \frac{b_2}{2} + b_s\left(\left(\hat{k}\cdot\hat{q}\right)^2 - \frac{1}{3}\right)\right] \tag{C.13}$$

This leaves no k dependence to alter the spectrum shape from $Z_1(\mathbf{k})P_L(k)Z_1(\mathbf{q})P_L(q)$. This result is often quoted as the squeezed matter bispectrum $\lim_{q\to 0} B_m(k,k,q) \propto P_L(k)P_L(q)$, as the first term vanishes in the absence of b_2. Note that this computation does not consider the stochastic and counterterm contributions to the spectra, which will de-correlate the shapes of MPS and squeezed bispectrum from P_L.

D Covariance matrix

Here, we describe the calculation for the variance of M and covariance between P and M, supplementing the material in section 3.

D.1 Covariance of M

As introduced in the main text, the $k-k'$ covariance of a marked spectrum can be calculated as

$$\mathrm{Cov}\left(\widehat{M}_k, \widehat{M}_{k'}\right) \equiv \left\langle \widehat{M}_k\widehat{M}_{k'}\right\rangle - \left\langle \widehat{M}_k\right\rangle\left\langle \widehat{M}_{k'}\right\rangle \tag{D.1}$$

Expanding this expression we obtain

$$\mathrm{Cov}\left(\widehat{M}_k, \widehat{M}_{k'}\right) = V^{-2}\int_{V_k}\frac{d^3q}{V_k}\int_{V_{k'}}\frac{d^3q'}{V_{k'}}\int_{\mathbf{p}}\int_{\mathbf{p}'} W_R(p)W_R(p') \tag{D.2}$$

$$\left(\left\langle\delta_{-\mathbf{q}}\delta_{\mathbf{p}}\delta_{\mathbf{q}-\mathbf{p}}\delta_{-\mathbf{q}'}\delta_{\mathbf{p}'}\delta_{\mathbf{q}'-\mathbf{p}'}\right\rangle - \left\langle\delta_{-\mathbf{q}}\delta_{\mathbf{p}}\delta_{\mathbf{q}-\mathbf{p}}\right\rangle\left\langle\delta_{-\mathbf{q}'}\delta_{\mathbf{p}'}\delta_{\mathbf{q}'-\mathbf{p}'}\right\rangle\right) \tag{D.3}$$

$$= \mathcal{I}[W_R(p)W_R(p')\left\langle\delta_{-\mathbf{q}}\delta_{\mathbf{p}}\delta_{\mathbf{q}-\mathbf{p}}\delta_{-\mathbf{q}'}\delta_{\mathbf{p}'}\delta_{\mathbf{q}'-\mathbf{p}'}\right\rangle] - M_kM_{k'} \tag{D.4}$$

where we define[18]

$$\mathcal{I}[f] = V^{-2}\int_{V_k}\frac{d^3q}{V_k}\int_{V_{k'}}\frac{d^3q'}{V_{k'}}\int_{\mathbf{p}}\int_{\mathbf{p}'} f(\mathbf{q},\mathbf{q}',\mathbf{p},\mathbf{p}') \tag{D.5}$$

As discussed in the main text, there are three possible categories of contractions: $(2,2,2)$ (e.g. $\langle ab\rangle\langle cd\rangle\langle ef\rangle$), $(3,3)$ (e.g. $\langle abc\rangle\langle def\rangle$), $(4,2)$ (e.g. $\langle abcd\rangle\langle ef\rangle$), and (6) (e.g. $\langle abcdef\rangle$),

[17] Even if $q > 0$, all terms odd in $\hat{k}$ will vanish under angular integration due to symmetry.

[18] This differs from that in ref. [21] by factors of W_R, due to the aforementioned loss of symmetry between fields in $\tilde{\delta}_M$.

JCAP08(2026)003

for fields $a = \delta_{-\mathbf{q}}$, $b = \delta_{\mathbf{p}}$, $c = \delta_{\mathbf{q}-\mathbf{p}}$, $d = \delta_{-\mathbf{q}'}$, $e = \delta_{\mathbf{p}'}$, $f = \delta_{\mathbf{q}'-\mathbf{p}'}$. We will enumerate all possible contractions for each category.

Before we start, it will be useful to point out that $b = \delta_{\mathbf{p}}$ and $c = \delta_{\mathbf{k}-\mathbf{p}}$ are equivalent by the transformation $\mathbf{p} \leftrightarrow \mathbf{q} - \mathbf{p}$, which will help simplify our calculations. Note that the full expression does not adhere to this symmetry, as $W(p)$ will transform to $W(|\mathbf{q} - \mathbf{p}|)$. A similar symmetry is present for $e = \delta_{\mathbf{p}'}$ and $f = \delta_{\mathbf{q}'-\mathbf{p}'}$, with the transformation $\mathbf{p}' \leftrightarrow \mathbf{q}' - \mathbf{p}'$.

For the $(2, 2, 2)$ contractions there are 15 terms

$$\begin{aligned}\langle abcdef\rangle_{222} = &\langle ab\rangle\langle cd\rangle\langle ef\rangle + \langle ab\rangle\langle ce\rangle\langle df\rangle + \langle ab\rangle\langle cf\rangle\langle de\rangle + \\ &\langle ac\rangle\langle bd\rangle\langle ef\rangle + \langle ac\rangle\langle be\rangle\langle df\rangle + \langle ac\rangle\langle bf\rangle\langle de\rangle + \\ &\langle ad\rangle\langle bc\rangle\langle ef\rangle + \langle ad\rangle\langle be\rangle\langle cf\rangle + \langle ad\rangle\langle bf\rangle\langle ce\rangle + \\ &\langle ae\rangle\langle cd\rangle\langle bf\rangle + \langle ae\rangle\langle bc\rangle\langle df\rangle + \langle ae\rangle\langle cf\rangle\langle bd\rangle + \\ &\langle af\rangle\langle cd\rangle\langle eb\rangle + \langle af\rangle\langle ce\rangle\langle db\rangle + \langle af\rangle\langle bc\rangle\langle de\rangle \end{aligned} \tag{D.6}$$

Notice that the three fields involved in constructing each M (abc and def) necessarily have zero net momentum. As correlators enforce that the sum of field momenta are zero, if two of these fields are contracted (e.g. $\langle ab\rangle$) the momentum of the third field (e.g. c) must be zero, and thus contributes no power. This reduces the number of terms to 6

$$\begin{aligned}\langle abcdef\rangle_{222} = &\langle ad\rangle\langle be\rangle\langle cf\rangle + \langle ad\rangle\langle bf\rangle\langle ce\rangle + \langle ae\rangle\langle cd\rangle\langle bf\rangle + \\ &\langle ae\rangle\langle cf\rangle\langle bd\rangle + \langle af\rangle\langle cd\rangle\langle eb\rangle + \langle af\rangle\langle ce\rangle\langle db\rangle \end{aligned} \tag{D.7}$$

The first two terms that contract the two external momenta ($\langle ab\rangle = \langle \delta_{-\mathbf{q}}\delta_{-\mathbf{q}'}\rangle$) contribute diagonal terms of the covariance. Let us start with the first term

$$\langle ad\rangle\langle be\rangle\langle cf\rangle = \mathcal{I}\left[W_R(p)W_R(p')\left\langle\delta_{-\mathbf{k}}\delta_{-\mathbf{k}'}\right\rangle\left\langle\delta_{\mathbf{p}}\delta_{\mathbf{p}'}\right\rangle\left\langle\delta_{\mathbf{k}-\mathbf{p}}\delta_{\mathbf{k}'-\mathbf{p}'}\right\rangle\right] \tag{D.8}$$

$$\begin{aligned}&= (2\pi)^9 V^{-2}\int_{V_k}\frac{d^3q}{V_k}\int_{V_{k'}}\frac{d^3q'}{V_{k'}}\int_{\mathbf{p}}\int_{\mathbf{p}'}W_R(p)W_R(p') \\ &\quad \delta_D(\mathbf{q}+\mathbf{q}')\delta_D(\mathbf{p}+\mathbf{p}')\delta_D(\mathbf{q}-\mathbf{p}+\mathbf{q}'-\mathbf{p}')P(-\mathbf{q})P(\mathbf{p})P(\mathbf{q}-\mathbf{p})\end{aligned} \tag{D.9}$$

One of the Dirac delta functions yield a $\delta_D(0) = V/(2\pi)^3$, which cancels a volume factor. The other yields $\int_{V_k} d^3k'\delta_D(\mathbf{k}-\mathbf{k}') = \delta^K_{k,k'}$. Using the third delta function to eliminate one integral over an internal momentum and approximating $P(\mathbf{q}) \approx P(\mathbf{k})$ we obtain

$$\langle ad\rangle\langle be\rangle\langle cf\rangle = (2\pi)^3 V^{-1}V_k^{-1}\delta^K_{k,k'}\int_{\mathbf{p}}W_R^2(p)P(\mathbf{k})P(\mathbf{p})P(\mathbf{k}-\mathbf{p}) \tag{D.10}$$

The second term has a similar calculation. By taking advantage of the transformation $\mathbf{p}' \leftrightarrow \mathbf{q}' - \mathbf{p}'$, we simplify the expression to be identical to the first term, with the exception of the argument of W_R

$$\langle ad\rangle\langle bf\rangle\langle ce\rangle = \mathcal{I}\left[W_R(p)W_R(p')\left\langle\delta_{-\mathbf{q}}\delta_{-\mathbf{q}'}\right\rangle\left\langle\delta_{\mathbf{p}}\delta_{\mathbf{q}'-\mathbf{p}'}\right\rangle\left\langle\delta_{\mathbf{q}-\mathbf{p}}\delta_{\mathbf{p}'}\right\rangle\right] \tag{D.11}$$

$$= \mathcal{I}\left[W_R(p)W_R(|\mathbf{q}'-\mathbf{p}'|)\left\langle\delta_{-\mathbf{q}}\delta_{-\mathbf{q}'}\right\rangle\left\langle\delta_{\mathbf{p}}\delta_{\mathbf{p}'}\right\rangle\left\langle\delta_{\mathbf{q}-\mathbf{p}}\delta_{\mathbf{q}'-\mathbf{p}'}\right\rangle\right] \tag{D.12}$$

$$= \langle ad\rangle\langle be\rangle\langle cf\rangle \times \frac{W_R(|\mathbf{k}'-\mathbf{p}'|)}{W_R(p')}$$

$$= (2\pi)^3 V^{-1}V_k^{-1}\delta^K_{k,k'}\int_{\mathbf{p}}W_R(|\mathbf{k}-\mathbf{p}|)W_R(p)P(\mathbf{k})P(\mathbf{p})P(\mathbf{k}-\mathbf{p}) \tag{D.13}$$

JCAP08(2026)003

The other four terms contribute non-diagonal terms by contracting the external and internal momenta. We can simplify these terms using the transformations $\mathbf{p} \leftrightarrow \mathbf{k} - \mathbf{p}$ and $\mathbf{p}' \leftrightarrow \mathbf{k}' - \mathbf{p}'$ that we have pointed out above.

$$\langle ae\rangle\langle cd\rangle\langle bf\rangle = \mathcal{I}\left[W_R(p)W_R(p')\left\langle\delta_{-\mathbf{q}}\delta_{\mathbf{p}'}\right\rangle\left\langle\delta_{\mathbf{q}-\mathbf{p}}\delta_{-\mathbf{q}'}\right\rangle\left\langle\delta_{\mathbf{p}}\delta_{\mathbf{q}'-\mathbf{p}'}\right\rangle\right] \tag{D.14}$$

$$\langle ae\rangle\langle cf\rangle\langle bd\rangle = \mathcal{I}\left[W_R(p)W_R(p')\left\langle\delta_{-\mathbf{q}}\delta_{\mathbf{p}'}\right\rangle\left\langle\delta_{\mathbf{q}-\mathbf{p}}\delta_{\mathbf{q}'-\mathbf{p}'}\right\rangle\left\langle\delta_{\mathbf{p}}\delta_{-\mathbf{q}'}\right\rangle\right] \tag{D.15}$$

$$= \mathcal{I}\left[W_R(|\mathbf{q}-\mathbf{p}|)W_R(p')\left\langle\delta_{-\mathbf{q}}\delta_{\mathbf{p}'}\right\rangle\left\langle\delta_{\mathbf{q}-\mathbf{p}}\delta_{-\mathbf{q}'}\right\rangle\left\langle\delta_{\mathbf{p}}\delta_{\mathbf{q}'-\mathbf{p}'}\right\rangle\right] \tag{D.16}$$

$$\langle af\rangle\langle cd\rangle\langle be\rangle = \mathcal{I}\left[W_R(p)W_R(p')\left\langle\delta_{-\mathbf{q}}\delta_{\mathbf{q}'-\mathbf{p}'}\right\rangle\left\langle\delta_{\mathbf{q}-\mathbf{p}}\delta_{-\mathbf{q}'}\right\rangle\left\langle\delta_{\mathbf{p}}\delta_{\mathbf{p}'}\right\rangle\right] \tag{D.17}$$

$$= \mathcal{I}\left[W_R(p)W_R(|\mathbf{q}'-\mathbf{p}'|)\left\langle\delta_{-\mathbf{q}}\delta_{\mathbf{p}'}\right\rangle\left\langle\delta_{\mathbf{q}-\mathbf{p}}\delta_{-\mathbf{q}'}\right\rangle\left\langle\delta_{\mathbf{p}}\delta_{\mathbf{q}'-\mathbf{p}'}\right\rangle\right] \tag{D.18}$$

$$\langle af\rangle\langle ce\rangle\langle bd\rangle = \mathcal{I}\left[W_R(p)W_R(p')\left\langle\delta_{-\mathbf{q}}\delta_{\mathbf{q}'-\mathbf{p}'}\right\rangle\left\langle\delta_{\mathbf{q}-\mathbf{p}}\delta_{\mathbf{p}'}\right\rangle\left\langle\delta_{\mathbf{p}}\delta_{-\mathbf{q}'}\right\rangle\right] \tag{D.19}$$

$$= \mathcal{I}\left[W_R(|\mathbf{q}-\mathbf{p}|)W_R(|\mathbf{q}'-\mathbf{p}'|)\left\langle\delta_{-\mathbf{q}}\delta_{\mathbf{p}'}\right\rangle\left\langle\delta_{\mathbf{q}-\mathbf{p}}\delta_{-\mathbf{q}'}\right\rangle\left\langle\delta_{\mathbf{p}}\delta_{\mathbf{q}'-\mathbf{p}'}\right\rangle\right] \tag{D.20}$$

Thus, the sum of the four terms are

$$\tilde{N}_{222} = \langle ae\rangle\langle cd\rangle\langle bf\rangle + \langle ae\rangle\langle cf\rangle\langle bd\rangle + \langle af\rangle\langle cd\rangle\langle be\rangle + \langle af\rangle\langle ce\rangle\langle bd\rangle \tag{D.21}$$

$$= \mathcal{I}\Big[\left[W_R(p) + W_R(|\mathbf{q}-\mathbf{p}|)\right]\left[W_R(p') + W_R(|\mathbf{q}'-\mathbf{p}'|)\right]\left\langle\delta_{-\mathbf{q}}\delta_{\mathbf{p}'}\right\rangle\left\langle\delta_{\mathbf{q}-\mathbf{p}}\delta_{-\mathbf{q}'}\right\rangle\left\langle\delta_{\mathbf{p}}\delta_{\mathbf{q}'-\mathbf{p}'}\right\rangle\Big] \tag{D.22}$$

showing that we only need to evaluate one contraction. Here, we have used $\tilde{N}_{222}$ to summarize the terms, as these are the non-diagonal contribution from the $(2,2,2)$ contractions. Expanding this term, we find

$$\tilde{N}_{222}$$
$$= \mathcal{I}\Big[\left[W_R(p) + W_R(|\mathbf{q}-\mathbf{p}|)\right]\left[W_R(p') + W_R(|\mathbf{q}'-\mathbf{p}'|)\right]\left\langle\delta_{-\mathbf{q}}\delta_{-\mathbf{p}'}\right\rangle\left\langle\delta_{\mathbf{q}-\mathbf{p}}\delta_{-\mathbf{q}'}\right\rangle\left\langle\delta_{\mathbf{p}}\delta_{\mathbf{q}'-\mathbf{p}'}\right\rangle\Big] \tag{D.23}$$

$$= (2\pi)^9 V^{-2}\int_{V_k}\frac{d^3q}{V_k}\int_{V_k}\frac{d^3q'}{V_k}\int_{\mathbf{p}}\int_{\mathbf{p}'}\left[W_R(p) + W_R(|\mathbf{q}-\mathbf{p}|)\right]\left[W_R(p') + W_R(|\mathbf{q}'-\mathbf{p}'|)\right]$$
$$\delta_D(\mathbf{q}-\mathbf{p}')\delta_D(\mathbf{q}-\mathbf{p}-\mathbf{q}')\delta_D(\mathbf{p}+\mathbf{q}'-\mathbf{p}')P(-\mathbf{q})P(\mathbf{p})P(\mathbf{q}-\mathbf{p}) \tag{D.24}$$

$$= V^{-1}\int_{V_k}\frac{d^3q}{V_k}\int_{V_k}\frac{d^3q'}{V_k}\left[W_R(q) + W_R(|\mathbf{q}-\mathbf{q}'|)\right]\left[W_R(q') + W_R(|\mathbf{q}-\mathbf{q}'|)\right]$$
$$P(\mathbf{q})P(\mathbf{q}')P(\mathbf{q}-\mathbf{q}') \tag{D.25}$$

making the $(2,2,2)$ covariance

$$\mathrm{Cov}\left(\widehat{P}_k, \widehat{M}_{k'}\right)_{222} = (2\pi)^3\delta^K_{k,k'}V^{-1}V_k^{-1}D_{222} + V^{-1}N_{222} \tag{D.26}$$

$$D_{222}(\mathbf{k}) = \int_{\mathbf{p}} W_R(p)\left[W_R(p) + W_R(|\mathbf{k}-\mathbf{p}|)\right]P(\mathbf{k})P(\mathbf{p})P(\mathbf{k}-\mathbf{p}) \tag{D.27}$$

$$N_{222}(\mathbf{k},\mathbf{k}') = V\tilde{N}_{222} \tag{D.28}$$

$$= \left\langle\left[W_R(|\mathbf{k}-\mathbf{k}'|) + W_R(k)\right]\left[W_R(|\mathbf{k}-\mathbf{k}'|) + W_R(k')\right]P(\mathbf{k})P(\mathbf{k}')P(\mathbf{k}-\mathbf{k}')\right\rangle_{\varphi} \tag{D.29}$$

JCAP08(2026)003

Now let us consider the $(3,3)$ contractions. There are 10 possible terms

$$\begin{aligned}\langle abcdef\rangle_{33} = \langle abc\rangle\langle def\rangle + \langle abd\rangle\langle cef\rangle + \langle abe\rangle\langle cdf\rangle + \langle abf\rangle\langle cde\rangle \; + \langle acd\rangle\langle bef\rangle + \\ \langle ace\rangle\langle bdf\rangle + \langle acf\rangle\langle bde\rangle \; + \langle ade\rangle\langle bcf\rangle + \langle adf\rangle\langle bce\rangle \; + \langle aef\rangle\langle bcd\rangle \end{aligned} \tag{D.30}$$

Immediately we notice that the first term involving $\langle abc\rangle$ cancels with the product of the spectra. Let us attempt to group the remaining terms together by invoking the same transformation of the internal momenta $\mathbf{p}$ and $\mathbf{p}'$ as we did above. Specifically, we look for terms that are identical when performing $b \leftrightarrow c$ or $e \leftrightarrow f$, or both. First we recognize the following pair of terms

$$\langle abd\rangle\langle cef\rangle = \mathcal{I}\big[W_R(p)W_R(p')\,\langle\delta_{-\mathbf{q}}\delta_{\mathbf{p}}\delta_{-\mathbf{q}'}\rangle\,\langle\delta_{\mathbf{q}-\mathbf{p}}\delta_{\mathbf{p}'}\delta_{\mathbf{q}'-\mathbf{p}'}\rangle\,\big] \tag{D.31}$$

$$\langle acd\rangle\langle bef\rangle = \mathcal{I}\big[W_R(p)W_R(p')\,\langle\delta_{-\mathbf{q}}\delta_{\mathbf{q}-\mathbf{p}}\delta_{-\mathbf{q}'}\rangle\,\langle\delta_{\mathbf{p}}\delta_{\mathbf{p}'}\delta_{\mathbf{q}'-\mathbf{p}'}\rangle\,\big] \tag{D.32}$$

$$= \mathcal{I}\big[W_R(|\mathbf{q}-\mathbf{p}|)W_R(p')\,\langle\delta_{-\mathbf{q}}\delta_{\mathbf{p}}\delta_{-\mathbf{q}'}\rangle\,\langle\delta_{\mathbf{q}-\mathbf{p}}\delta_{\mathbf{p}'}\delta_{\mathbf{q}'-\mathbf{p}'}\rangle\,\big] \tag{D.33}$$

$$\langle abd\rangle\langle cef\rangle + \langle acd\rangle\langle bef\rangle = \mathcal{I}\Big[\big[W_R(|\mathbf{q}-\mathbf{p}|) + W_R(p)\big]W_R(p')\,\langle\delta_{-\mathbf{q}}\delta_{\mathbf{p}}\delta_{-\mathbf{q}'}\rangle\,\langle\delta_{\mathbf{q}-\mathbf{p}}\delta_{\mathbf{p}'}\delta_{\mathbf{q}'-\mathbf{p}'}\rangle\Big] \tag{D.34}$$

This is nearly identical to the pair

$$\langle ade\rangle\langle bcf\rangle = \mathcal{I}\big[W_R(p)W_R(p')\,\langle\delta_{-\mathbf{q}}\delta_{-\mathbf{q}'}\delta_{\mathbf{p}'}\rangle\,\langle\delta_{\mathbf{p}}\delta_{\mathbf{q}-\mathbf{p}}\delta_{\mathbf{q}'-\mathbf{p}'}\rangle\,\big] \tag{D.35}$$

$$\langle adf\rangle\langle bce\rangle = \mathcal{I}\big[W_R(p)W_R(p')\,\langle\delta_{-\mathbf{q}}\delta_{-\mathbf{q}'}\delta_{\mathbf{q}'-\mathbf{p}'}\rangle\,\langle\delta_{\mathbf{p}}\delta_{\mathbf{q}-\mathbf{p}}\delta_{\mathbf{p}'}\rangle\,\big] \tag{D.36}$$

$$= \mathcal{I}\big[W_R(p)W_R(|\mathbf{q}'-\mathbf{p}'|)\,\langle\delta_{-\mathbf{q}}\delta_{-\mathbf{q}'}\delta_{\mathbf{p}'}\rangle\,\langle\delta_{\mathbf{p}}\delta_{\mathbf{q}-\mathbf{p}}\delta_{\mathbf{q}'-\mathbf{p}'}\rangle\,\big] \tag{D.37}$$

$$\langle ade\rangle\langle bcf\rangle + \langle adf\rangle\langle bce\rangle = \mathcal{I}\Big[\big[W_R(|\mathbf{q}'-\mathbf{p}'|) + W_R(p')\big]W_R(p)\,\langle\delta_{-\mathbf{q}}\delta_{-\mathbf{q}'}\delta_{\mathbf{p}'}\rangle\,\langle\delta_{\mathbf{p}}\delta_{\mathbf{q}-\mathbf{p}}\delta_{\mathbf{q}'-\mathbf{p}'}\rangle\Big] \tag{D.38}$$

other than the substitution $\mathbf{q} \leftrightarrow \mathbf{q}'$. Then, the quartet of terms

$$\langle abe\rangle\langle cdf\rangle = \mathcal{I}\big[W_R(p)W_R(p')\,\langle\delta_{-\mathbf{q}}\delta_{\mathbf{p}}\delta_{\mathbf{p}'}\rangle\,\langle\delta_{\mathbf{q}-\mathbf{p}}\delta_{-\mathbf{q}'}\delta_{\mathbf{q}'-\mathbf{p}'}\rangle\,\big] \tag{D.39}$$

$$\langle abf\rangle\langle cde\rangle = \mathcal{I}\big[W_R(p)W_R(p')\,\langle\delta_{-\mathbf{q}}\delta_{\mathbf{p}}\delta_{\mathbf{q}'-\mathbf{p}'}\rangle\,\langle\delta_{\mathbf{q}-\mathbf{p}}\delta_{-\mathbf{q}'}\delta_{\mathbf{p}'}\rangle\,\big] \tag{D.40}$$

$$= \mathcal{I}\big[W_R(p)W_R(|\mathbf{q}'-\mathbf{p}'|)\,\langle\delta_{-\mathbf{q}}\delta_{\mathbf{p}}\delta_{\mathbf{p}'}\rangle\,\langle\delta_{\mathbf{q}-\mathbf{p}}\delta_{-\mathbf{q}'}\delta_{\mathbf{q}'-\mathbf{p}'}\rangle\,\big] \tag{D.41}$$

$$\langle ace\rangle\langle bdf\rangle = \mathcal{I}\big[W_R(p)W_R(p')\,\langle\delta_{-\mathbf{q}}\delta_{\mathbf{q}-\mathbf{p}}\delta_{\mathbf{p}'}\rangle\,\langle\delta_{\mathbf{p}}\delta_{-\mathbf{q}'}\delta_{\mathbf{q}'-\mathbf{p}'}\rangle\,\big] \tag{D.42}$$

$$= \mathcal{I}\big[W_R(|\mathbf{q}-\mathbf{p}|)W_R(p')\,\langle\delta_{-\mathbf{q}}\delta_{\mathbf{p}}\delta_{\mathbf{p}'}\rangle\,\langle\delta_{\mathbf{q}-\mathbf{p}}\delta_{-\mathbf{q}'}\delta_{\mathbf{q}'-\mathbf{p}'}\rangle\,\big] \tag{D.43}$$

$$\langle acf\rangle\langle bde\rangle = \mathcal{I}\big[W_R(p)W_R(p')\,\langle\delta_{-\mathbf{q}}\delta_{\mathbf{q}-\mathbf{p}}\delta_{\mathbf{q}'-\mathbf{p}'}\rangle\,\langle\delta_{\mathbf{p}}\delta_{-\mathbf{q}'}\delta_{\mathbf{p}'}\rangle\,\big] \tag{D.44}$$

$$= \mathcal{I}\big[W_R(|\mathbf{q}-\mathbf{p}|)W_R(|\mathbf{q}'-\mathbf{p}'|)\,\langle\delta_{-\mathbf{q}}\delta_{\mathbf{p}}\delta_{\mathbf{p}'}\rangle\,\langle\delta_{\mathbf{q}-\mathbf{p}}\delta_{-\mathbf{q}'}\delta_{\mathbf{q}'-\mathbf{p}'}\rangle\,\big] \tag{D.45}$$

$$\begin{aligned}&\langle abe\rangle\langle cdf\rangle + \langle abf\rangle\langle cde\rangle + \langle ace\rangle\langle bdf\rangle + \langle acf\rangle\langle bde\rangle \\ &\quad = \mathcal{I}\Big[\big[W_R(p) + W_R(|\mathbf{q}-\mathbf{p}|)\big]\big[W_R(p') + W_R(|\mathbf{q}'-\mathbf{p}'|)\big]\,\langle\delta_{-\mathbf{q}}\delta_{\mathbf{p}}\delta_{\mathbf{p}'}\rangle\,\langle\delta_{\mathbf{q}-\mathbf{p}}\delta_{-\mathbf{q}'}\delta_{\mathbf{q}'-\mathbf{p}'}\rangle\Big]\end{aligned} \tag{D.46}$$

Finally, the last term $\langle aef\rangle\,\langle bcd\rangle$ has a unique contraction.

JCAP08(2026)003

Let us start our evaluation with the last term $\langle aef\rangle\langle bcd\rangle$

$$\langle aef\rangle\langle bcd\rangle = \mathcal{I}[W_R(p)W_R(p')\langle\delta_{-\mathbf{q}}\delta_{\mathbf{p}'}\delta_{\mathbf{q}'-\mathbf{p}'}\rangle\langle\delta_{\mathbf{p}}\delta_{\mathbf{q}-\mathbf{p}}\delta_{-\mathbf{q}'}\rangle] \tag{D.47}$$

$$= (2\pi)^6 V^{-2}\int_{V_k}\frac{d^3q}{V_k}\int_{V_{k'}}\frac{d^3q'}{V_{k'}}\int_{\mathbf{p}}\int_{\mathbf{p}'}W_R(p)W_R(p')$$

$$\delta_D^2(-\mathbf{q}+\mathbf{q}')B(-\mathbf{q},\mathbf{p}',\mathbf{q}'-\mathbf{p}')B(\mathbf{p},\mathbf{q}-\mathbf{p},-\mathbf{q}') \tag{D.48}$$

$$= (2\pi)^3 V^{-1}V_k^{-1}\delta^K_{k,k'}M(\mathbf{k})M(\mathbf{k}') \tag{D.49}$$

which, we find, is the diagonal contribution. Next, the pairs of terms in eq. (D.34) and (D.38). As they are equivalent under $\mathbf{q}\leftrightarrow\mathbf{q}'$ ($\mathbf{k}\leftrightarrow\mathbf{k}'$), we only need to solve for the first pair

$$\mathcal{I}\Big[[W_R(|\mathbf{q}-\mathbf{p}|)+W_R(p)]W_R(p')\langle\delta_{-\mathbf{q}}\delta_{\mathbf{p}}\delta_{-\mathbf{q}'}\rangle\langle\delta_{\mathbf{q}-\mathbf{p}}\delta_{\mathbf{p}'}\delta_{\mathbf{q}'-\mathbf{p}'}\rangle\Big] \tag{D.50}$$

$$= (2\pi)^6 V^{-2}\int_{V_k}\frac{d^3q}{V_k}\int_{V_{k'}}\frac{d^3q'}{V_{k'}}\int_{\mathbf{p}}\int_{\mathbf{p}'}[W_R(|\mathbf{q}-\mathbf{p}|)+W_R(p)]W_R(p')$$

$$\delta_D^2(\mathbf{q}+\mathbf{q}'-\mathbf{p})B(-\mathbf{q},\mathbf{p},-\mathbf{q}')B(\mathbf{q}-\mathbf{p},\mathbf{p}',\mathbf{q}'-\mathbf{p}') \tag{D.51}$$

$$= V^{-1}\int_{V_k}\frac{d^3q}{V_k}\int_{V_{k'}}\frac{d^3q'}{V_{k'}}\int_{\mathbf{p}'}[W_R(q')+W_R(|\mathbf{q}+\mathbf{q}'|)]W_R(p')$$

$$B(-\mathbf{q},\mathbf{q}+\mathbf{q}',-\mathbf{q}')B(-\mathbf{q}',\mathbf{p}',\mathbf{q}'-\mathbf{p}') \tag{D.52}$$

$$= V^{-1}\langle[W_R(k')+W_R(|\mathbf{k}+\mathbf{k}'|)]B(-\mathbf{k},\mathbf{k}+\mathbf{k}',-\mathbf{k}')M(\mathbf{k}')\rangle_\varphi \tag{D.53}$$

Finally the quartet of terms in eq. (D.46)

$$\mathcal{I}\Big[[W_R(p)+W_R(|\mathbf{q}-\mathbf{p}|)][W_R(p')+W_R(|\mathbf{q}'-\mathbf{p}'|)]\langle\delta_{-\mathbf{q}}\delta_{\mathbf{p}}\delta_{\mathbf{p}'}\rangle\langle\delta_{\mathbf{q}-\mathbf{p}}\delta_{-\mathbf{q}'}\delta_{\mathbf{q}'-\mathbf{p}'}\rangle\Big] \tag{D.54}$$

$$= (2\pi)^6 V^{-2}\int_{V_k}\frac{d^3q}{V_k}\int_{V_{k'}}\frac{d^3q'}{V_{k'}}\int_{\mathbf{p}}\int_{\mathbf{p}'}[W_R(p)+W_R(|\mathbf{q}-\mathbf{p}|)][W_R(p')+W_R(|\mathbf{q}'-\mathbf{p}'|)]$$

$$\delta_D^2(-\mathbf{q}+\mathbf{p}+\mathbf{p}')B(-\mathbf{q},\mathbf{p},\mathbf{p}')B(\mathbf{q}-\mathbf{p},-\mathbf{q}',\mathbf{q}'-\mathbf{p}') \tag{D.55}$$

$$= V^{-1}\Big\langle\int_{\mathbf{p}}[W_R(p)+W_R(|\mathbf{k}-\mathbf{p}|)][W_R(p)+W_R(|\mathbf{k}'-\mathbf{p}|)]$$

$$B(-\mathbf{k},\mathbf{p},\mathbf{k}-\mathbf{p})B(\mathbf{p},-\mathbf{k}',\mathbf{k}'-\mathbf{p})\Big\rangle_\varphi \tag{D.56}$$

where, in the last line we apply the transformation $\mathbf{p}\to\mathbf{k}-\mathbf{p}$ along with the thin-shell approximation.

Let us now consider the $(4,2)$ contributions. There are 15 possible terms, that can be reduced to 9 terms using the same method we had used above for $(2,2,2)$

$$\langle abcdef\rangle_{42} = \langle ad\rangle\langle bcef\rangle + \langle ae\rangle\langle bcdf\rangle + \langle af\rangle\langle bcde\rangle + \langle bd\rangle\langle acef\rangle + \langle be\rangle\langle acdf\rangle +$$
$$\langle bf\rangle\langle acde\rangle + \langle cd\rangle\langle abef\rangle + \langle ce\rangle\langle abdf\rangle + \langle cf\rangle\langle abde\rangle \tag{D.57}$$

Using the transformation $\mathbf{p}\to\mathbf{q}-\mathbf{p}$ we can group terms together

$$\langle ae\rangle\langle bcdf\rangle + \langle af\rangle\langle bcde\rangle = \mathcal{I}\Big[W_R(p)\,[W_R(p')+W_R(|\mathbf{q}'-\mathbf{p}'|)]\,\langle\delta_{-\mathbf{q}}\delta_{\mathbf{p}'}\rangle\langle\delta_{\mathbf{p}}\delta_{\mathbf{q}-\mathbf{p}}\delta_{-\mathbf{q}'}\delta_{\mathbf{q}'-\mathbf{p}'}\rangle\Big] \tag{D.58}$$

JCAP08(2026)003

Similarly, $\langle bd\rangle\langle acef\rangle + \langle cd\rangle\langle abef\rangle$ yield the equivalent expression with $\mathbf{k} \leftrightarrow \mathbf{k}'$. The terms

$$\begin{aligned}&\langle be\rangle\langle acdf\rangle + \langle bf\rangle\langle acde\rangle + \langle ce\rangle\langle abdf\rangle + \langle cf\rangle\langle abde\rangle \\ &= \mathcal{I}\Big[\left[W_R(p) + W_R(|\mathbf{q}-\mathbf{p}|)\right]\left[W_R(p') + W_R(|\mathbf{q}'-\mathbf{p}'|)\right]\langle\delta_{\mathbf{p}}\delta_{\mathbf{p}'}\rangle\langle\delta_{-\mathbf{q}}\delta_{\mathbf{q}-\mathbf{p}}\delta_{-\mathbf{q}'}\delta_{\mathbf{q}'-\mathbf{p}'}\rangle\Big]\end{aligned} \tag{D.59}$$

also arise from equivalent contractions. With the unique term $\langle ad\rangle\langle bcef\rangle$, we have three unique contractions to evaluate. Each of these terms become

$$\langle ad\rangle\langle bcef\rangle = (2\pi)^3 V^{-1} V_k^{-1}\delta^K_{k,k'}\int_{\mathbf{p}}\int_{\mathbf{p}'} W_R(p)W_R(p')P(\mathbf{k})T(\mathbf{p},\mathbf{k}-\mathbf{p},\mathbf{p}',-\mathbf{k}-\mathbf{p}') \tag{D.60}$$

$$\begin{aligned}&\langle ae\rangle\langle bcdf\rangle + \langle af\rangle\langle bcde\rangle \\ &= V^{-1}\left\langle\int_{\mathbf{p}} W_R(p)[W_R(k) + W_R(|\mathbf{k}-\mathbf{k}'|)]P(\mathbf{k})T(\mathbf{p},\mathbf{k}-\mathbf{p},-\mathbf{k}',\mathbf{k}'-\mathbf{k})\right\rangle_\varphi\end{aligned} \tag{D.61}$$

$$\begin{aligned}&\langle be\rangle\langle acdf\rangle + \langle bf\rangle\langle acde\rangle + \langle ce\rangle\langle abdf\rangle + \langle cf\rangle\langle abde\rangle \\ &= V^{-1}\left\langle\int_{\mathbf{p}}[W_R(k) + W_R(|\mathbf{k}-\mathbf{k}'|)][W_R(k') + W_R(|\mathbf{k}-\mathbf{k}'|)]P_{\mathbf{p}}T(-\mathbf{k},\mathbf{k}-\mathbf{p},-\mathbf{k}',\mathbf{k}'+\mathbf{p})\right\rangle_\varphi\end{aligned} \tag{D.62}$$

This leaves us to evaluate the final term in the covariance, contracting all 6 fields together. This yields

$$\mathrm{Cov}\left(\widehat{M}_k,\widehat{M}'_k\right)_6 = \langle abcdef\rangle = V^{-1}\left\langle\int_{\mathbf{p}}\int_{\mathbf{p}'} W_R(p)W_R(p')P_6(-\mathbf{k},\mathbf{p},\mathbf{k}-\mathbf{p},-\mathbf{k}',\mathbf{p}',\mathbf{k}'-\mathbf{p}')\right\rangle_\varphi \tag{D.63}$$

The final expression for the covariance becomes

$$\mathrm{Cov}\left(\widehat{M}_k,\widehat{M}'_k\right) = (2\pi)^3\delta^K_{k,k'}V^{-1}V_k^{-1}\left[D_{22}+D_{33}+D_{42}\right] + V^{-1}[N_{222}+N_{33}+N_{42}+N_6] \tag{D.64}$$

with diagonal terms

$$D_{222}(\mathbf{k}) = \int_{\mathbf{p}} W_R(p)\big[W_R(p) + W_R(|\mathbf{k}-\mathbf{p}|)\big]P(\mathbf{k})P(\mathbf{p})P(\mathbf{k}-\mathbf{p}) \tag{D.65}$$

$$D_{33}(\mathbf{k}) = M(\mathbf{k})^2 \tag{D.66}$$

$$D_{42}(\mathbf{k}) = \int_{\mathbf{p}}\int_{\mathbf{p}'} W_R(p)W_R(p')P(\mathbf{k})T(\mathbf{p},\mathbf{k}-\mathbf{p},\mathbf{p}',-\mathbf{k}-\mathbf{p}') \tag{D.67}$$

and non-diagonal terms

$$N_{222}(\mathbf{k},\mathbf{k}') = \big\langle\big[W_R(|\mathbf{k}-\mathbf{k}'|) + W_R(k)\big]\big[W_R(|\mathbf{k}-\mathbf{k}'|) + W_R(k')\big]P(\mathbf{k})P(\mathbf{k}')P(\mathbf{k}-\mathbf{k}')\big\rangle_\varphi \tag{D.68}$$

$$\begin{aligned}N_{33}(\mathbf{k},\mathbf{k}') = &\big\langle\big[W_R(k') + W_R(|\mathbf{k}+\mathbf{k}'|)\big]B(-\mathbf{k},\mathbf{k}+\mathbf{k}',-\mathbf{k}')M(\mathbf{k}')\big\rangle_\varphi \\ &+\Big\langle\int_{\mathbf{p}}\big[W_R(p) + W_R(|\mathbf{k}-\mathbf{p}|)\big]\big[W_R(p) + W_R(|\mathbf{k}'-\mathbf{p}|)\big] \\ &B(-\mathbf{k},\mathbf{p},\mathbf{k}-\mathbf{p})B(\mathbf{p},-\mathbf{k}',\mathbf{k}'-\mathbf{p})\Big\rangle_\varphi\end{aligned} \tag{D.69}$$

JCAP08(2026)003

$$N_{42}(\mathbf{k},\mathbf{k}') = \left\langle \int_{\mathbf{p}} W_R(p)[W_R(k) + W_R(|\mathbf{k}-\mathbf{k}'|)]P(\mathbf{k})T(\mathbf{p},\mathbf{k}-\mathbf{p},-\mathbf{k}',\mathbf{k}'-\mathbf{k}) \right\rangle_{\varphi}$$
$$+ \left\langle \int_{\mathbf{p}} [W_R(k) + W_R(|\mathbf{k}-\mathbf{k}'|)][W_R(k') + W_R(|\mathbf{k}-\mathbf{k}'|)]P_{\mathbf{p}}T(-\mathbf{k},\mathbf{k}-\mathbf{p},-\mathbf{k}',\mathbf{k}'+\mathbf{p}) \right\rangle_{\varphi} \tag{D.70}$$

$$N_6(\mathbf{k},\mathbf{k}') = \left\langle \int_{\mathbf{p}}\int_{\mathbf{p}'} W_R(p)W_R(p')P_6(-\mathbf{k},\mathbf{p},\mathbf{k}-\mathbf{p},-\mathbf{k}',\mathbf{p}',\mathbf{k}'-\mathbf{p}') \right\rangle_{\varphi} \tag{D.71}$$

Before moving on to the P-M covariance, let us compare this with the result under the Gaussian approximation. The Gaussian (disconnected) covariance can be computed by assuming that each field involved (δ and $\tilde{\delta}_M$) is Gaussian. Invoking Wick's theorem [86], one finds

$$\mathrm{Cov}_G\left(\widehat{P}_k, \widehat{M}_{k'}\right) = V^{-2}\int_{V_k}\frac{d^3q}{V_k}\int_{V_k}\frac{d^3q'}{V_k}\left(\left\langle \tilde{\delta}_{M,-q}\delta_q\tilde{\delta}_{M,-q'}\delta_{q'}\right\rangle - \left\langle \tilde{\delta}_{M,-q}\delta_q\right\rangle\left\langle \tilde{\delta}_{M,-q'}\delta_{q'}\right\rangle\right) \tag{D.72}$$

$$= V^{-2}\int_{V_k}\frac{d^3q}{V_k}\int_{V_k}\frac{d^3q'}{V_k}\left\langle \tilde{\delta}_{M,-q}\delta_{q'}\right\rangle\left\langle \delta_k\tilde{\delta}_{M,-q'}\right\rangle + \left\langle \tilde{\delta}_{M,-q}\tilde{\delta}_{M,-q'}\right\rangle\langle\delta_q\delta_{q'}\rangle \tag{D.73}$$

$$= (2\pi)^3\delta^K_{k,k'}V^{-1}V_k^{-1}\left[M^2(\mathbf{k}) + P(\mathbf{k})V^{-1}\left\langle \tilde{\delta}_{M,k}\tilde{\delta}_{M,-k}\right\rangle\right] \tag{D.74}$$

The only term above that cannot be expressed as P or M is $\left\langle \tilde{\delta}_{M,k}\tilde{\delta}_{M,-k}\right\rangle$. Expanding this, we find

$$\left\langle \tilde{\delta}_{M,-k}\tilde{\delta}_{M,-k}\right\rangle = \int_{V_k}\frac{d^3q}{V_k}\left\langle \tilde{\delta}_{M,\mathbf{q}}\tilde{\delta}_{M,-\mathbf{q}}\right\rangle \tag{D.75}$$

$$= \int_{V_k}\frac{d^3q}{V_k}\int_{\mathbf{p}}\int_{\mathbf{p}'} W_R(p)W_R(p')\left\langle \delta_{\mathbf{p}}\delta_{\mathbf{q}-\mathbf{p}}\delta_{\mathbf{p}'}\delta_{-\mathbf{q}-\mathbf{p}'}\right\rangle \tag{D.76}$$

The four-point correlator contributes both $(2,2)$ disconnected contractions and a connected four-point contraction. One of the disconnected terms, involving $\langle\delta_{\mathbf{p}}\delta_{\mathbf{q}-\mathbf{p}}\rangle$ vanishes. The other disconnected terms contribute

$$\left\langle \tilde{\delta}_{M,-k}\tilde{\delta}_{M,-k}\right\rangle_{22} = V\int_{V_k}\frac{d^3q}{V_k}\int_{\mathbf{p}} W_R(p)\left[W_R(p) + W_R(|\mathbf{q}-\mathbf{p}|)\right]P(\mathbf{q})P(\mathbf{q}-\mathbf{p}) \tag{D.77}$$

$$= V\int_{\mathbf{p}} W_R(p)\left[W_R(p) + W_R(|\mathbf{k}-\mathbf{p}|)\right]P(\mathbf{k})P(\mathbf{k}-\mathbf{p}) \tag{D.78}$$

while the connected piece yields

$$\left\langle \tilde{\delta}_{M,-k}\tilde{\delta}_{M,-k}\right\rangle_4 = \int_{V_k}\frac{d^3q}{V_k}\int_{\mathbf{p}}\int_{\mathbf{p}'} W_R(p)W_R(p')T(\mathbf{p},\mathbf{q}-\mathbf{p},\mathbf{p}',-\mathbf{q}-\mathbf{p}') \tag{D.79}$$

Thus the Gaussian covariance becomes

$$\mathrm{Cov}_G\left(\widehat{P}_k, \widehat{M}_{k'}\right) = (2\pi)^3\delta^K_{k,k'}V^{-1}V_k^{-1}[D_{222}(k) + D_{42}(k) + D_{33}(k)] \tag{D.80}$$

indicating that the Gaussian approximation correctly captures the full diagonal contribution, which is dominant at large scales.

JCAP08(2026)003

D.2 Covariance of *P*-*M*

We now calculate the cross-covariance between P and M (section 3.2). The expression for the cross-covariance is

$$\mathrm{Cov}\left(\widehat{P}_k, \widehat{M}_{k'}\right) = V^{-2}\int_{V_k}\frac{d^3q}{V_k}\int_{V_k}\frac{d^3q'}{V_k}\int_{\mathbf{p}'} W_R(p') \tag{D.81}$$

$$\left(\left\langle\delta_{-\mathbf{q}}\delta_{\mathbf{q}}\delta_{-\mathbf{q}'}\delta_{\mathbf{p}'}\delta_{\mathbf{q}'-\mathbf{p}'}\right\rangle - \left\langle\delta_{-\mathbf{q}}\delta_{\mathbf{q}}\right\rangle\left\langle\delta_{-\mathbf{q}'}\delta_{\mathbf{p}'}\delta_{\mathbf{q}'-\mathbf{p}'}\right\rangle\right) \tag{D.82}$$

$$= \mathcal{I}[W_R(p')\left\langle\delta_{-\mathbf{q}}\delta_{\mathbf{q}}\delta_{-\mathbf{q}'}\delta_{\mathbf{p}'}\delta_{\mathbf{q}'-\mathbf{p}'}\right\rangle] - P_k M_{k'} \tag{D.83}$$

where $\mathcal{I}$ is now

$$\mathcal{I}[f] = V^{-2}\int_{V_k}\frac{d^3q\,d^3q'}{V_k^2}\int_{\mathbf{p}'} f(\mathbf{q},\mathbf{q}',\mathbf{p}') \tag{D.84}$$

We again start by considering the possible permutations of the five-point correlator. As $\langle\delta\rangle = 0$, we are restricted to considering contractions of the form $(3,2)$ or (5). There are 10 possible contractions of the (32) form

$$\langle abcde\rangle_{32} = \langle ab\rangle\langle cde\rangle + \langle ac\rangle\langle bde\rangle + \langle ad\rangle\langle bce\rangle + \langle ae\rangle\langle bcd\rangle + \langle bc\rangle\langle ade\rangle$$
$$+ \langle bd\rangle\langle ace\rangle + \langle be\rangle\langle acd\rangle + \langle cd\rangle\langle abe\rangle + \langle ce\rangle\langle abd\rangle + \langle de\rangle\langle abc\rangle \tag{D.85}$$

$$\text{with} \qquad a = \delta_{-\mathbf{q}}, \quad b = \delta_{\mathbf{q}}, \quad c = \delta_{-\mathbf{q}'}, \quad d = \delta_{\mathbf{p}'}, \quad e = \delta_{\mathbf{q}'-\mathbf{p}'} \tag{D.86}$$

Immediately we notice that the first term involving $\langle ab\rangle$ cancels with the product of the spectra. Of the remaining terms, we again take advantage of the property that the momenta of $c = \delta_{-\mathbf{q}'}$, $d = \delta_{\mathbf{p}'}$, and $e = \delta_{\mathbf{q}'-\mathbf{p}'}$ sum to zero, hence having an ensemble average over two of the fields results in the third field being evaluated at zero momentum. This eliminates the last three terms, leaving us with six. Then, notice that $a = \delta_{-\mathbf{q}}$ and $b = \delta_{\mathbf{q}}$ merely differ by the sign of $\mathbf{q}$, which can be dropped by symmetry. Thus the final contribution will be identical to

$$2\left[\langle ac\rangle\langle bde\rangle + \langle ad\rangle\langle bce\rangle + \langle ae\rangle\langle bcd\rangle\right] \tag{D.87}$$

Let us start the evaluation with

$$\mathcal{I}[W_R(p')\langle ac\rangle\langle bde\rangle] = \mathcal{I}[W_R(p')\langle\delta_{-\mathbf{q}}\delta_{-\mathbf{q}'}\rangle\langle\delta_{\mathbf{q}}\delta_{\mathbf{p}'}\delta_{\mathbf{q}'-\mathbf{p}'}\rangle] \tag{D.88}$$

$$= \mathcal{I}[(2\pi)^6 W_R(p')\delta_D^2(\mathbf{q}+\mathbf{q}')B(\mathbf{p}',\mathbf{q}-\mathbf{p}',-\mathbf{q})P(-\mathbf{q})] \tag{D.89}$$

$$= V^{-1}V_k^{-1}\delta^K_{k,-k'}P(-\mathbf{k})\int_{\mathbf{p}'} W_R(p')B(\mathbf{p}',\mathbf{k}-\mathbf{p}',-\mathbf{k}) \tag{D.90}$$

$$= V^{-1}V_k^{-1}\delta^K_{k,k'}P(\mathbf{k})M(\mathbf{k}) \tag{D.91}$$

where in the last line we drop the sign of $\mathbf{k}$ due to symmetry. We notice that this is the diagonal contribution to the covariance. Next, we have

$$\mathcal{I}[W_R(p')\langle ad\rangle\langle bce\rangle] = \mathcal{I}[W_R(p')\langle\delta_{-\mathbf{q}}\delta_{\mathbf{p}'}\rangle\langle\delta_{\mathbf{q}}\delta_{-\mathbf{q}'}\delta_{\mathbf{q}'-\mathbf{p}'}\rangle] \tag{D.92}$$

$$= \mathcal{I}[(2\pi)^6 W_R(p')\delta_D^2(\mathbf{q}-\mathbf{p}')B(\mathbf{q},-\mathbf{q}',\mathbf{q}'-\mathbf{p}')P(-\mathbf{q})] \tag{D.93}$$

$$= V^{-1}V_k^{-2}\int_{V_k} d^3q\int_{V_k} d^3q' W_R(q)P(\mathbf{q})B(\mathbf{q},-\mathbf{q}',\mathbf{q}'-\mathbf{q}) \tag{D.94}$$

$$= V^{-1}W_R(k)P(\mathbf{k})\left\langle B(\mathbf{k},-\mathbf{k}',\mathbf{k}'-\mathbf{k})\right\rangle_\varphi \tag{D.95}$$

JCAP08(2026)003

Finally,

$$\mathcal{I}[W_R(p')\langle ae\rangle\langle bcd\rangle] = \mathcal{I}[W_R(p')\langle\delta_{-\mathbf{q}}\delta_{\mathbf{q}'-\mathbf{p}'}\rangle\langle\delta_{\mathbf{q}}\delta_{-\mathbf{q}'}\delta_{\mathbf{p}'}\rangle] \tag{D.96}$$

$$= \mathcal{I}[(2\pi)^6 W_R(p')\delta_D^2(\mathbf{q}-\mathbf{q}'+\mathbf{p}')B(\mathbf{q},-\mathbf{q}',\mathbf{p}')P(-\mathbf{q})] \tag{D.97}$$

$$= V^{-1}V_k^{-2}\int_{V_k} d^3q \int_{V_k} d^3q' W_R(|\mathbf{q}'-\mathbf{q}|)P(\mathbf{q})B(\mathbf{q},-\mathbf{q}',\mathbf{q}'-\mathbf{q}) \tag{D.98}$$

$$= V^{-1}P(\mathbf{k})\left\langle W_R(|\mathbf{k}'-\mathbf{k}|)B(\mathbf{k},-\mathbf{k}',\mathbf{k}'-\mathbf{k})\right\rangle_\varphi \tag{D.99}$$

Now we consider the only other contribution, contracting all five operators

$$\mathcal{I}[W_R(p')\langle abcde\rangle] = \mathcal{I}[W_R(p')\left\langle\delta_{-\mathbf{q}}\delta_{\mathbf{q}}\delta_{-\mathbf{q}'}\delta_{\mathbf{p}'}\delta_{\mathbf{q}'-\mathbf{p}'}\right\rangle] \tag{D.100}$$

$$= \mathcal{I}\left[(2\pi)^3 W_R(p')\delta_D(0)P_5\left(-\mathbf{q},\mathbf{q},-\mathbf{q}',\mathbf{p}',\mathbf{q}'-\mathbf{p}'\right)\right] \tag{D.101}$$

$$= V^{-1}V_k^{-2}\int_{V_k} d^3q \int_{V_k} d^3q' \int_{\mathbf{p}'} W_R(p')P_5\left(-\mathbf{q},\mathbf{q},-\mathbf{q}',\mathbf{p}',\mathbf{q}'-\mathbf{p}'\right) \tag{D.102}$$

$$= V^{-1}\left\langle\int_{\mathbf{p}'} W_R(p')P_5\left(-\mathbf{k},\mathbf{k},-\mathbf{k}',\mathbf{p}',\mathbf{k}'-\mathbf{p}'\right)\right\rangle_\varphi \tag{D.103}$$

Thus the final contributions are

$$\mathrm{Cov}\left(\widehat{P}_k,\widehat{M}_{k'}\right) = (2\pi)^3\delta^K_{k,k'}V^{-1}V_k^{-1}D_{23} + V^{-1}\left[N_{23}+N_5\right] \tag{D.104}$$

$$D_{23}(\mathbf{k}) = 2P(\mathbf{k})M(\mathbf{k}) \tag{D.105}$$

$$N_{23}(\mathbf{k},\mathbf{k}') = 2P(\mathbf{k})\left\langle(W_R(k)+W_R(|\mathbf{k}'-\mathbf{k}|))B(\mathbf{k},-\mathbf{k}',\mathbf{k}'-\mathbf{k})\right\rangle_\varphi \tag{D.106}$$

$$N_5 = 2\left\langle\int_{\mathbf{p}'} W_R(p')P_5\left(-\mathbf{k},\mathbf{k},-\mathbf{k}',\mathbf{p}',\mathbf{k}'-\mathbf{p}'\right)\right\rangle_\varphi \tag{D.107}$$

As with the auto-covariance, we observe that the disconnected term dominates at large-scales due to the V_k^{-1} scaling, but the non-diagonal terms become important as we go to small scales.

Now we compute the cross-covariance assuming that the fields involved are Gaussian. Using Wick's theorem [21, 86],

$$\mathrm{Cov}_G\left(\widehat{P}_k,\widehat{M}_{k'}\right) = V^{-2}\int_{V_k}\frac{d^3q}{V_k}\int_{V_k}\frac{d^3q'}{V_k}\left(\left\langle\delta_{-q}\delta_q\tilde{\delta}_{M,-q'}\delta_{q'}\right\rangle - \langle\delta_{-q}\delta_q\rangle\left\langle\tilde{\delta}_{M,-q'}\delta_{q'}\right\rangle\right) \tag{D.108}$$

$$= V^{-2}\int_{V_k}\frac{d^3q}{V_k}\int_{V_k}\frac{d^3q'}{V_k}\langle\delta_{-q}\delta_{q'}\rangle\left\langle\delta_q\tilde{\delta}_{M,-q'}\right\rangle + \left\langle\delta_{-q}\tilde{\delta}_{M,-q'}\right\rangle\langle\delta_q\delta_{q'}\rangle \tag{D.109}$$

$$= (2\pi)^3\delta^K_{k,k'}V^{-1}V_k^{-1}2P(\mathbf{k})M(\mathbf{k}') \tag{D.110}$$

$$= (2\pi)^3\delta^K_{k,k'}V^{-1}V_k^{-1}D_{23} \tag{D.111}$$

Thus we find again that the Gaussian approximation recovers the diagonal entries of the covariance matrix. At the large-scales that M is applicable, these terms are expected to dominate.

One can extend these calculations to redshift space by utilizing the multipole expansion of power spectra when invoking the thin-shell approximation: $P(\mathbf{q}) = \sum_\ell P(q)\mathcal{L}_\ell(\mu) \approx \sum_\ell P(k)\mathcal{L}_\ell(\mu) = P(\mathbf{k})$.

JCAP08(2026)003

References

[1] F. Bernardeau, S. Colombi, E. Gaztanaga and R. Scoccimarro, *Large scale structure of the universe and cosmological perturbation theory*, *Phys. Rept.* **367** (2002) 1 [astro-ph/0112551] [INSPIRE].

[2] D. Baumann, *Cosmology*, Cambridge University Press (2022).

[3] S. Ferraro, N. Sailer, A. Slosar and M. White, *Snowmass2021 Cosmic Frontier White Paper: Cosmology and Fundamental Physics from the three-dimensional Large Scale Structure*, arXiv:2203.07506 [INSPIRE].

[4] DESI collaboration, *Data Release 1 of the Dark Energy Spectroscopic Instrument*, *Astron. J.* **171** (2026) 285 [arXiv:2503.14745] [INSPIRE].

[5] DESI collaboration, *DESI DR2 results. II. Measurements of baryon acoustic oscillations and cosmological constraints*, *Phys. Rev. D* **112** (2025) 083515 [arXiv:2503.14738] [INSPIRE].

[6] DESI collaboration, *DESI 2024 V: Full-Shape galaxy clustering from galaxies and quasars*, *JCAP* **09** (2025) 008 [*Erratum ibid.* **02** (2026) E02] [arXiv:2411.12021] [INSPIRE].

[7] E.J. Groth and P.J.E. Peebles, *Statistical analysis of catalogs of extragalactic objects. V. Three-point correlation function for the galaxy distribution in the Zwicky catalog*, *Astrophys. J.* **196** (1975) 1.

[8] J.N. Fry and M. Seldner, *Transform analysis of the high-resolution Shane-Wirtanen Catalog — The power spectrum and the bispectrum*, *Astrophys. J.* **259** (1982) 474.

[9] J.N. Fry and P.J.E. Peebles, *Statistical analysis of catalogs of extragalactic objects. IX. The four-point galaxy correlation function*, *Astrophys. J.* **221** (1978) 19.

[10] W. Hu, *Angular trispectrum of the CMB*, *Phys. Rev. D* **64** (2001) 083005 [astro-ph/0105117] [INSPIRE].

[11] M. Schmittfull, T. Baldauf and U. Seljak, *Near optimal bispectrum estimators for large-scale structure*, *Phys. Rev. D* **91** (2015) 043530 [arXiv:1411.6595] [INSPIRE].

[12] J. Hou et al., *Cosmological constraints from the redshift-space galaxy skew spectra*, *Phys. Rev. D* **109** (2024) 103528 [arXiv:2401.15074] [INSPIRE].

[13] S. Cheng, Y.-S. Ting, B. Ménard and J. Bruna, *A new approach to observational cosmology using the scattering transform*, *Mon. Not. Roy. Astron. Soc.* **499** (2020) 5902 [arXiv:2006.08561] [INSPIRE].

[14] G. Valogiannis and C. Dvorkin, *Towards an optimal estimation of cosmological parameters with the wavelet scattering transform*, *Phys. Rev. D* **105** (2022) 103534 [arXiv:2108.07821] [INSPIRE].

[15] M. Eickenberg et al., *Wavelet Moments for Cosmological Parameter Estimation*, arXiv:2204.07646 [INSPIRE].

[16] S. Cheng et al., *Cosmological constraints from weak lensing scattering transform using HSC Y1 data*, *JCAP* **01** (2025) 006 [arXiv:2404.16085] [INSPIRE].

[17] DES collaboration, *Weak lensing by galaxy troughs in DES Science Verification data*, *Mon. Not. Roy. Astron. Soc.* **455** (2016) 3367 [arXiv:1507.05090] [INSPIRE].

[18] DES collaboration, *Density Split Statistics: Joint Model of Counts and Lensing in Cells*, *Phys. Rev. D* **98** (2018) 023508 [arXiv:1710.05162] [INSPIRE].

JCAP08(2026)003

[19] E. Paillas, Y.-C. Cai, N. Padilla and A.G. Sánchez, *Redshift-space distortions with split densities*, *Mon. Not. Roy. Astron. Soc.* **505** (2021) 5731 [arXiv:2101.09854] [INSPIRE].

[20] E. Paillas et al., *Cosmological constraints from density-split clustering in the BOSS CMASS galaxy sample*, *Mon. Not. Roy. Astron. Soc.* **531** (2024) 898 [arXiv:2309.16541] [INSPIRE].

[21] L. Harscouet et al., *Fast Projected Bispectra: the filter-square approach*, arXiv:2409.07980 [DOI:10.33232/001c.128309] [INSPIRE].

[22] F. Verdiani et al., *Cosmological constraints from the angular power spectrum and bispectrum of luminous red galaxies and CMB lensing*, *JCAP* **06** (2026) 027 [arXiv:2510.17796] [INSPIRE].

[23] BEYOND-2PT collaboration, *A Parameter-Masked Mock Data Challenge for Beyond-Two-Point Galaxy Clustering Statistics*, *Astrophys. J.* **990** (2025) 99 [arXiv:2405.02252] [INSPIRE].

[24] R. Takahashi et al., *Fitting the nonlinear matter bispectrum by the Halofit approach*, *Astrophys. J.* **895** (2020) 113 [arXiv:1911.07886] [INSPIRE].

[25] O.H.E. Philcox, *Probing parity violation with the four-point correlation function of BOSS galaxies*, *Phys. Rev. D* **106** (2022) 063501 [arXiv:2206.04227] [INSPIRE].

[26] U. Giri, M. Münchmeyer and K.M. Smith, *Robust neural network-enhanced estimation of local primordial non-Gaussianity*, *Phys. Rev. D* **107** (2023) L061301 [arXiv:2205.12964] [INSPIRE].

[27] U. Giri, M. Münchmeyer and K.M. Smith, *Constraining fNL using the large-scale modulation of small-scale statistics*, *Phys. Rev. D* **112** (2025) 123544 [arXiv:2305.03070] [INSPIRE].

[28] F. Bernardeau, L. Van Waerbeke and Y. Mellier, *Weak lensing statistics as a probe of Omega and power spectrum*, *Astron. Astrophys.* **322** (1997) 1 [astro-ph/9609122] [INSPIRE].

[29] S. Pires, A. Leonard and J.-L. Starck, *Cosmological constraints from the capture of non-Gaussianity in Weak Lensing data*, *Mon. Not. Roy. Astron. Soc.* **423** (2012) 983 [arXiv:1203.2877] [INSPIRE].

[30] C. Hahn and F. Villaescusa-Navarro, *Constraining M_ν with the bispectrum. Part II. The information content of the galaxy bispectrum monopole*, *JCAP* **04** (2021) 029 [arXiv:2012.02200] [INSPIRE].

[31] M. White, *A marked correlation function for constraining modified gravity models*, *JCAP* **11** (2016) 057 [arXiv:1609.08632] [INSPIRE].

[32] H. Ebina and M. White, *An analytically tractable marked power spectrum*, *JCAP* **01** (2025) 150 [arXiv:2409.17133] [INSPIRE].

[33] J.A. Cowell et al., *First constraints from marked angular power spectra with Subaru Hyper Suprime-Cam Survey First-Year Data*, *Mon. Not. Roy. Astron. Soc.* **546** (2026) stag033 [arXiv:2507.12315] [INSPIRE].

[34] E. Massara et al., *Using the Marked Power Spectrum to Detect the Signature of Neutrinos in Large-Scale Structure*, *Phys. Rev. Lett.* **126** (2021) 011301 [arXiv:2001.11024] [INSPIRE].

[35] E. Massara et al., *Cosmological Information in the Marked Power Spectrum of the Galaxy Field*, *Astrophys. J.* **951** (2023) 70 [arXiv:2206.01709] [INSPIRE].

[36] 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** (2020) 043516 [arXiv:2006.10055] [INSPIRE].

[37] O.H.E. Philcox, A. Aviles and E. Massara, *Modeling the Marked Spectrum of Matter and Biased Tracers in Real- and Redshift-Space*, *JCAP* **03** (2021) 038 [arXiv:2010.05914] [INSPIRE].

JCAP08(2026)003

[38] J.A. Cowell, D. Alonso and J. Liu, *Optimizing marked power spectra for cosmology*, *Mon. Not. Roy. Astron. Soc.* **535** (2024) 3129 [arXiv:2409.05695] [INSPIRE].

[39] C. Beisbart and M. Kerscher, *Luminosity- and morphology-dependent clustering of galaxies*, *Astrophys. J.* **545** (2000) 6 [astro-ph/0003358] [INSPIRE].

[40] R.K. Sheth, *The halo-model description of marked statistics*, *Mon. Not. Roy. Astron. Soc.* **364** (2005) 796 [astro-ph/0511772] [INSPIRE].

[41] Z. Gao, A. Moradinezhad Dizgah and Z. Vlah, *Parity in composite-field galaxy correlators*, *JCAP* **03** (2026) 027 [arXiv:2509.13207] [INSPIRE].

[42] A. Cooray, *Squared temperature-temperature power spectrum as a probe of the CMB bispectrum*, *Phys. Rev. D* **64** (2001) 043516 [astro-ph/0105415] [INSPIRE].

[43] D. Munshi and A. Heavens, *A New Approach to Probing Primordial Non-Gaussianity*, *Mon. Not. Roy. Astron. Soc.* **401** (2010) 2406 [arXiv:0904.4478] [INSPIRE].

[44] A. Moradinezhad Dizgah, H. Lee, M. Schmittfull and C. Dvorkin, *Capturing non-Gaussianity of the large-scale structure with weighted skew-spectra*, *JCAP* **04** (2020) 011 [arXiv:1911.05763] [INSPIRE].

[45] M.M. Ivanov, *Effective Field Theory for Large-Scale Structure*, in *Handbook of Quantum Gravity*, Springer, Singapore (2023) [DOI:10.1007/978-981-19-3079-9_5-1] [arXiv:2212.08488] [INSPIRE].

[46] N. Kaiser, *Clustering in real space and in redshift space*, *Mon. Not. Roy. Astron. Soc.* **227** (1987) 1 [INSPIRE].

[47] S.-F. Chen, Z. Vlah and M. White, *Consistent Modeling of Velocity Statistics and Redshift-Space Distortions in One-Loop Perturbation Theory*, *JCAP* **07** (2020) 062 [arXiv:2005.00523] [INSPIRE].

[48] M.M. Ivanov et al., *Precision analysis of the redshift-space galaxy bispectrum*, *Phys. Rev. D* **105** (2022) 063512 [arXiv:2110.10161] [INSPIRE].

[49] 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 [INSPIRE].

[50] A. Chudaykin, M.M. Ivanov and O.H.E. Philcox, *Reanalyzing DESI DR1. I. ΛCDM constraints from the power spectrum and bispectrum*, *Phys. Rev. D* **113** (2026) 063502 [arXiv:2507.13433] [INSPIRE].

[51] M. Maus et al., *An analysis of parameter compression and Full-Modeling techniques with Velocileptors for DESI 2024 and beyond*, *JCAP* **01** (2025) 138 [arXiv:2404.07312] [INSPIRE].

[52] M.M. Ivanov et al., *Full-shape analysis with simulation-based priors: Constraints on single field inflation from BOSS*, *Phys. Rev. D* **110** (2024) 063538 [arXiv:2402.13310] [INSPIRE].

[53] M. Biagetti, L. Castiblanco, J. Noreña and E. Sefusatti, *The covariance of squeezed bispectrum configurations*, *JCAP* **09** (2022) 009 [arXiv:2111.05887] [INSPIRE].

[54] J. Hou, A. Moradinezhad Dizgah, C.H. Hahn and E. Massara, *Cosmological information in skew spectra of biased tracers in redshift space*, *JCAP* **03** (2023) 045 [arXiv:2210.12743] [INSPIRE].

[55] J. Salvalaggio et al., *Bispectrum non-Gaussian covariance in redshift space*, *JCAP* **08** (2024) 046 [arXiv:2403.08634] [INSPIRE].

[56] DESI collaboration, *FolpsD: combining EFT and phenomenological approaches for joint power spectrum and bispectrum analyses*, arXiv:2604.08895 [DOI:10.5281/zenodo.19302052] [INSPIRE].

JCAP08(2026)003

[57] R. Scoccimarro et al., *Nonlinear evolution of the bispectrum of cosmological perturbations*, *Astrophys. J.* **496** (1998) 586 [astro-ph/9704075] [INSPIRE].

[58] J.-P. Dai, L. Verde and J.-Q. Xia, *What Can We Learn by Combining the Skew Spectrum and the Power Spectrum?*, *JCAP* **08** (2020) 007 [arXiv:2002.09904] [INSPIRE].

[59] H.A. Feldman, N. Kaiser and J.A. Peacock, *Power spectrum analysis of three-dimensional redshift surveys*, *Astrophys. J.* **426** (1994) 23 [astro-ph/9304022] [INSPIRE].

[60] K. Yamamoto et al., *A Measurement of the quadrupole power spectrum in the clustering of the 2dF QSO Survey*, *Publ. Astron. Soc. Jap.* **58** (2006) 93 [astro-ph/0505115] [INSPIRE].

[61] E. Castorina and M. White, *Beyond the plane-parallel approximation for redshift surveys*, *Mon. Not. Roy. Astron. Soc.* **476** (2018) 4403 [arXiv:1709.09730] [INSPIRE].

[62] F. Beutler and P. McDonald, *Unified galaxy power spectrum measurements from 6dFGS, BOSS, and eBOSS*, *JCAP* **11** (2021) 031 [arXiv:2106.06324] [INSPIRE].

[63] E. Chaussidon et al., *Constraining primordial non-Gaussianity with DESI 2024 LRG and QSO samples*, *JCAP* **06** (2025) 029 [arXiv:2411.17623] [INSPIRE].

[64] S.-F. Chen, P. Chakraborty and C. Dvorkin, *Analysis of BOSS galaxy data with weighted skew-spectra*, *JCAP* **05** (2024) 011 [arXiv:2401.13036] [INSPIRE].

[65] R. de Belsunce, J.M. Sullivan and P. McDonald, *Compressed 3D Lyman-α forest bispectrum*, *Phys. Rev. D* **113** (2026) 063550 [arXiv:2510.23597] [INSPIRE].

[66] N. Hand, Y. Li, Z. Slepian and U. Seljak, *An optimal FFT-based anisotropic power spectrum estimator*, *JCAP* **07** (2017) 002 [arXiv:1704.02357] [INSPIRE].

[67] N.A. Maksimova et al., *AbacusSummit: a massive set of high-accuracy, high-resolution N-body simulations*, *Mon. Not. Roy. Astron. Soc.* **508** (2021) 4017 [arXiv:2110.11398] [INSPIRE].

[68] L.H. Garrison et al., *The Abacus Cosmos: A Suite of Cosmological N-body Simulations*, *Astrophys. J. Suppl.* **236** (2018) 43 [arXiv:1712.05768] [INSPIRE].

[69] L.H. Garrison et al., *The abacus cosmological N-body code*, *Mon. Not. Roy. Astron. Soc.* **508** (2021) 575 [arXiv:2110.11392] [INSPIRE].

[70] Z. Zheng, A.L. Coil and I. Zehavi, *Galaxy Evolution from Halo Occupation Distribution Modeling of DEEP2 and SDSS Galaxy Clustering*, *Astrophys. J.* **667** (2007) 760 [astro-ph/0703457] [INSPIRE].

[71] S. Yuan et al., *AbacusHOD: a highly efficient extended multitracer HOD framework and its application to BOSS and eBOSS data*, *Mon. Not. Roy. Astron. Soc.* **510** (2022) 3301 [arXiv:2110.11412] [INSPIRE].

[72] S. Yuan et al., *The DESI one-per cent survey: exploring the halo occupation distribution of luminous red galaxies and quasi-stellar objects with AbacusSummit*, *Mon. Not. Roy. Astron. Soc.* **530** (2024) 947 [arXiv:2306.06314] [INSPIRE].

[73] DESI collaboration, *DESI 2024 IV: Baryon Acoustic Oscillations from the Lyman alpha forest*, *JCAP* **01** (2025) 124 [arXiv:2404.03001] [INSPIRE].

[74] 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** (2025) A253 [arXiv:2406.02504] [INSPIRE].

[75] J. Torrado and A. Lewis, *Cobaya: Code for Bayesian Analysis of hierarchical physical models*, *JCAP* **05** (2021) 057 [arXiv:2005.05290] [INSPIRE].

JCAP08(2026)003

[76] J. Torrado and A. Lewis, *Cobaya: Bayesian analysis in cosmology*, Astrophysics Source Code Library, record ascl:1910.019 (2019).

[77] J.A. Nelder and R. Mead, *A Simplex Method for Function Minimization*, *Comput. J.* **7** (1965) 308 [INSPIRE].

[78] DESI collaboration, *DESI 2024 II: sample definitions, characteristics, and two-point clustering statistics*, *JCAP* **07** (2025) 017 [arXiv:2411.12020] [INSPIRE].

[79] BOSS collaboration, *Velocity Bias from the Small Scale Clustering of SDSS-III BOSS Galaxies*, *Mon. Not. Roy. Astron. Soc.* **446** (2015) 578 [arXiv:1407.4811] [INSPIRE].

[80] C. Alcock and B. Paczynski, *An evolution free test for non-zero cosmological constant*, *Nature* **281** (1979) 358 [INSPIRE].

[81] N. Padmanabhan and M.J. White, *Constraining Anisotropic Baryon Oscillations*, *Phys. Rev. D* **77** (2008) 123540 [arXiv:0804.0799] [INSPIRE].

[82] H.E. Noriega, A. Aviles, S. Fromenteau and M. Vargas-Magaña, *Fast computation of non-linear power spectrum in cosmologies with massive neutrinos*, *JCAP* **11** (2022) 038 [arXiv:2208.02791] [INSPIRE].

[83] M. White and N. Padmanabhan, *Including parameter dependence in the data and covariance for cosmological inference*, *JCAP* **12** (2015) 058 [arXiv:1508.00566] [INSPIRE].

[84] A. Chudaykin, M.M. Ivanov and O.H.E. Philcox, *Reanalyzing DESI DR1. II. Constraints on dark energy, spatial curvature, and neutrino masses*, *Phys. Rev. D* **113** (2026) 123506 [arXiv:2511.20757] [INSPIRE].

[85] A.J.S. Hamilton, *Uncorrelated modes of the nonlinear power spectrum*, *Mon. Not. Roy. Astron. Soc.* **312** (2000) 257 [astro-ph/9905191] [INSPIRE].

[86] G.C. Wick, *The Evaluation of the Collision Matrix*, *Phys. Rev.* **80** (1950) 268 [INSPIRE].

JCAP08(2026)003