Near optimal bispectrum estimators for large-scale structure

Marcel Schmittfull, Tobias Baldauf, and Uroš Seljak

Phys. Rev. D **91**, 043530 — Published 26 February 2015

DOI: 10.1103/PhysRevD.91.043530

Near optimal bispectrum estimators for large-scale structure

Marcel Schmittfull,[1] Tobias Baldauf,[2] and Uroš Seljak[3]

[1] *Berkeley Center for Cosmological Physics, Department of Physics and Lawrence Berkeley National Laboratory, University of California, Berkeley, CA 94720, USA*

[2] *School of Natural Sciences, Institute for Advanced Study, Einstein Drive, Princeton, NJ 08540, USA*

[3] *Department of Physics, Department of Astronomy and Lawrence Berkeley National Laboratory, University of California, Berkeley, CA 94720, USA*

(Dated: February 2, 2015)

Clustering of large-scale structure provides significant cosmological information through the power spectrum of density perturbations. Additional information can be gained from higher-order statistics like the bispectrum, especially to break the degeneracy between the linear halo bias b_1 and the amplitude of fluctuations σ_8. We propose new simple, computationally inexpensive bispectrum statistics that are near optimal for the specific applications like bias determination. Corresponding to the Legendre decomposition of nonlinear halo bias and gravitational coupling at second order, these statistics are given by the cross-spectra of the density with three quadratic fields: the squared density, a tidal term, and a shift term. For halos and galaxies the first two have associated nonlinear bias terms b_2 and b_{s^2}, respectively, while the shift term has none in the absence of velocity bias (valid in the $k \to 0$ limit). Thus the linear bias b_1 is best determined by the shift cross-spectrum, while the squared density and tidal cross-spectra mostly tighten constraints on b_2 and b_{s^2} once b_1 is known. Since the form of the cross-spectra is derived from optimal maximum-likelihood estimation, they contain the full bispectrum information on bias parameters. Perturbative analytical predictions for their expectation values and covariances agree with simulations on large scales, $k \lesssim 0.09h/\mathrm{Mpc}$ at $z = 0.55$ with Gaussian $R = 20h^{-1}\mathrm{Mpc}$ smoothing, for matter-matter-matter, and matter-matter-halo combinations. For halo-halo-halo cross-spectra the model also needs to include corrections to the Poisson stochasticity.

I. INTRODUCTION

Large-scale structure (LSS) surveys constrain cosmology with ever higher precision (e.g. [1] and references therein). However, the unkown bias between invisible dark matter and observable tracers introduces degeneracies which weaken constraints on cosmological parameters. In particular, the normalization of the matter power spectrum σ_8 is degenerate with the unknown linear bias b_1 if only the 2-point correlation function or the power spectrum in Fourier space are measured, because they only depend on the combination $b_1^2\sigma_8^2$ (on large scales). This degeneracy is reduced if anisotropic redshift-space distortions are included. Further improvements can be obtained by probing statistics beyond the 2-point level such as the 3-point correlation function or the bispectrum in Fourier space, exploiting the fact that different combinations of cosmological and bias parameters are associated with different functional dependencies of the bispectrum on triangle shape. This has been demonstrated on galaxy survey data in [2–9]. Recently, Gil-Marin et al. [10, 11] measured the bispectrum of SDSS DR11 BOSS galaxies with unprecedented accuracy obtaining constraints on the growth rate f and the clustering amplitude σ_8 from galaxy clustering data alone. An additional important motivation to measure large-scale structure bispectra is to constrain primordial non-Gaussianity that can help to distinguish inflation models.

All analyses of real data to date have estimated the bispectrum or 3-point correlation function in a brute-force approach, going manually over triangle configurations and orientations. The number of these possible triangles is very large because it is cubic in the number of grid points per dimension (e.g. choosing ten k-bins in every direction gives $\mathcal{O}(10^3)$ triangles). This leads to several problems. First, the estimation of the bispectrum itself is computationally expensive, which is typically overcome by considering only a subset of possible triangles which comes however at the expense of losing a potentially significant fraction of the signal in the data. Second, the covariance between thousands or more triangles is hard to estimate from simulations because it would require a large number of independent realizations.[1] Third, the task of testing theoretical models of the bispectrum against simulations is rather complex because it is not obvious which triangle configurations are relevant for final parameter constraints, making it nontrivial to find a suitable metric for comparing models against simulations. Lastly, modeling anisotropic effects due to redshift-space distortions or survey geometry at the level of triangles can be cumbersome.

[1] Analytical covariances can only partially overcome this problem because it is hard to include real-world effects from e.g. the survey window function. A possible work-around is to neglect non-trivial bispectrum covariances in the estimation procedure and only estimate the covariance between final cosmological and bias parameters from simulations [10–12]. However, this leads to a sub-optimal estimator that is only unbiased in the limit of infinitely many observations.

To avoid some of these issues, the goal of this paper is to propose alternative simpler statistics that do not require direct bispectrum estimation for individual triangles but still include the full information of the bispectrum on bias and cosmological parameters in an optimal way. The basic idea to achieve this is to square the density and cross-correlate it with the density itself. This cross-spectrum then depends on the bispectrum.

In fact, the separability of dark matter (DM) and halo bispectra can be exploited to show that optimal bispectrum estimation is equivalent to cross-correlating three fields that are quadratic in the density with the density itself. Due to the particular form of DM and halo bispectra, these three quadratic fields are the squared density in configuration space $\delta^2(\mathbf{x})$, a shift term that contracts the density gradient with the displacement field, $\Psi^i(\mathbf{x})\partial_i\delta(\mathbf{x})$, and a tidal term, $s^2(\mathbf{x})$. The cross-spectra of these quadratic fields with the density encode the full bispectrum information but require only the computational cost of typical power spectrum analyses, which is much faster than brute-force bispectrum estimation. Additionally, since the cross-spectra only depend on a single scale k, covariances are much simpler to estimate from simulations than covariances between all triangle configurations.

The quadratic fields themselves can be computed efficiently as the product of fields in configuration space at the same location $\mathbf{x}$. The factors are given by the density itself or related quantities like e.g. the density gradient $\partial_i\delta$ or the displacement field $\Psi_i = -(\partial_i\partial^{-2})\delta$. The displacement field is already conventionally computed from real data when employing reconstruction to sharpen the BAO peak, so it should be straightforward to compute the other fields we propose in a similar fashion.

One of the three proposed statistics, the cross-spectrum between the squared density and the density, has been studied previously in the literature [13–15], but mostly in configuration space rather than Fourier space and in the large separation limit. Some of these studies speculated that this cross-spectrum contains information not present in the full three point function. The same statistic was also analysed in Fourier space in [16]. We will point out that the inverse variance weighted integral over this statistic in Fourier space is in fact an optimal estimator for the amplitude of the angle-independent part of the bispectrum, which is sensitive to a combination of b_1 and b_2, and we calculate this statistic consistently using one loop perturbation theory. Furthermore, we are extending these studies in considering a more realistic bias model and will show in detail that considering the other two cross-spectra involving other transformations than squaring are instrumental in constraining the bias parameters.

Our proposed procedure is also related to (but different from) a method recently proposed by Chiang et al. [17], which similarly aims at compressing bispectrum information to simpler observable quantities. Their work considers the correlation between locally measured power spectra with the local mean density, which is determined by the squeezed limit of the bispectrum. However, the squeezed limit of bispectra generated by nonlinear gravity is suppressed in absence of primordial non-Gaussianity. In contrast, we systematically derive the types of cross-spectra that explicitly probe bispectrum shapes that are optimal to estimate bias parameters, arriving at different statistics than the one proposed in [17].

To some extent our method can be regarded as a simplification of the separable expansion method to estimate general bispectra [18–20] by tailoring it to estimate bias parameters from large-scale structure bispectra, keeping the scale-dependence more obvious by not integrating over scales k. The method proposed here could be combined with separable expansion estimators to probe bispectrum contributions beyond those included in our modeling, but we leave this for future work.

In the context of the CMB, the fact that separable bispectra and trispectra can be estimated more efficiently than non-separable bispectra has been exploited for a long time. For example, the KSW estimator [21] is often used to estimate separable primordial bispectra directly (e.g. $f_{\rm NL}^{\rm local}$), primordial bispectra are often approximated by separable templates to facilitate their estimation, the ISW-lensing bispectrum can be measured by cross-correlating the lensing reconstruction field (which is quadratic in the CMB temperature) with the CMB temperature [22, 23], and primordial and non-primordial general bispectra can be estimated by expanding in separable basis functions [24–26]. Similarly, the separability of the lensing-induced trispectrum of the CMB temperature is exploited when computing the auto-power spectrum of the CMB lensing reconstruction field, which is a quadratic function in the CMB temperature with derivative operations dictated by the form of the CMB trispectrum [27–29]. However, the explicit separability of bispectra or trispectra induced by nonlinear gravity and bias relations has so far not been exploited in the field of large-scale structure. The aim of this paper is to provide a theoretical framework for this and test it with simulations.

We restrict ourselves to real space in this paper, but note that redshift space distortions (RSDs) should be included before applying the proposed technique to real data. Since RSDs change the expected bispectrum signal, in particular rendering it non-isotropic, the specific estimators we derive in real space should be modified and extended in order to be optimal in redshift space. While this will likely increase the number of cross-spectra, making the analysis somewhat more complicated, we do not anticipate any new conceptual challenges because the leading-order RSDs and certain Fingers-of-God models are still product-separable. Since venturing into redshift space is beyond the scope of this paper we leave it for future work.

The paper is organized as follows. We start with general definitions of quadratic fields and bispectra in Section II. Section III describes the relationship between optimal bispectrum estimation and cross-spectra of quadratic fields.

Theoretical predictions for the cross-spectra are computed in Section IV. Section V describes simulation results in comparison with theoretical predictions. An extension to primordial non-Gaussianity is briefly discussed in Section VI. Finally we conclude in Section VII. Two appendices provide technical details of large-scale limits and analytical covariances.

II. QUADRATIC FIELDS AND BISPECTRUM DECOMPOSITION

A. Quadratic fields

As will be shown in Section III, maximum-likelihood bispectrum estimators for bias parameters can be cast in form of cross-spectra of the density field with three fields that are quadratic in the configuration-space density, with different dependencies on the cosine μ between the Fourier space wavevectors $\mathbf{q}$ and $\mathbf{k}-\mathbf{q}$,

$$\mu \equiv \frac{\mathbf{q}\cdot(\mathbf{k}-\mathbf{q})}{q|\mathbf{k}-\mathbf{q}|}. \tag{1}$$

Explicitly, these three quadratic fields are:

- The *squared density* $\delta^2(\mathbf{x})$, which can be written as a convolution in Fourier space,

$$\delta^2(\mathbf{x}) = \int \frac{\mathrm{d}^3k}{(2\pi)^3} e^{i\mathbf{k}\mathbf{x}} \int \frac{\mathrm{d}^3q}{(2\pi)^3} \mathsf{P}_0(\mu)\delta(\mathbf{q})\delta(\mathbf{k}-\mathbf{q}), \tag{2}$$

  where $\mathsf{P}_0(\mu)=1$ is the Legendre polynomial for $l=0$.[2]

- The *shift-term*

$$-\Psi^i(\mathbf{x})\partial_i\delta(\mathbf{x}) = -\mathbf{\Psi}(\mathbf{x})\cdot\nabla\delta(\mathbf{x}) = -\int \frac{\mathrm{d}^3k}{(2\pi)^3} e^{i\mathbf{k}\mathbf{x}} \int \frac{\mathrm{d}^3q}{(2\pi)^3} F_2^1(q,|\mathbf{k}-\mathbf{q}|)\mathsf{P}_1(\mu)\delta(\mathbf{q})\delta(\mathbf{k}-\mathbf{q}), \tag{3}$$

  which depends on the $l=1$ Legendre polynomial $\mathsf{P}_1(\mu)=\mu$ and is obtained by contracting the density gradient $\nabla\delta$ with the displacement field

$$\mathbf{\Psi}(\mathbf{k}) = -\frac{i\mathbf{k}}{k^2}\delta(\mathbf{k}). \tag{4}$$

  The symmetric kernel F_2^1 in Eq. (3) is defined as

$$F_2^1(k_1,k_2) = \frac{1}{2}\left(\frac{k_1}{k_2}+\frac{k_2}{k_1}\right). \tag{5}$$

- The *tidal term*[3]

$$s^2(\mathbf{x}) \equiv \frac{3}{2}s_{ij}(\mathbf{x})s_{ij}(\mathbf{x}) = \int \frac{\mathrm{d}^3k}{(2\pi)^3} e^{i\mathbf{k}\mathbf{x}} \int \frac{\mathrm{d}^3q}{(2\pi)^3} \mathsf{P}_2(\mu)\delta(\mathbf{q})\delta(\mathbf{k}-\mathbf{q}), \tag{6}$$

  which is defined by contracting the tidal tensor

$$s_{ij}(\mathbf{k}) = \left(\frac{k_ik_j}{k^2} - \frac{1}{3}\delta_{ij}^{(\mathrm{K})}\right)\delta(\mathbf{k}) \tag{7}$$

  with itself. $\delta_{ij}^{(\mathrm{K})}$ denotes the Kronecker delta. The corresponding convolution kernel in Eq. (6) is given by the $l=2$ Legendre polynomial

$$\mathsf{P}_2(\mu) = \frac{3}{2}\left(\mu^2-\frac{1}{3}\right). \tag{8}$$

---

[2] $\mathsf{P}_l(\mu)$ with $l\in\{0,1,2\}$ always denotes Legendre polynomials in this paper and should not be confused with power spectra $P(k)$.
[3] An overall factor of 3/2 is absorbed compared to e.g. [30], i.e. $s^2_{\mathrm{here}}(\mathbf{x}) = \frac{3}{2}s^2_{\mathrm{there}}(\mathbf{x})$ and $\mathsf{P}_2^{\mathrm{here}}(\mu) = \frac{3}{2}S_2^{\mathrm{there}}(\mathbf{q},\mathbf{k}-\mathbf{q})$.

While these three quadratic fields will be derived more rigorously below, some intuition for their appearance can be gained as follows. In standard perturbation theory (see [31] for a review) the dark matter density is expanded in powers of the linear perturbation δ_0. In configuration space, truncating at second order, this can be written as (e.g. [30, 32, 33])

$$\delta_{\rm m}(\mathbf{x}) = \delta_0(\mathbf{x}) + \frac{17}{21}\delta_0^2(\mathbf{x}) + \mathbf{\Psi}_0(\mathbf{x})\cdot\nabla\delta_0(\mathbf{x}) + \frac{4}{21}s_0^2(\mathbf{x}). \tag{9}$$

The biased halo density can be modeled by [30, 34, 35]

$$\delta_{\rm h}(\mathbf{x}) = b_1\delta_{\rm m}(\mathbf{x}) + b_2\left[\delta_{\rm m}^2(\mathbf{x}) - \langle\delta_{\rm m}^2(\mathbf{x})\rangle\right] + \frac{2}{3}b_{s^2}\left[s_{\rm m}^2(\mathbf{x}) - \langle s_{\rm m}^2(\mathbf{x})\rangle\right], \tag{10}$$

where the normalization of the bias parameters b_1, b_2 and b_{s^2} is the same as in [30]. Since the linear perturbation δ_0 is assumed to be Gaussian, the leading order 3-point function is due to expectation values of the form $\langle\delta^{(2)}\delta_0\delta_0\rangle$, where $\delta^{(2)}$ can be any of the terms in Eqs. (9) and (10) that are quadratic in δ_0. There are only three types of such terms; the squared density (2), the shift term (3) and the tidal term (6). The corresponding k-dependencies that these fields imprint on the bispectra will later be used to reduce optimal bispectrum estimation to cross-spectra of these quadratic fields with the density.

To simplify calculations in the rest of the paper we define a general quadratic field $D[\delta](\mathbf{k})$ obtained from a density realization $\delta(\mathbf{k})$ and some kernel $D(\mathbf{q},\mathbf{k}-\mathbf{q})$, always assumed to be symmetric in its arguments, by[4]

$$D[\delta](\mathbf{k}) \equiv \int\frac{{\rm d}^3q}{(2\pi)^3}D(\mathbf{q},\mathbf{k}-\mathbf{q})\delta(\mathbf{q})\delta(\mathbf{k}-\mathbf{q}). \tag{11}$$

The squared density (2) corresponds to the identity kernel $D(\mathbf{q},\mathbf{k}-\mathbf{q}) = \mathsf{P}_0(\mu) = 1$, the shift term (3) to $D(\mathbf{q},\mathbf{k}-\mathbf{q}) = -F_2^1(q,|\mathbf{k}-\mathbf{q}|)\mathsf{P}_1(\mu)$ and the tidal term (6) to $D(\mathbf{q},\mathbf{k}-\mathbf{q}) = \mathsf{P}_2(\mu)$. We will usually suppress the arguments and just write $D \in \{\mathsf{P}_0, -F_2^1\mathsf{P}_1, \mathsf{P}_2\}$.

B. Cross-spectra

The cross-spectrum of a quadratic field $D[\delta]$ with the density is

$$\langle D[\delta](\mathbf{k})\delta(\mathbf{k}')\rangle \equiv (2\pi)^3\delta_D(\mathbf{k}+\mathbf{k}')P_{D[\delta],\delta}(k), \tag{12}$$

where δ_D always denotes the Dirac delta function. Given a realization of the density, an unbiased estimator for the cross-spectrum is given by

$$\hat{P}_{D[\delta],\delta}(k) = \frac{1}{4\pi L^3}\int{\rm d}\Omega_{\hat{\mathbf{k}}}D[\delta](\mathbf{k})\delta(-\mathbf{k}), \tag{13}$$

where L is the box size. The normalization is fixed by noting that $(2\pi)^3\delta_D(\mathbf{0}) = L^3$ and requiring

$$\langle\hat{P}_{D[\delta],\delta}(k)\rangle = P_{D[\delta],\delta}(k). \tag{14}$$

On a discrete grid the angular integral is replaced by a sum over discrete $\mathbf{k}$ vectors with wavenumber $|\mathbf{k}|$ belonging to the bin centered at k,

$$\frac{1}{4\pi}\int{\rm d}\Omega_{\hat{\mathbf{k}}} \quad\to\quad \frac{1}{N_{\rm modes}(k)}\sum_{\mathbf{k},[k-\Delta k/2\le|\mathbf{k}|\le k+\Delta k/2]}, \tag{15}$$

where $N_{\rm modes}(k) = 4\pi(k/\Delta k)^2$ at high k if the binning width is Δk. In practice, we count the number of modes manually in the code because this is more accurate at low k.

---

[4] $D[\delta]$ with square brackets denotes the functional that turns some density realization $\delta(\mathbf{k})$ into the r.h.s. of Eq. (11). $D(\mathbf{q},\mathbf{k}-\mathbf{q})$ denotes the corresponding kernel.

C. Legendre decomposition of gravitational bispectra

At leading order in perturbation theory the DM bispectrum is given by[5]

$$B_{\rm mmm}(k_1,k_2,k_3) = 2P^{\rm lin}_{\rm mm}(k_1)P^{\rm lin}_{\rm mm}(k_2)F_2(\mathbf{k}_1,\mathbf{k}_2) + \text{ 2 perms in } k_1,k_2,k_3, \tag{16}$$

where $P^{\rm lin}_{\rm mm}$ is the linear DM power spectrum and F_2 denotes the symmetrized kernel for the second order density perturbation,

$$F_2(\mathbf{k}_1,\mathbf{k}_2) = \frac{17}{21} + \frac{1}{2}\left(\frac{k_1}{k_2}+\frac{k_2}{k_1}\right)\hat{\mathbf{k}}_1\cdot\hat{\mathbf{k}}_2 + \frac{4}{21}\frac{3}{2}\left((\hat{\mathbf{k}}_1\cdot\hat{\mathbf{k}}_2)^2 - \frac{1}{3}\right), \tag{17}$$

where $\hat{\mathbf{k}}_i = \mathbf{k}_i/k_i$. The decomposition of this kernel in Legendre polynomials in the cosine $\hat{\mathbf{k}}_1\cdot\hat{\mathbf{k}}_2$ is

$$F_2(\mathbf{k}_1,\mathbf{k}_2) = \sum_{l=0}^{2} F_2^l(k_1,k_2)\mathsf{P}_l(\hat{\mathbf{k}}_1\cdot\hat{\mathbf{k}}_2) \tag{18}$$

with coefficients F_2^l given by

$$F_2^0(k_1,k_2) = \frac{17}{21}, \tag{19}$$

$$F_2^1(k_1,k_2) = \frac{1}{2}\left(\frac{k_1}{k_2}+\frac{k_2}{k_1}\right), \tag{20}$$

$$F_2^2(k_1,k_2) = \frac{4}{21}. \tag{21}$$

The bispectrum (16) can thus be split in parts that depend on different Legendre polynomials in the angle $\hat{\mathbf{k}}_1\cdot\hat{\mathbf{k}}_2$ (and permutations),

$$B_{\rm mmm}(k_1,k_2,k_3) = \sum_{l=0}^{2} B^{(l)}_{\rm mmm}(k_1,k_2,k_3), \tag{22}$$

where

$$B^{(l)}_{\rm mmm}(k_1,k_2,k_3) = 2P^{\rm lin}_{\rm mm}(k_1)P^{\rm lin}_{\rm mm}(k_2)F_2^l(k_1,k_2)\mathsf{P}_l(\hat{\mathbf{k}}_1\cdot\hat{\mathbf{k}}_2) + 2 \text{ perms}. \tag{23}$$

The leading-order unsymmetrized matter-matter-halo bispectrum can be obtained from the bias relation (10),

$$B^{\rm unsym}_{\rm mmh}(k_1,k_2,k_3) = b_1 B_{\rm mmm}(k_1,k_2,k_3) + 2b_2 P^{\rm lin}_{\rm mm}(k_1)P^{\rm lin}_{\rm mm}(k_2) + \frac{4}{3}b_{s^2}P^{\rm lin}_{\rm mm}(k_1)P^{\rm lin}_{\rm mm}(k_2)\mathsf{P}_2(\hat{\mathbf{k}}_1\cdot\hat{\mathbf{k}}_2), \tag{24}$$

where k_3 is for the halo density. This can similarly be decomposed into $l=0,1,2$ Legendre polynomials:

$$B^{\rm unsym,(0)}_{\rm mmh}(k_1,k_2,k_3) = b_1 B^{(0)}_{\rm mmm}(k_1,k_2,k_3) + 2b_2 P^{\rm lin}_{\rm mm}(k_1)P^{\rm lin}_{\rm mm}(k_2), \tag{25}$$

$$B^{\rm unsym,(1)}_{\rm mmh}(k_1,k_2,k_3) = b_1 B^{(1)}_{\rm mmm}(k_1,k_2,k_3), \tag{26}$$

$$B^{\rm unsym,(2)}_{\rm mmh}(k_1,k_2,k_3) = b_1 B^{(2)}_{\rm mmm}(k_1,k_2,k_3) + \frac{4}{3}b_{s^2}P^{\rm lin}_{\rm mm}(k_1)P^{\rm lin}_{\rm mm}(k_2)\mathsf{P}_2(\hat{\mathbf{k}}_1\cdot\hat{\mathbf{k}}_2). \tag{27}$$

Similarly, the halo-halo-halo bispectrum

$$B_{\rm hhh}(k_1,k_2,k_3) = 2P^{\rm lin}_{\rm mm}(k_1)P^{\rm lin}_{\rm mm}(k_2)\left[b_1^3 F_2(\mathbf{k}_1,\mathbf{k}_2) + b_1^2 b_2 + \frac{2}{3}b_1^2 b_{s^2}\mathsf{P}_2(\hat{\mathbf{k}}_1\cdot\hat{\mathbf{k}}_2)\right] + 2 \text{ perms} \tag{28}$$

---

[5] The $\mathbf{k}_i$ must be such that they form a closed triangle, $\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3=0$. Up to the overall orientation, the triangle can be specified e.g. by three sidelengths k_1, k_2 and k_3, which fixes the angle between any two sides, e.g. $\mathbf{k}_1\cdot\mathbf{k}_2 = (k_3^2-k_1^2-k_2^2)/2$. We will parametrize triangles by sidelengths and/or angles, using whichever is most convenient in the context.

can be decomposed into

$$B^{(0)}_{\rm hhh}(k_1,k_2,k_3) = b_1^3 B^{(0)}_{\rm mmm}(k_1,k_2,k_3) + 2b_1^2 b_2 \left[P^{\rm lin}_{\rm mm}(k_1)P^{\rm lin}_{\rm mm}(k_2) + 2 \text{ perms}\right], \tag{29}$$

$$B^{(1)}_{\rm hhh}(k_1,k_2,k_3) = b_1^3 B^{(1)}_{\rm mmm}(k_1,k_2,k_3), \tag{30}$$

$$B^{(2)}_{\rm hhh}(k_1,k_2,k_3) = b_1^3 B^{(2)}_{\rm mmm}(k_1,k_2,k_3) + \frac{4}{3} b_1^2 b_{s^2} \left[P^{\rm lin}_{\rm mm}(k_1)P^{\rm lin}_{\rm mm}(k_2)\mathsf{P}_2(\hat{\mathbf{k}}_1\cdot\hat{\mathbf{k}}_2) + 2 \text{ perms}\right]. \tag{31}$$

Since the $l = 1$ part depends only on b_1 and is not contaminated by the nonlinear bias parameters b_2 and b_{s^2} we expect it to be most powerful for determining b_1. The $l = 0$ contribution depends on a mixture of b_1 and b_2, so one could constrain b_2 if b_1 was known. Similarly, the $l = 2$ part depends on b_1 and b_{s^2}, so b_{s^2} can be constrained once b_1 is known.

Note that all bispectrum contributions are product-separable and only depend on $l = 0, 1, 2$ Legendre polynomials. In the following section we will discuss how this can be used to simplify estimators for the amplitude of bispectrum contributions, which can in turn be used to estimate bias parameters.

III. MAXIMUM LIKELIHOOD BISPECTRUM ESTIMATION

A. General bispectra

Assuming a fiducial theoretical power spectrum $P_\delta(k)$ and bispectrum $f_{\rm NL}^{B_{\rm theo}} B_\delta^{\rm theo}$ for the density perturbation δ, the maximum likelihood estimator (in the limit of weak non-Gaussianity) for the amplitude[6] $f_{\rm NL}^{B_{\rm theo}}$ of the bispectrum is given by [18, 19]

$$\hat{f}_{\rm NL}^{B^{\rm theo}} = \frac{(2\pi)^3}{N_{\rm theo}} \int \frac{{\rm d}^3k}{(2\pi)^3} \int \frac{{\rm d}^3q}{(2\pi)^3} \frac{B_\delta^{\rm theo}(\mathbf{q},\mathbf{k}-\mathbf{q},-\mathbf{k})[\delta(\mathbf{q})\delta(\mathbf{k}-\mathbf{q})\delta(-\mathbf{k}) - 3\langle\delta(\mathbf{q})\delta(\mathbf{k}-\mathbf{q})\rangle\delta(-\mathbf{k})]}{P_\delta(q)P_\delta(|\mathbf{k}-\mathbf{q}|)P_\delta(k)} \tag{32}$$

where P_δ factors in the denominator represent inverse-variance weighting of the observed density perturbation δ, and $N_{\rm theo}$ is a normalization factor depending on the theoretical bispectrum whose amplitude we aim to measure (see [20] for the explicit definition). The term linear in the observed density will be omitted in the following because it is only relevant if the field is statistically inhomogeneous (e.g. due to inhomogenous noise).

The estimator of Eq. (32) is unbiased and includes information from the full bispectrum, i.e. all triangle configurations and orientations. It is optimal in the limit of weak non-Gaussianity since it is derived by Edgeworth-expanding the likelihood around a Gaussian pdf truncating higher orders of $f_{\rm NL}$ and connected n-point functions beyond the bispectrum. The inverse-variance weighting of the density in Eq. (32) is optimal if higher-order contributions to the bispectrum covariance are neglected. The estimator can be improved by relaxing these assumptions, but we leave this to future work.[7]

B. Estimating separable bispectra by cross-correlating linear and quadratic fields

The bispectra discussed above are given by sums of terms that have a product-separable form.[8] For simplicity, let us start with a bispectrum given by just one such product-separable term

$$B_\delta^{\rm theo}(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3) = f(\mathbf{k}_1)g(\mathbf{k}_2)h(\mathbf{k}_3) \tag{33}$$

for some functions f, g and h. Then the estimator (32) becomes

$$\hat{f}_{\rm NL}^{B^{\rm theo}} = \frac{(2\pi)^3}{N_{\rm theo}} \int \frac{{\rm d}^3k}{(2\pi)^3} \left[\int \frac{{\rm d}^3q}{(2\pi)^3} \frac{f(\mathbf{q})\delta(\mathbf{q})}{P_\delta(q)} \frac{g(\mathbf{k}-\mathbf{q})\delta(\mathbf{k}-\mathbf{q})}{P_\delta(|\mathbf{k}-\mathbf{q}|)}\right] \frac{h(-\mathbf{k})\delta(-\mathbf{k})}{P_\delta(k)}. \tag{34}$$

The integral over $\mathbf{q}$ is a convolution of two filtered densities $f\delta/P_\delta$ and $g\delta/P_\delta$, which we denote as

$$\left[\frac{f\delta}{P_\delta} * \frac{g\delta}{P_\delta}\right](\mathbf{k}) \equiv \int \frac{{\rm d}^3q}{(2\pi)^3} \frac{f(\mathbf{q})\delta(\mathbf{q})}{P_\delta(q)} \frac{g(\mathbf{k}-\mathbf{q})\delta(\mathbf{k}-\mathbf{q})}{P_\delta(|\mathbf{k}-\mathbf{q}|)}. \tag{35}$$

---

[6] In our notation $f_{\rm NL}^B$ denotes the nonlinearity amplitude of an arbitrary bispectrum B, no matter if it is generated by primordial non-Gaussianity or nonlinear gravity.

[7] Alternatively, constraints can be tightened by pushing the analysis to smaller scales, which could be achieved by improving the theory modeling of the bispectrum (see e.g. [12, 36–39]), or by applying clipping or logarithmic density transforms [40, 41].

[8] At leading order this is the case because the second order density (9) and the halo density (10) depend on products of fields evaluated at the same location $\mathbf{x}$ which can be written as convolutions with separable kernels in Fourier space.

Then, the estimator (34) simplifies to

$$\begin{aligned}\hat{f}_{\rm NL}^{B^{\rm theo}} &= \frac{1}{N_{\rm theo}}\int \mathrm{d}k\frac{k^2}{P_\delta(k)}\int \mathrm{d}\Omega_{\hat{\mathbf{k}}}\left[\frac{f\delta}{P_\delta}*\frac{g\delta}{P_\delta}\right](\mathbf{k})\,[h\delta](-\mathbf{k})\\ &= \frac{4\pi L^3}{N_{\rm theo}}\int \mathrm{d}k\frac{k^2}{P_\delta(k)}\hat{P}_{\frac{f\delta}{P_\delta}*\frac{g\delta}{P_\delta},\,h\delta}(k),\end{aligned} \tag{36}$$

where we used Eq. (13). This is a (weighted) integral over the estimated cross-spectrum between the quadratic field (35) and the filtered density

$$[h\delta](\mathbf{k}) \equiv h(\mathbf{k})\delta(\mathbf{k}). \tag{37}$$

Rather than performing the integration of the cross-spectrum over wavenumbers k in Eq. (36), it is useful to consider the cross-spectrum as the fundamental observable. This can simplify comparisons of theory, simulations and observations as a function of scale k, and offers the possibility to incorporate covariances between cross-spectra.

The convolution in Eq. (35) can be calculated efficiently by filtering the density in k-space with two filters f/P and g/P, Fourier transforming to configuration space, multiplying the fields with each other in configuration space, and Fourier transforming back to k-space (e.g. for $f = g = P_\delta$, Eq. (35) is the Fourier transform of $\delta^2(\mathbf{x})$).

If the theoretical bispectrum consists of a sum of separable terms, each term gives rise to a cross-spectrum of accordingly filtered densities. Each cross-spectrum can be measured to analyse individual contributions to the bispectrum, or they can be combined to constrain the overall bispectrum amplitude.

C. Gravitational bispectrum and bias estimators

Eqs. (23), (25)-(27) and (29)-(31) show that matter-matter-matter, matter-matter-halo and halo-halo-halo bispectra can be constructed from contributions of the form[9]

$$B^{(l)}_{\rm unsym}(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3) \equiv 2P^{\rm lin}_{\rm mm}(k_1)P^{\rm lin}_{\rm mm}(k_2)F_2^l(k_1,k_2)\mathsf{P}_l(\hat{\mathbf{k}}_1\cdot\hat{\mathbf{k}}_2), \tag{38}$$

which are not symmetric in the k_i. Note that every $F_2^l\mathsf{P}_l$ kernel is separable (or a sum of separable terms), because squaring the triangle condition $\mathbf{k}_1+\mathbf{k}_2=-\mathbf{k}_3$ implies $\mathbf{k}_1\cdot\mathbf{k}_2 = \frac{1}{2}(k_3^2-k_1^2-k_2^2)$. Consequently, the maximum likelihood estimator for the amplitude of a bispectrum contribution (38) can be expressed as an integral over a cross-spectrum between a quadratic field and the density. Explicitly, the maximum likelihood estimator (32) for the amplitude of a bispectrum (38) involving the densities δ_a, δ_a and δ_b (for $a,b\in\{\mathrm{m},\mathrm{h}\}$ labeling DM or halo densities) gives[10]

$$\hat{f}_{\rm NL}^{B^{(l)}_{\rm unsym}} = \frac{8\pi L^3}{N_{B^{(l)}_{\rm unsym}}}\int \mathrm{d}k\frac{k^2}{P_{bb}(k)}\hat{P}_{\tilde{F}_2^l\mathsf{P}_l[\delta_a],\,\delta_b}(k) \tag{39}$$

where we defined

$$\tilde{F}_2^l\mathsf{P}_l[\delta_a](\mathbf{k}) \equiv \int\frac{\mathrm{d}^3q}{(2\pi)^3}\frac{P^{\rm lin}_{\rm mm}(q)P^{\rm lin}_{\rm mm}(|\mathbf{k}-\mathbf{q}|)}{P_{aa}(q)P_{aa}(|\mathbf{k}-\mathbf{q}|)}F_2^l(q,|\mathbf{k}-\mathbf{q}|)\mathsf{P}_l(\mu)\delta_a(\mathbf{q})\delta_a(\mathbf{k}-\mathbf{q}), \tag{40}$$

where μ is the cosine (1) between $\mathbf{q}$ and $\mathbf{k}-\mathbf{q}$. The power spectrum ratio in the integrand serves as a weight which becomes unity on large scales for dark matter. To gain intuition, we choose a slightly less optimal weight by setting the power spectrum ratio to unity on all scales[11] (dropping the tilde on F_2),

$$F_2^l\mathsf{P}_l[\delta_a](\mathbf{k}) \equiv \int\frac{\mathrm{d}^3q}{(2\pi)^3}F_2^l(q,|\mathbf{k}-\mathbf{q}|)\mathsf{P}_l(\mu)\delta_a(\mathbf{q})\delta_a(\mathbf{k}-\mathbf{q}). \tag{41}$$

---

[9] In all equations l is fixed to $l=0,1,2$, unless we explicitly write $\sum_l$ to sum over l. We never use Einstein summation notation for l.

[10] Note that the estimators for the $l=0,1,2$ contributions to the full symmetric gravitational bispectrum $B^{(l)}$ contain an additional factor of 3 because Eq. (38) picked one out of three permutations in the k_i.

[11] If δ_a is the halo density this implies a sub-optimal weighting at high k, but results are not biased because we treat simulations and theory consistently. Also, high k modes are suppressed by smoothing, which we apply before squaring any fields to suppress nonlinear mode coupling, similarly to cutting off bispectrum analyses at some maximum wavenumber $k_{\rm max}$. We do this because we do not have a reliable way to model these high k modes. In our simulations, the power ratio $P^{\rm lin}_{\rm mm}/P_{\rm hh}$ drops by 30% or less between $k\to 0$ and $k = 0.3h/\mathrm{Mpc}$. All smoothing kernels we use asymptote to 0 much faster for increasing k, so that corrections due to the power spectrum ratio weight are likely small.

In configuration space, the quadratic fields (41) are proportional to

$$\mathsf{P}_0[\delta](\mathbf{x}) = \delta^2(\mathbf{x}) \tag{42}$$
$$-F_2^1\mathsf{P}_1[\delta](\mathbf{x}) = -\Psi^i(\mathbf{x})\partial_i\delta(\mathbf{x}) \tag{43}$$
$$\mathsf{P}_2[\delta](\mathbf{x}) = s^2(\mathbf{x}). \tag{44}$$

According to Eq. (39) the (integrated) cross-spectra of these three quadratic fields with the density are optimal estimators for the amplitude of the $l = 0, 1, 2$ contributions to the gravitational bispectrum,

$$\hat{f}_{\rm NL}^{B^{(0)}_{\rm unsym}} = \frac{17}{21}\frac{8\pi L^3}{N_{B^{(0)}_{\rm unsym}}}\int \mathrm{d}k\,\frac{k^2}{P_{bb}(k)}\hat{P}_{\delta_a^2,\,\delta_b}(k) \tag{45}$$

$$\hat{f}_{\rm NL}^{B^{(1)}_{\rm unsym}} = -\frac{8\pi L^3}{N_{B^{(1)}_{\rm unsym}}}\int \mathrm{d}k\,\frac{k^2}{P_{bb}(k)}\hat{P}_{-\Psi_a^i\partial_i\delta_a,\,\delta_b}(k) \tag{46}$$

$$\hat{f}_{\rm NL}^{B^{(2)}_{\rm unsym}} = \frac{4}{21}\frac{8\pi L^3}{N_{B^{(2)}_{\rm unsym}}}\int \mathrm{d}k\,\frac{k^2}{P_{bb}(k)}\hat{P}_{s_a^2,\,\delta_b}(k). \tag{47}$$

These nonlinearity amplitudes $\hat{f}_{\rm NL}$ depend on bias parameters, e.g. on b_1^3, $b_1^2b_2$ and $b_1^2b_{s^2}$ for halo-halo-halo cross-spectra. Bias parameters can therefore be estimated by combining the measured nonlinearity amplitudes appropriately. Similarly, the bias parameters can be obtained by jointly fitting them to the three measured cross-spectra $\hat{P}_{D[\delta_a]\delta_b}(k)$ for $D \in \{\mathsf{P}_0, -F_2^1\mathsf{P}_1, \mathsf{P}_2\}$.

These cross-spectra contain the entire information that a full bispectrum analysis would yield, if we are only interested in the amplitudes of fixed shape contributions to the bispectrum, which is often the case, e.g. for estimating bias parameters.

As will be discussed later, the $k^2/P_{bb}(k)$ weighting of the cross-spectra corresponds to inverse-variance weighting in the limit of $k \to 0$ assuming Gaussian bispectrum covariance. This weighting is improved if cross-spectrum variances and covariances obtained from N-body simulations or mock catalogues are used to fit cross-spectrum models to measurements. Estimating covariances from simulations is computationally much less expensive for cross-spectra than for bispectra because the three cross-spectra depend only on a single argument k whereas bispectra depend on triangle configurations (k_1, k_2, k_3), of which there can be many thousands, especially if fine k-binning is used to distinguish different bispectrum shapes.

D. Configuration space estimators

We mainly worked in Fourier space so far which has the advantage that different modes are uncorrelated at leading order. Sometimes it is more convenient to work in configuration space because it may be easier to include effects that are localized in configuration space (e.g. the survey selection function). It turns out that the optimal bispectrum estimator (36) for the amplitude of a generic separable bispectrum (33) can be rewritten in configuration space instead of Fourier space as[12]

$$\hat{f}_{\rm NL}^{B^{\rm theo}} = \frac{(2\pi)^3}{N_{\rm theo}}\int \mathrm{d}^3x\,\frac{f\delta}{P_\delta}(\mathbf{x})\frac{g\delta}{P_\delta}(\mathbf{x})\frac{h\delta}{P_\delta}(\mathbf{x}), \tag{48}$$

where the filtered densities are defined as

$$\frac{f\delta}{P_\delta}(\mathbf{x}) = \int\frac{\mathrm{d}^3k}{(2\pi)^3}e^{i\mathbf{k}\mathbf{x}}\frac{f(\mathbf{k})\delta(\mathbf{k})}{P_\delta(k)}, \tag{49}$$

and similarly if f is replaced by g and h. If one wants to avoid Fourier space entirely, the filtering can be performed with convolutions in configuration space.

---

[12] To see this, express the quadratic field in (36) as $\int \mathrm{d}^3x e^{-i\mathbf{k}\mathbf{x}}\frac{f\delta}{P}(\mathbf{x})\frac{g\delta}{P}(\mathbf{x})$ and the linear field as $\int \mathrm{d}^3y e^{i\mathbf{k}\mathbf{y}}\frac{h\delta}{P}(\mathbf{y})$. Then, integrating over $\mathbf{k}$ gives $\mathbf{y} = \mathbf{x}$ and the result (48) follows. Alternatively, Eq. (48) can be derived by introducing $\int \mathrm{d}^3q'\delta_D(\mathbf{q}' - \mathbf{k} + \mathbf{q})$ in Eq. (34) and using $(2\pi)^3\delta_D(\mathbf{k}) = \int \mathrm{d}^3x e^{i\mathbf{k}\mathbf{x}}$.

For the gravitational bispectrum contributions $B^{(l)}_{\rm unsym}$ we get[13]

$$\hat{f}_{\rm NL}^{B^{(0)}_{\rm unsym}} = \frac{34}{21}\frac{(2\pi)^3}{N_{B^{(0)}_{\rm unsym}}}\int {\rm d}^3x\, \delta_a^2(\mathbf{x})\frac{\delta_b}{P_{bb}}(\mathbf{x}), \tag{50}$$

$$\hat{f}_{\rm NL}^{B^{(1)}_{\rm unsym}} = 2\frac{(2\pi)^3}{N_{B^{(1)}_{\rm unsym}}}\int {\rm d}^3x\, \Psi_a^i(\mathbf{x})[\partial_i\delta_a(\mathbf{x})]\frac{\delta_b}{P_{bb}}(\mathbf{x}), \tag{51}$$

$$\hat{f}_{\rm NL}^{B^{(2)}_{\rm unsym}} = \frac{8}{21}\frac{(2\pi)^3}{N_{B^{(2)}_{\rm unsym}}}\int {\rm d}^3x\, s_a^2(\mathbf{x})\frac{\delta_b}{P_{bb}}(\mathbf{x}). \tag{52}$$

IV. THEORY CROSS-SPECTRA

A. Smoothing

To suppress small-scale modes, we apply a smoothing filter $W_R(k)$ to the nonlinear (DM or halo) density,

$$\delta^R(\mathbf{k}) \equiv W_R(k)\delta(\mathbf{k}), \tag{53}$$

where R is the smoothing radius. For Gaussian smoothing,

$$W_R^{\rm Gauss}(k) = e^{-\frac{1}{2}k^2R^2}. \tag{54}$$

The power spectrum and bispectrum of the smoothed field δ^R are

$$P_{\delta^R\delta^R}(k) = P^R_{\delta\delta}(k) = W_R^2(k)P_{\delta\delta}(k), \tag{55}$$

$$B_{\delta^R\delta^R\delta^R}(k_1,k_2,k_3) = W_R(k_1)W_R(k_2)W_R(k_3)B_{\delta\delta\delta}(k_1,k_2,k_3). \tag{56}$$

More small-scale modes can be included by choosing smaller smoothing scale R. This will generally increase the signal to noise in the observables at the expense of worse agreement with models, so in practice some trade-off smoothing scale R should be chosen. Note that the window function is applied to every density field, so an isotropic survey window function could be included in the model predictions below by simply modifying the smoothing kernel W_R appropriately.

Haloes are finite size objects and thus the correlators of halo centers should not have structure on scales below the halo scale. This is equivalent to modelling the halo density field in terms of the density field smoothed on the halo scale. There is indeed evidence for a cutoff of the protohalo power spectrum in Lagrangian space [42]. Gravitational evolution, in particular the collapse, shrinks this scale and generates non-linear contributions to the clustering statistics. While a self consistent theoretical understanding of the combined effects of Lagrangian smoothing and non-linear gravitational evolution does not exist, there are hints from simulation for the existence of a finite smoothing scale in the halo density field in Eulerian space [16]. Such a smoothing scale would also affect our perturbation theory calculations of the correlators of the squared field. In particular, the integral over the smoothed power spectrum would decrease the amplitude of the low-k limit of the I^R_{DE} terms (Eq. (62) below) by a few percent for realistic smoothing scales and thus increase the inferred bias parameters by this amount. At higher wavenumbers, the smoothing scale enters more explicitly, also in the $I^{{\rm bare},R}_{DE}$ (Eq. (63) below) and can thus lead to larger changes. We have calculated these effects and found that the additional fitting parameter did not improve our fits of halo-halo-halo statistics. With no apparent improvements and the lack of a theoretical model, we decide to neglect the smoothing corrections in this study but remark that they should be understood and included in the future.

B. General bispectrum

To compute general theoretical predictions for cross-spectra we work with two smoothed fields δ^R_a and δ^R_b that can be dark matter or halo densities, $a, b \in \{{\rm m}, {\rm h}\}$. The expectation value of the cross-spectrum of a quadratic field $D[\delta^R_a]$

[13] Again, we assume $P_\delta = P^{\rm lin}_{\rm mm}$ to get a simple weighting, but it would be straightforward to include the optimal $P^{\rm lin}_{\rm mm}/P_\delta$ weighting.

for $D \in \{\mathsf{P}_0, -F_2^1\mathsf{P}_1, \mathsf{P}_2\}$ with the density δ_b^R is given by an integral over the bispectrum B,

$$\langle D[\delta_a^R](\mathbf{k})\delta_b^R(\mathbf{k}')\rangle = \int \frac{\mathrm{d}^3q}{(2\pi)^3} D(\mathbf{q},\mathbf{k}-\mathbf{q})\langle\delta_a^R(\mathbf{q})\delta_a^R(\mathbf{k}-\mathbf{q})\delta_b^R(\mathbf{k}')\rangle \tag{57}$$

$$= (2\pi)^3\delta_D(\mathbf{k}+\mathbf{k}')\int \frac{\mathrm{d}^3q}{(2\pi)^3} D(\mathbf{q},\mathbf{k}-\mathbf{q})B_{\delta_a^R\delta_a^R\delta_b^R}(\mathbf{q},\mathbf{k}-\mathbf{q},-\mathbf{k}). \tag{58}$$

For $a \neq b$, the bispectrum B is not symmetric in its arguments and the last argument $-\mathbf{k}$ is associated with δ_b. Writing the smoothing kernels explicitly, we have

$$P_{D[\delta_a^R],\delta_b^R}(k) = W_R(k)\int \frac{\mathrm{d}^3q}{(2\pi)^3} W_R(q)W_R(|\mathbf{k}-\mathbf{q}|)D(\mathbf{q},\mathbf{k}-\mathbf{q})B_{\delta_a\delta_a\delta_b}(\mathbf{q},\mathbf{k}-\mathbf{q},-\mathbf{k}). \tag{59}$$

C. Matter-matter-matter cross-spectra

For cross-spectra of smoothed dark matter fields ($a = b = \mathrm{m}$), the DM bispectrum (16) gives

$$P_{D[\delta_\mathrm{m}^R],\delta_\mathrm{m}^R}(k) = \int \frac{\mathrm{d}^3q}{(2\pi)^3} D(\mathbf{q},\mathbf{k}-\mathbf{q})B_{\delta_\mathrm{m}^\mathrm{R}\delta_\mathrm{m}^\mathrm{R}\delta_\mathrm{m}^\mathrm{R}}(\mathbf{q},\mathbf{k}-\mathbf{q},-\mathbf{k}) \tag{60}$$

$$= 2I_{DF_2}^R(k) + 4I_{DF_2}^{\mathrm{bare},R}(k), \tag{61}$$

where we used that the kernel $D(\mathbf{q},\mathbf{k}-\mathbf{q})$ is assumed to be symmetric in its arguments and we defined for kernels D and E

$$I_{DE}^R(k) \equiv W_R(k)\int \frac{\mathrm{d}^3q}{(2\pi)^3} W_R(q)W_R(|\mathbf{k}-\mathbf{q}|)P_\mathrm{mm}^\mathrm{lin}(q)P_\mathrm{mm}^\mathrm{lin}(|\mathbf{k}-\mathbf{q}|)D(\mathbf{q},\mathbf{k}-\mathbf{q})E(\mathbf{q},\mathbf{k}-\mathbf{q}), \tag{62}$$

which is symmetric under $D \leftrightarrow E$, and

$$I_{DE}^{\mathrm{bare},R}(k) \equiv W_R(k)P_\mathrm{mm}^\mathrm{lin}(k)\int \frac{\mathrm{d}^3q}{(2\pi)^3} W_R(q)W_R(|\mathbf{k}-\mathbf{q}|)P_\mathrm{mm}^\mathrm{lin}(q)D(\mathbf{q},\mathbf{k}-\mathbf{q})E(\mathbf{q},-\mathbf{k}), \tag{63}$$

which is not symmetric under $D \leftrightarrow E$. Explicit predictions for the cross-spectra $P_{\delta_\mathrm{m}^2,\delta_\mathrm{m}}$, $P_{-\Psi_\mathrm{m}^i\partial_i\delta_\mathrm{m},\delta_\mathrm{m}}$ and $P_{s_\mathrm{m}^2,\delta_\mathrm{m}}$ can be obtained from Eq. (61) by plugging in $D = \mathsf{P}_0$, $D = -F_2^1\mathsf{P}_1$ and $D = \mathsf{P}_2$, respectively, and setting $E = F_2$. Note that there are three powers of the smoothing kernel because we smooth the nonlinear rather than the linear field.

The integrals in Eqs. (62) and (63) are similar to typical 1-loop expressions and can be reduced to two-dimensional integrals over scale q and cosine $\hat{\mathbf{q}}\cdot\hat{\mathbf{k}}$, which can be evaluated numerically with little computational cost (see e.g. [43, 44] for public codes that compute similar integrals). The factors $W_R(|\mathbf{k}-\mathbf{q}|)$ and $P_\mathrm{mm}^\mathrm{lin}(|\mathbf{k}-\mathbf{q}|)$ introduce a non-trivial angle dependence so that the angular integration generally needs to be performed numerically.

The only ingredient for the theory prediction of Eq. (61) is the model for the DM bispectrum. Improved bispectrum models that have the same form as Eq. (16) could easily be included, e.g. by replacing the perturbation theory F_2 kernel by an effective F_2 kernel fitted to N-body simulations [12, 36, 37].

D. Matter-matter-halo cross-spectra

From the unsymmetric unsmoothed matter-matter-halo bispectrum of Eq. (24) we find for the cross-spectrum of a quadratic matter field with the halo density

$$P_{D[\delta_\mathrm{m}^R],\delta_\mathrm{h}^R}(k) = \int \frac{\mathrm{d}^3q}{(2\pi)^3} D(\mathbf{q},\mathbf{k}-\mathbf{q})B^\mathrm{unsym}_{\delta_\mathrm{m}^\mathrm{R}\delta_\mathrm{m}^\mathrm{R}\delta_\mathrm{h}^\mathrm{R}}(\mathbf{q},\mathbf{k}-\mathbf{q},-\mathbf{k}) \tag{64}$$

$$= 2b_1I_{DF_2}^R(k) + 4b_1I_{DF_2}^{\mathrm{bare},R}(k) + 2b_2I_{D\mathsf{P}_0}^R(k) + \frac{4}{3}b_{s^2}I_{D\mathsf{P}_2}^R(k). \tag{65}$$

The part depending on b_1 can be expressed in terms of the matter-matter-matter cross-spectrum so that

$$P_{D[\delta_\mathrm{m}^R],\delta_\mathrm{h}^R}(k) - b_1P_{D[\delta_\mathrm{m}^R],\delta_\mathrm{m}^R}(k) = 2b_2I_{D\mathsf{P}_0}^R(k) + \frac{4}{3}b_{s^2}I_{D\mathsf{P}_2}^R(k). \tag{66}$$

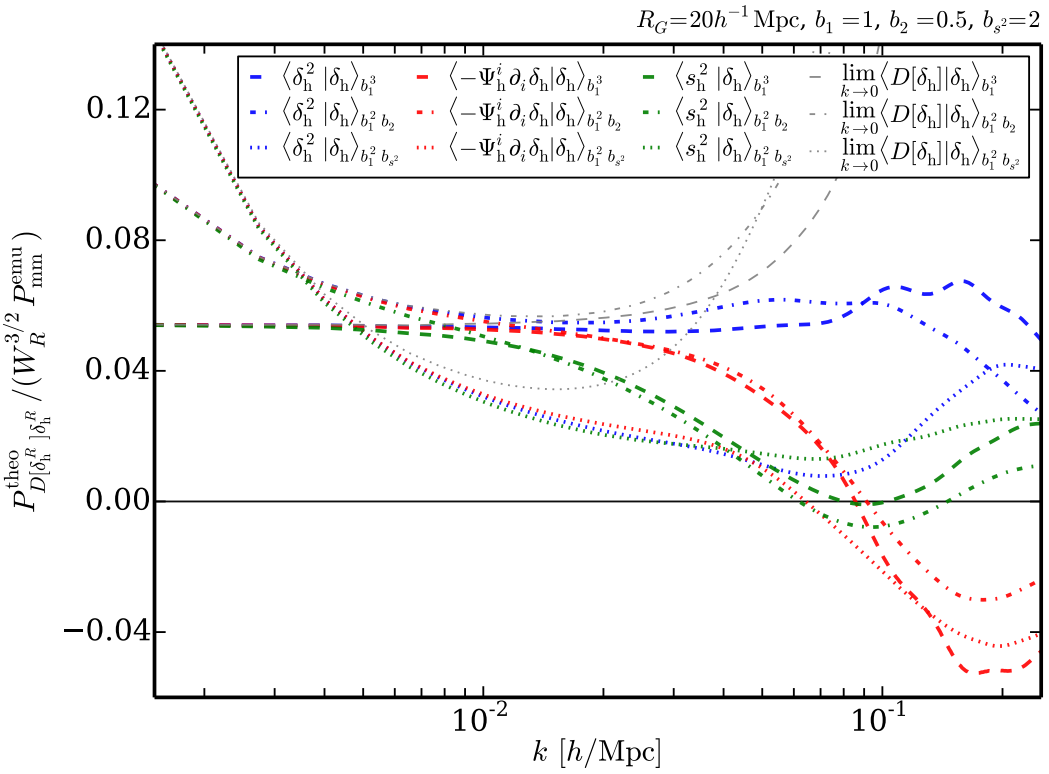

FIG. 1. Theory contributions (67) to halo-halo-halo cross-spectra scaling like b_1^3 (dashed), $b_1^2 b_2$ (dash-dotted) and $b_1^2 b_{s^2}$ (dotted) for squared density $\delta_{\rm h}^2(\mathbf{x})$ (blue), shift term $-\Psi_{\rm h}^i(\mathbf{x})\partial_i\delta_{\rm h}(\mathbf{x})$ (red) and tidal term $s_{\rm h}^2(\mathbf{x})$ (green), evaluated for fixed bias parameters $b_1 = 1$, $b_2 = 0.5$ and $b_{s^2} = 2$, Gaussian smoothing with $R_G = 20h^{-1}$Mpc, at $z = 0.55$, with linear matter power spectra in integrands. Thin gray lines show the large-scale (low k) limit given by Eq. (70). The cross-spectra are divided by the partially smoothed FrankenEmu emulator matter power spectrum $W_R^{3/2}P_{\rm mm}^{\rm emu}$ [45–48] for plotting convenience.

E. Halo-halo-halo cross-spectra

The halo-halo-halo bispectrum (28) gives for the halo-halo-halo cross-spectra

$$P_{D[\delta_{\rm h}^R],\delta_{\rm h}^R}(k) = 2b_1^3\left[I^R_{DF_2}(k) + 2I^{{\rm bare},R}_{DF_2}(k)\right] + 2b_1^2 b_2\left[I^R_{D\mathsf{P}_0}(k) + 2I^{{\rm bare},R}_{D\mathsf{P}_0}(k)\right] + \frac{4}{3}b_1^2 b_{s^2}\left[I^R_{D\mathsf{P}_2}(k) + 2I^{{\rm bare},R}_{D\mathsf{P}_2}(k)\right]. \quad (67)$$

Contributions depending on b_1^3 also appear in matter-matter-halo cross spectra, so that

$$P_{D[\delta_{\rm h}^R],\delta_{\rm h}^R}(k) - b_1^2 P_{D[\delta_{\rm m}^R],\delta_{\rm h}^R}(k) = 4b_1^2\left[b_2 I^{{\rm bare},R}_{D\mathsf{P}_0}(k) + \frac{2}{3}b_{s^2} I^{{\rm bare},R}_{D\mathsf{P}_2}(k)\right]. \quad (68)$$

Decomposing the F_2 kernel in Legendre polynomials as in Eq. (18), Eq. (67) can be rewritten as

$$\begin{aligned} P_{D[\delta_{\rm h}^R],\delta_{\rm h}^R}(k) = &\left(\frac{34}{21}b_1^3 + 2b_1^2 b_2\right)\left[I^R_{D\mathsf{P}_0}(k) + 2I^{{\rm bare},R}_{D\mathsf{P}_0}(k)\right] \\ &+ 2b_1^3\left[I^R_{D,F_2^1\mathsf{P}_1}(k) + 2I^{{\rm bare},R}_{D,F_2^1\mathsf{P}_1}(k)\right] \\ &+ \left(\frac{8}{21}b_1^3 + \frac{4}{3}b_1^2 b_{s^2}\right)\left[I^R_{D\mathsf{P}_2}(k) + 2I^{{\rm bare},R}_{D\mathsf{P}_2}(k)\right]. \end{aligned} \quad (69)$$

The contributions to the theory expression of Eq. (67) are shown in Fig. 1 for Gaussian smoothing with $R = 20h^{-1}$Mpc (see Fig. 10 in the appendix for $R = 10h^{-1}$Mpc). Different colors describe different cross-spectra, $D \in \{\mathsf{P}_0, -F_2^1\mathsf{P}_1, \mathsf{P}_2\}$, while different line styles correspond to the contributions with different dependencies on

bias parameters, scaling like b_1^3, $b_1^2 b_2$ or $b_1^2 b_{s^2}$. The characteristic k-dependencies of the different contributions can be exploited to fit b_1, b_2 and b_{s^2} to the three cross-spectra at the same time. In practice, the fitted bias parameters can still be degenerate because sampling variance at low k and modeling uncertainty at high k limit the usable k range. In particular, in the range $0.01h/\mathrm{Mpc} \lesssim k \lesssim 0.1h/\mathrm{Mpc}$, every cross-spectrum depends rather similarly on b_1^3 and $b_1^2 b_2$ leading to a degeneracy where larger b_1 can be compensated by a smaller b_2, which is also present when considering individual bispectrum triangles rather than cross-spectra. Consequently, models that extend leading-order PT to higher k are expected to improve bias constraints significantly.

In the large-scale limit, $k \ll q$, the three halo-halo-halo cross-spectra of Eq. (67) equal each other (see Appendix A for details),

$$\lim_{k\to 0} P_{D[\delta_\mathrm{h}^R],\delta_\mathrm{h}^R}(k) = W_R(k)\left[b_1^3 P_\mathrm{mm}^\mathrm{lin}(k)\left(\frac{68}{21}\sigma_R^2 - \frac{1}{3}\sigma_{R,P'}^2\right) + 2b_1^2 b_2\left(\tau_R^4 + 2P_\mathrm{mm}^\mathrm{lin}(k)\sigma_R^2\right) + \frac{4}{3}b_1^2 b_{s^2}\tau_R^4\right], \quad D \in \left\{\mathsf{P}_0, -F_2^1\mathsf{P}_1, \mathsf{P}_2\right\}, \quad (70)$$

where $W_R(k) \to 1$ for $k \to 0$, and $\sigma_{R,P'}^2$ and τ_R^4 are defined in Eqs. (A20) and (A16), respectively. In this limit, the b_1^3 term is proportional to the linear matter power spectrum, whereas the term involving b_2 scales like the linear matter power spectrum plus a k-independent correction and the term involving b_{s^2} is entirely k-independent (see thin gray lines in Figs. 1 and 10 for the the ratio of these limits over the matter power spectrum). Due to large sampling variance at low k, this scale-dependence is expected to be less powerful in distinguishing bias parameters than constraints obtained from the different scale-dependencies at high k.

F. Shot noise

The bispectrum of the smoothed halo density has an additional stochasticity contribution,

$$\langle \hat{B}_\mathrm{hhh}^R(k_1,k_2,k_3)\rangle = B_\mathrm{hhh}^R(k_1,k_2,k_3) + B_\mathrm{hhh}^{R,\mathrm{shot}}(k_1,k_2,k_3), \quad (71)$$

whose Poissonian prediction is (see e.g. [49])

$$B_\mathrm{hhh}^{R,\mathrm{shot}}(k_1,k_2,k_3) = W_R(k_1)W_R(k_2)W_R(k_3)\left\{\frac{1}{\bar{n}_\mathrm{h}}\left[P_\mathrm{hh}(k_1) + 2 \text{ perms}\right] + \frac{1}{\bar{n}_\mathrm{h}^2}\right\}. \quad (72)$$

Here, $\bar{n}_\mathrm{h}$ is the mean halo number density, and P_hh is the power spectrum of the unsmoothed continuous halo density field, which we approximate by the ensemble-averaged, CIC- and shot-noise-corrected power spectrum of the unsmoothed halo density measured in the simulations. The stochasticity bispectrum contributes to halo cross-spectra as

$$P_{D[\delta_\mathrm{h}^R],\delta_\mathrm{h}^R}^\mathrm{shot}(k) = \left[\frac{1}{\bar{n}_\mathrm{h}^2} + \frac{P_\mathrm{hh}(k)}{\bar{n}_\mathrm{h}}\right] J_D^R(k) + \frac{2}{\bar{n}_\mathrm{h}}\tilde{J}_D^R(k), \quad (73)$$

where we defined

$$J_D^R(k) \equiv W_R(k)\int \frac{\mathrm{d}^3q}{(2\pi)^3} W_R(q)W_R(|\mathbf{k}-\mathbf{q}|)D(\mathbf{q},\mathbf{k}-\mathbf{q}) \quad (74)$$

and

$$\tilde{J}_D^R(k) \equiv W_R(k)\int \frac{\mathrm{d}^3q}{(2\pi)^3} W_R(q)W_R(|\mathbf{k}-\mathbf{q}|)D(\mathbf{q},\mathbf{k}-\mathbf{q})P_\mathrm{hh}(q), \quad (75)$$

which depends on the mass bin through P_hh. The full model is

$$\langle \hat{P}_{D[\delta_\mathrm{h}^R],\delta_\mathrm{h}^R}(k)\rangle = P_{D[\delta_\mathrm{h}^R],\delta_\mathrm{h}^R}(k) + P_{D[\delta_\mathrm{h}^R],\delta_\mathrm{h}^R}^\mathrm{shot}(k), \quad (76)$$

where the first term on the r.h.s. depends on bias parameters to be fitted from data, while the stochasticity (shot noise) term does not explicitly depend on bias parameters because we use the measured ensemble-averaged halo power spectrum there.

The Poisson stochasticity should be corrected for exclusion and clustering effects, similarly to the power spectrum results of [50]. Phenomenologically, we can model these shot noise corrections with two scale-independent parameters, Δ_1 and Δ_2, by adding $\Delta_1[P_\mathrm{hh}(k_1) + 2\mathrm{perms}]$ to $\bar{n}_\mathrm{h}^{-1}[P_\mathrm{hh}(k_1) + 2\mathrm{perms}]$ and Δ_2 to $\bar{n}_\mathrm{h}^{-2}$ in Eq. (72), so that

$$P_{D[\delta_\mathrm{h}^R],\delta_\mathrm{h}^R}^\mathrm{shot}(k) = \left[\left(\bar{n}_\mathrm{h}^{-2} + \Delta_2\right) + \left(\bar{n}_\mathrm{h}^{-1} + \Delta_1\right)P_\mathrm{hh}(k)\right] J_D^R(k) + 2\left(\bar{n}_\mathrm{h}^{-1} + \Delta_1\right)\tilde{J}_D^R(k). \quad (77)$$

For $\Delta_2 = \Delta_1/\bar{n}_\mathrm{h}$, this is equivalent to rescaling the Poisson shot noise by an overall scale-independent amplitude as done in e.g. [12]. The Poisson prediction is recovered for $\Delta_1 = \Delta_2 = 0$.

G. Covariances

To estimate bias and cosmological parameters from cross-spectra we need to know their noise and covariance properties. Leading-order perturbation theory predicts for the covariance between two cross-spectra at the same wavenumber (see Appendix B)

$$\mathrm{cov}(\hat{P}_{D[\delta_a^R],\delta_b^R}(k), \hat{P}_{E[\delta_a^R],\delta_b^R}(k)) = \frac{2}{N_{\mathrm{modes}}(k)} P_{bb}^R(k) I_{DE}^{P_{aa}^R P_{aa}^R}(k), \tag{78}$$

where $I_{DE}^{P_{aa}^R P_{aa}^R}$ is defined in Eq. (B3). The correlation between two cross-spectra is therefore

$$\mathrm{correl}(\hat{P}_{D[\delta_a^R],\delta_b^R}(k), \hat{P}_{E[\delta_a^R],\delta_b^R}(k)) = \frac{I_{DE}^{P_{aa}^R P_{aa}^R}(k)}{\sqrt{I_{DD}^{P_{aa}^R P_{aa}^R}(k) I_{EE}^{P_{aa}^R P_{aa}^R}(k)}}, \tag{79}$$

which does not depend on the type b of the linear field. Note that the perturbative calculation in Appendix B predicts additional covariances between cross-spectra at different wavenumbers $k \neq k'$, but we neglect them here for simplicity. Eqs. (78) and (79) will be compared against simulations in Section V C and Fig. 7.

For sufficiently large smoothing scale R and low k, we can approximate $P_{\mathrm{hh}}^R \approx b_1^2 P_{\mathrm{mm}}^R$ in the integrand of Eq. (B3), so that the halo correlation ($a = \mathrm{h}$ in Eq. (79)) approaches the dark matter correlation ($a = \mathrm{m}$ in Eq. (79)). In the large scale limit $k \to 0$ the kernels $D, E \in \{\mathsf{P}_0, -F_2^1\mathsf{P}_1, \mathsf{P}_2\}$ become unity (see Appendix A), so that

$$\lim_{k\to 0} \mathrm{correl}(\hat{P}_{D[\delta_a^R],\delta_b^R}(k), \hat{P}_{E[\delta_a^R],\delta_b^R}(k)) = 1, \qquad D, E \in \{\mathsf{P}_0, -F_2^1\mathsf{P}_1, \mathsf{P}_2\} \tag{80}$$

i.e. the three cross-spectra are perfectly correlated on large scales. Therefore all information is already contained in any one of the three cross-spectra. On intermediate scales (higher k) the cross-spectra are less correlated (and have different expectation values), so that constraints are expected to improve if more than just one cross-spectrum is considered. On small scales, smoothing destroys clustering information and the cross-spectra become perfectly correlated or anti-correlated (see Fig. 7 below; this happens at higher k if smaller smoothing scale R is chosen).

On large scales, $k \to 0$, the variance of the cross-spectra scales like $P_{bb}^R(k)/k^2$, because the large-scale limit of Eq. (B3) is independent of k and $N_{\mathrm{modes}} \propto k^2$. This confirms that the k^2/P weighting in the optimal bispectrum estimators in Eqs. (36), (39), (45), (46) and (47) corresponds to inverse-variance weighting on large scales.

V. SIMULATIONS

A. Setup

We use ten realizations of N-body simulations that were also used in [51, 52]. The simulations were run with the TreePM code of [53]. Each realization has 2048^3 DM particles in a box of side length $L = 1380h^{-1}$Mpc. The cosmology is flat ΛCDM with $\Omega_b h^2 = 0.022, \Omega_m h^2 = 0.139, n_s = 0.965, h = 0.69$ and $\sigma_8 = 0.82$, and we only use the snapshot at $z = 0.55$. Details of the simulations can be found in [51, 52].

To obtain the DM density, for each realization the full set of 2048^3 DM particles is interpolated to a $N_g^3 = 512^3$ grid using the cloud-in-cell (CIC) scheme. Halos are identified using the FoF algorithm with linking length $b = 0.168$. The halo sample is split into four mass bins, each spanning a factor of three in mass, and interpolated to halo density grids using CIC. The CIC window is deconvolved from DM and halo densities. The inverse number density $1/\bar{n}$ in units of $h^{-3}\mathrm{Mpc}^3$ is 0.306 for dark matter and 351.5, 746.6, 2026.2 and 6561.3 for halo bins ordered by increasing mass.

Before squaring any fields, we apply a Gaussian smoothing filter (54) to the density. The squared density $\delta^2(\mathbf{x})$ in Eq. (2) is obtained by squaring this smoothed density in configuration space. To obtain the shift term $-\Psi^i(\mathbf{x})\partial_i\delta(\mathbf{x})$ of Eq. (3), the density is first Fourier-transformed to k space, where it is multiplied by $\mathbf{k}/k^2$ or $\mathbf{k}$ to get the displacement or density gradient fields in k space. Then both fields are Fourier-transformed back to configuration space, where they are multiplied and contracted as in Eq. (3). A similar procedure is used to obtain $s_{ij}(\mathbf{k})$ and $s^2(\mathbf{x})$ as defined in Eq. (6). Finally, the three quadratic fields $\delta^2(\mathbf{x})$, $-\Psi^i(\mathbf{x})\partial_i\delta(\mathbf{x})$ and $s^2(\mathbf{x})$ are Fourier-transformed back to k-space, where their cross-spectra with the density $\delta(\mathbf{k})$ are estimated using Eqs. (13) and (15).

The computational cost is dominated by Fourier transforms which can be evaluated efficiently as FFTs, requiring only $\mathcal{O}(N_g^3 \log N_g^3)$ operations. Therefore the cross-spectrum analysis with quadratic fields has the same complexity as a usual power spectrum analysis in k-space, but it is sensitive to the full bispectrum information. Note that

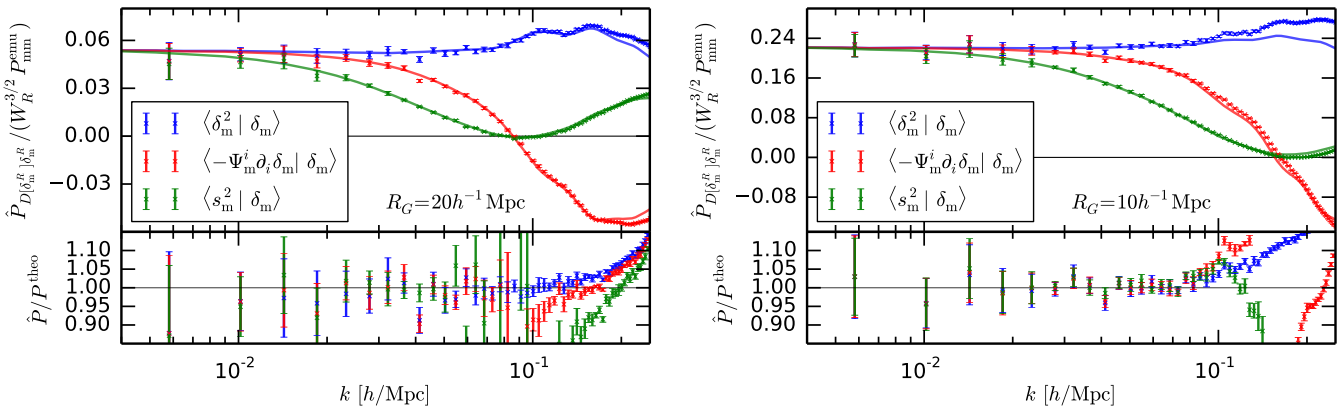

FIG. 2. Matter-matter-matter cross-spectra measured from 10 realizations at $z = 0.55$ (crosses with error bars), compared with leading order theory prediction of Eq. (61) (solid lines), neglecting shot noise. Upper panels show cross-spectra divided by the partially smoothed emulator matter power spectrum $W_R^{3/2}P_{\rm mm}^{\rm emu}$, lower panels show the ratio of measured cross-spectra over their theory expectation (61). Gaussian smoothing is applied with $R_G = 20h^{-1}$Mpc (left) and $R_G = 10h^{-1}$Mpc (right). Different colors represent different cross-spectra (squared density in blue, shift term in red and tidal term in green).

brute-force estimation of the bispectrum triangle by triangle is computationally more expensive by several orders of magnitude because it requires $\mathcal{O}(N_g^6)$ operations.

Theoretical expressions for expectation values and covariances use the linear matter-matter power spectrum computed by CAMB [54] at $z = 0.55$ for our fiducial cosmology. If theoretical expressions involve halo-halo power spectra, we use the estimated ensemble-averaged halo power spectrum corrected for shot noise and CIC. For plotting convenience, the cross-spectrum expectation values are typically divided by the partially smoothed nonlinear matter power spectrum $W_R^{3/2}P_{\rm mm}^{\rm emu}$ which is calculated with the FrankenEmu emulator [45–48]. Error bars in all plots of this paper show the standard error of the mean of the ten realizations, which is estimated as the sample standard deviation divided by $\sqrt{10}$. This corresponds to 1σ errors in the total volume $26.3h^{-3}$Gpc^3 of the ten realizations.[14]

B. Cross-spectrum expectation values

We test the consistency of the model by comparing theory expressions for matter-matter-matter, matter-matter-halo and halo-halo-halo cross-spectra with simulations. Since the goal of this section is to test the model, we consider statistics involving the dark matter field although they cannot directly be observed.

Fig. 2 compares matter-matter-matter cross-spectra measured in simulations against the theory expression of Eq. (61), finding agreement at the 5% level for $k \lesssim 0.09h$/Mpc for $R = 20h^{-1}$Mpc and $R = 10h^{-1}$Mpc. This demonstrates that the model for matter-matter-matter cross-spectra works well on large scales.

Figs. 3 and 11 test the model for matter-matter-halo cross-spectra by comparing the excess cross-spectra

$$\hat{P}_{D[\delta_m^R]\delta_h^R}(k) - \hat{b}_1\hat{P}_{D[\delta_m^R]\delta_m^R}(k) \tag{81}$$

to the theory expectation of Eq. (66), where $\hat{b}_1$ is obtained from large-scale $\hat{P}_{\rm hm}/\hat{P}_{\rm mm}$, while b_2 and b_{s^2} are jointly fitted to the three excess cross-spectra at $k \le 0.09h$/Mpc.[15] The plots show that for all mass bins there is a combination of b_2 and b_{s^2} that describes the simulations within statistical uncertainties for $k \lesssim 0.09h$/Mpc.[16] The fits are somewhat better for $R_G = 20h^{-1}$Mpc than for $R_G = 10h^{-1}$Mpc because the former excludes nonlinear mode coupling more efficiently.

[14] Due to the small number of realizations the estimated error bars are rather uncertain, see Fig. 7 below for a comparison with theoretical error bars.

[15] The fits in Figs. 3, 11 and 5 use estimated variances of the cross-spectra in the likelihood and neglect covariances for simplicity. In contrast, Fig. 9 uses theoretical covariances (78) between different cross-spectra with kernels $D \neq E$ at $k = k'$. Generally, error bars of bias parameters are consistent at the few percent level if theoretical instead of estimated variances are used. Including theoretical covariances (78) between different cross-spectra with kernels $D \neq E$ at $k = k'$ typically leads to fractional changes of bias parameter error bars by $\sim 10\%$ or less.

[16] In fact, the model seems to overfit the data in Figs. 3 and 11 because the reduced χ^2 is less than 1. This might be attributed to the fact that b_2 and b_{s^2} are degenerate in the excess matter-matter-halo cross-spectra, so that the fitting procedure can pick a parameter combination along the degeneracy that overfits the data. We do not address this issue further because the goal of this section is only to show that there are bias parameters for which the model of Eq. (66) agrees with simulations.

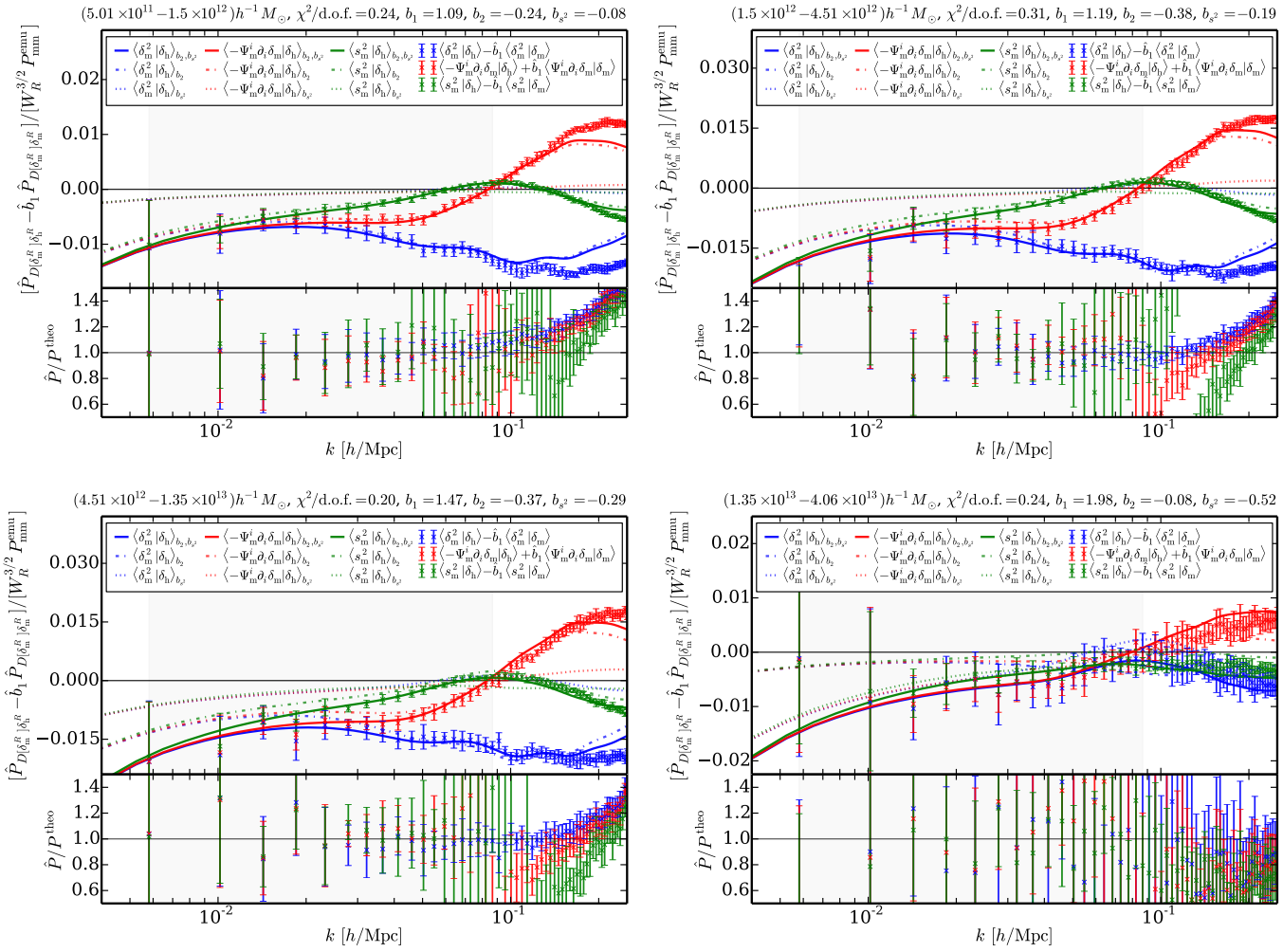

FIG. 3. Test of b_2 and b_{s^2} contributions to matter-matter-halo cross-spectra. First, $\hat{b}_1$ is obtained from $\hat{P}_{\rm hm}/\hat{P}_{\rm mm}$ at $k < 0.04h$/Mpc. Then b_2 and b_{s^2} are obtained by fitting the model (66) to the measured excess cross-spectra $\hat{P}_{D[\delta_{\rm m}^R]\delta_{\rm h}^R} - \hat{b}_1\hat{P}_{D[\delta_{\rm m}^R]\delta_{\rm m}^R}$ (crosses). The best-fit total model (solid lines) consists of the b_2 contribution (dash-dotted) and the b_{s^2} contribution (dotted), while shot noise is neglected. Different plots show different mass bins (increasing from upper left to lower right). The upper sub-panels show excess cross-spectra divided by the partially smoothed emulator matter power $W_R^{3/2}P_{\rm mm}^{\rm emu}$, the lower sub-panels show measured excess cross-spectra divided by their theory expectation. The fit is obtained from the grey shaded region, assuming estimated standard errors of the mean without any covariances. Best-fit bias parameters, reduced χ^2 and halo mass range are reported at the top of each plot. Gaussian smoothing with $R_G = 20h^{-1}$Mpc is applied to matter and halo densities.

Fig. 4 tests the theory prediction of Eq. (67) for halo-halo-halo cross-spectra for $R_G = 20h^{-1}$Mpc. The full measured halo-halo-halo cross-spectra are compared with their theory prediction (67) for bias parameters fixed to the values obtained from matter-halo statistics (see caption for details) and for halo-halo-halo shot noise (stochasticity) fixed to be Poissonian, corresponding to $\Delta_1 = \Delta_2 = 0$ in Eq. (77). While theory and simulations do not differ strongly for the lowest and highest mass bins, the simulations show a clear excess over theory for the two intermediate mass bins. Possible reasons for this could be that the perturbative treatment breaks down or shot noise is not Poissonian.

To test the latter, Fig. 5 shows the same cross-spectra if the shot noise correction Δ_1 in Eq. (77) is varied as a free parameter and fitted to the measured halo-halo-halo cross-spectra, imposing $\Delta_2 = \Delta_1/\bar{n}_{\rm h}$ and keeping bias parameters fixed to their values from matter-halo statistics. This clearly improves the agreement between simulations and theory, especially for the two intermediate mass bins.[17] For our fiducial four mass bins the reduced χ^2 does not

[17] The worst reduced χ^2 is 2.16, which would improve further if the lowest k-bin was removed (the theory of this bin is rather noisy because the estimated halo-halo power is used to compute the shot noise in Eq. (77)).

improve significantly if Δ_2 is treated as a free parameter, i.e. $\Delta_2 = \Delta_1/\bar{n}_\mathrm{h}$ seems to be an acceptable approximation within the error bars of the simulations.

To further test the shot noise corrections, we consider two additional mass bins above the fiducial four mass bins used in the rest of the paper, with linear bias $b_1 = 2.8$ and 4.3. Fitting Δ_1 and imposing $\Delta_2 = \Delta_1/\bar{n}_\mathrm{h}$ gives $\Delta_1 = -4.7h^{-3}\mathrm{Gpc}^3$ with $\chi^2/\mathrm{d.o.f.} = 1.8$ for the $b_1 = 2.8$ mass bin, and $\Delta_1 = -16.6h^{-3}\mathrm{Gpc}^3$ with $\chi^2/\mathrm{d.o.f.} = 3.4$ for the $b_1 = 4.3$ mass bin. For these two high-mass bins, the reduced χ^2 of the fits improve to 1.4 and 0.67 if Δ_2 is treated as a free parameter; see Fig. 6.

The shot noise correction Δ_1 is positive for the lowest four mass bins, but becomes negative for the two very high mass bins shown in Fig. 6. This mass dependence is qualitatively consistent with Fig. 11 of [50], where the shot noise correction to the power spectrum turns negative at around $3 \times 10^{13} h^{-1} M_\odot$, because the exclusion effect dominates at high mass. While alternative modifications of the model might be able to describe the simulations similarly well, the qualitative agreement of the mass dependence of the shot noise correction and the fact that the shot noise corrections are capable to capture the cross-spectrum measurements for all mass bins indicate that deviations from Poisson shot noise caused by exclusion and nonlinear biasing are indeed responsible for the disagreement in Fig. 4. If so one should be able to model these effects rather than treat them as a free parameter (see Fig. 11 of [50] for a theoretical model prediction that qualitatively agrees with our measurements). We note that a more detailed modeling in [50] predicts this stochasticity term to be constant (i.e. shot noise like) only for low k, and is expected to vanish at high k (with the transition given by the halo radius scale). A more detailed analysis is needed to investigate what the appropriate form is for the bispectrum analysis, and we expect that the phenomenological approach adopted here can be improved considerably with a more detailed modeling.

C. Cross-spectrum covariances

For the case of matter-matter-matter cross-spectra and $R_G = 20h^{-1}\mathrm{Mpc}$, Fig. 7 tests the predicted cross-spectrum covariances of Eq. (78) (counting the number of modes manually in the code for the given k binning) against estimates obtained from ten realizations. Due to the small number of realizations, the covariance estimates are rather noisy. Within this large uncertainty, the standard deviations of the three cross-spectra (Fig. 7 left) as well as the cross-correlations between the cross-spectra at the same scale $k = k'$ (Fig. 7 right) are consistent between simulations and theory at $k \lesssim 0.2h/\mathrm{Mpc}$.

For matter-matter-halo and halo-halo-halo cross-spectra, we find similar agreement; see Fig. 8 for the $b_1 = 1.98$ halo mass bin. Similar results are obtained for lower mass bins (not shown for brevity). In particular, the ratio of measured over theoretical standard deviations fluctuates between 0.5 and 2 for all mass bins for $k \lesssim 0.2h/\mathrm{Mpc}$.

More work is needed to test the covariances at higher precision, e.g. by running more realizations or by dividing simulation boxes into sub-boxes. We also leave it for future work to test the theoretical covariances (B4) between cross-spectra at different scales $k \neq k'$.

D. Bias estimation from halo-halo-halo cross-spectra

While the goal of this section thus far has been to test the theory predictions against simulations, we now aim to get a rough sense of how well bias parameters could be measured from observable halo-halo-halo cross spectra alone. This should be regarded mainly as a motivation to study these observables in more detail in the future rather than a realistic forecast, because we make a number of idealistic assumptions that are not valid in practice: the covariance between cross-spectra is assumed to be given by the leading-order theoretical expression of Eq. (79), neglecting any covariance between different wavenumbers $k' \neq k$; observations are assumed to be in periodic boxes in real space, neglecting redshift-space distortions; the shot noise correction is treated as a free parameter; and the cosmology is fixed to the fiducial cosmology of the N-body runs. Most of these assumptions are likely to impact the constraining power of the cross-spectra and should be addressed before considering real data.

Under the above idealistic assumptions, we use the Monte-Carlo sampler emcee [55] to fit the bias parameters b_1, b_2 and b_{s^2} as well as the dimensionless shot noise amplitude $A_\mathrm{shot} \equiv -\bar{n}_\mathrm{h}\Delta_1$ (imposing $\Delta_2 = \Delta_1/\bar{n}_\mathrm{h}$) to the three ensemble-averaged halo-halo-halo cross-spectra $P_{D[\delta_\mathrm{h}^R],\delta_\mathrm{h}^R}(k)$, $D \in \{\mathsf{P}_0, -F_2^1\mathsf{P}_1, \mathsf{P}_2\}$ for the halo mass bin with $b_1 = 1.98$. Colored contours in Fig. 9 show results if each cross-spectrum is fitted individually, while the black contours correspond to the joint fit to all three cross-spectra. Focusing on the leftmost column of Fig. 9, we see that the linear bias b_1 is best determined by the shift cross-spectrum. Adding the other two cross-spectra does not tighten the b_1 constraint much, i.e. the shift cross-spectrum contains almost the entire bispectrum information on b_1. Once b_1 is known, the constraint on b_2 is mostly tightened by the squared density cross-spectrum (blue degeneracy regions are thinner than green ones in the b_1 vs b_2 panel of Fig. 9), while the constraint on b_{s^2} is mostly tightened by the tidal

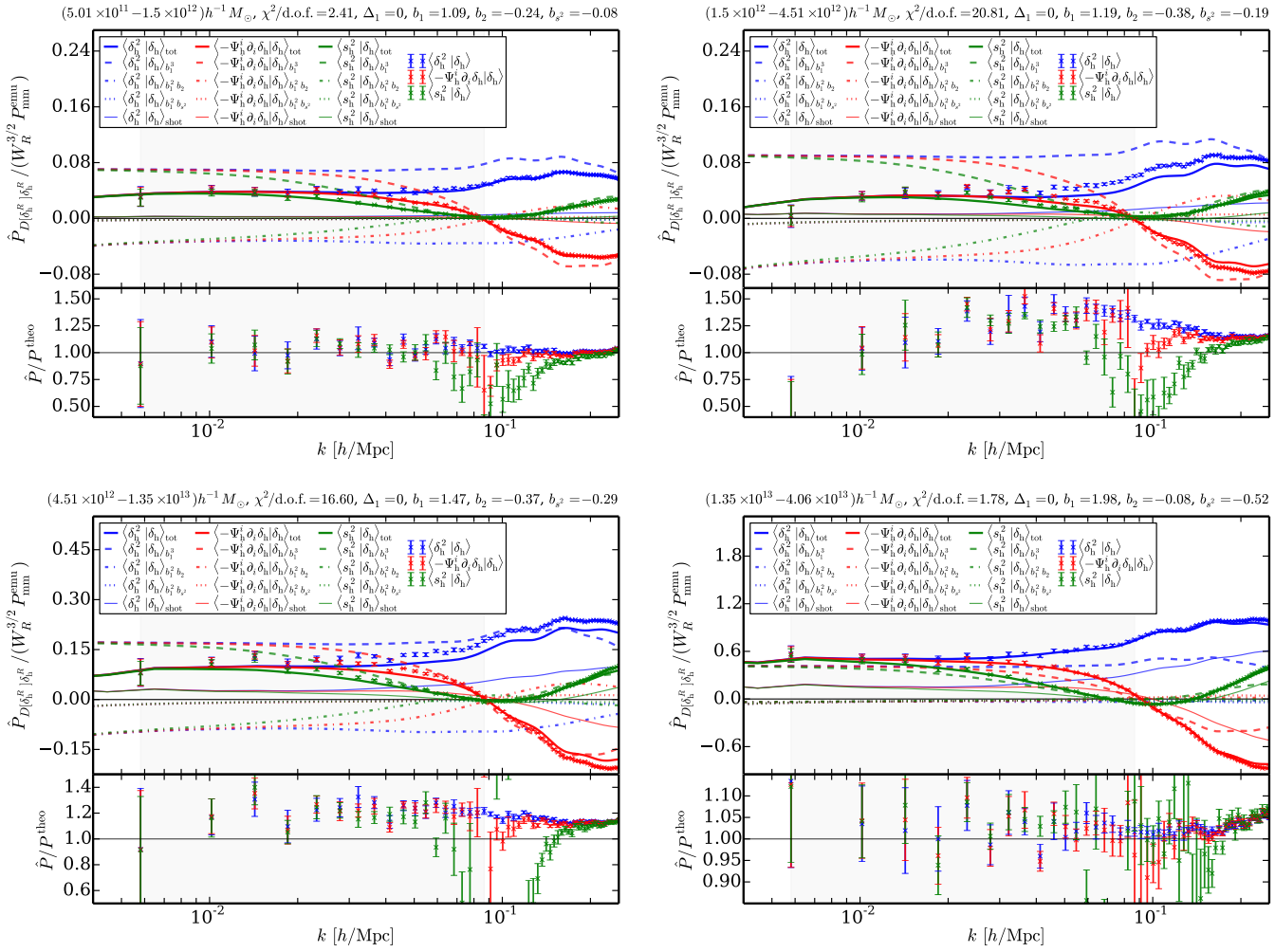

FIG. 4. Measured halo-halo-halo cross-spectra (crosses) compared against theory (thick solid, Eq. (67)) with bias parameters $\hat{b}_1$ from $\hat{P}_{\rm hm}/\hat{P}_{\rm mm}$ and b_2 and b_{s^2} from $\hat{P}_{D[\delta_{\rm h}^R]\delta_{\rm m}^R} - \hat{b}_1\hat{P}_{D[\delta_{\rm m}^R]\delta_{\rm m}^R}$ for $R_G = 20h^{-1}$Mpc smoothing. Upper panels also show theory contributions scaling like b_1^3 (dashed), $b_1^2 b_2$ (dash-dotted) and $b_1^2 b_{s^2}$ (dotted), as well as the halo-halo-halo shot noise contribution (thin solid), which is assumed to be Poissonian (i.e. $\Delta_1 = \Delta_2 = 0$ in Eq. (77)). The reduced χ^2 on top of the plots quantifies the (dis-)agreement between halo-halo-halo cross-spectra measurements and model for the fixed bias parameters. It is computed over the gray region, neglecting covariances. Note that the shot noise contribution fluctuates on very large scales because it is computed using the ensemble-averaged estimated halo-halo power spectrum.

cross-spectrum (green degeneracy regions are thinner than blue ones in the b_1 vs b_{s^2} panel). Intuitively, this can be understood from the Legendre decomposition of the halo-halo-halo bispectrum in Eqs. (29)-(31): The shift term corresponding to $l = 1$ is the only term that picks up b_1 without any contribution from nonlinear bias b_2 or b_{s^2} (in absence of velocity bias), while the $l = 0$ squared density term is the only one that gets contributions from b_2 and the $l = 2$ tidal term is the only term depending on b_{s^2}.[18]

Overall, the best-fit parameters from the combined fit, $(b_1, b_2, b_{s^2}, A_{\rm shot}) = (1.96, -0.02, -0.44, -0.007)$, agree with the values obtained from matter-halo statistics, $(1.98, -0.08, -0.52, -5\times10^{-5})$, within the uncertainties shown in Fig. 9, but the 1σ error of the mean of b_1 is relatively large ($\sim 5\%$, using the full $26.3h^{-3}$Gpc^3 volume). The uncertainties decrease significantly for higher $k_{\rm max}$ and smaller smoothing scales R, e.g. b_1 could be measured with $\sim 0.4\%$ precision for $k_{\rm max} = 0.24h$/Mpc and $R = 5h^{-1}$Mpc, but best-fit parameters become inconsistent with the true values, demonstrating that modeling needs to be improved or appropriate clipping or logarithmic density transforms

[18] This intuitive picture is not exact though: In practice, finite $k_{\rm max}$ and smoothing imply that each cross-spectrum actually picks up dependencies on all bias parameters, see Fig. 9.

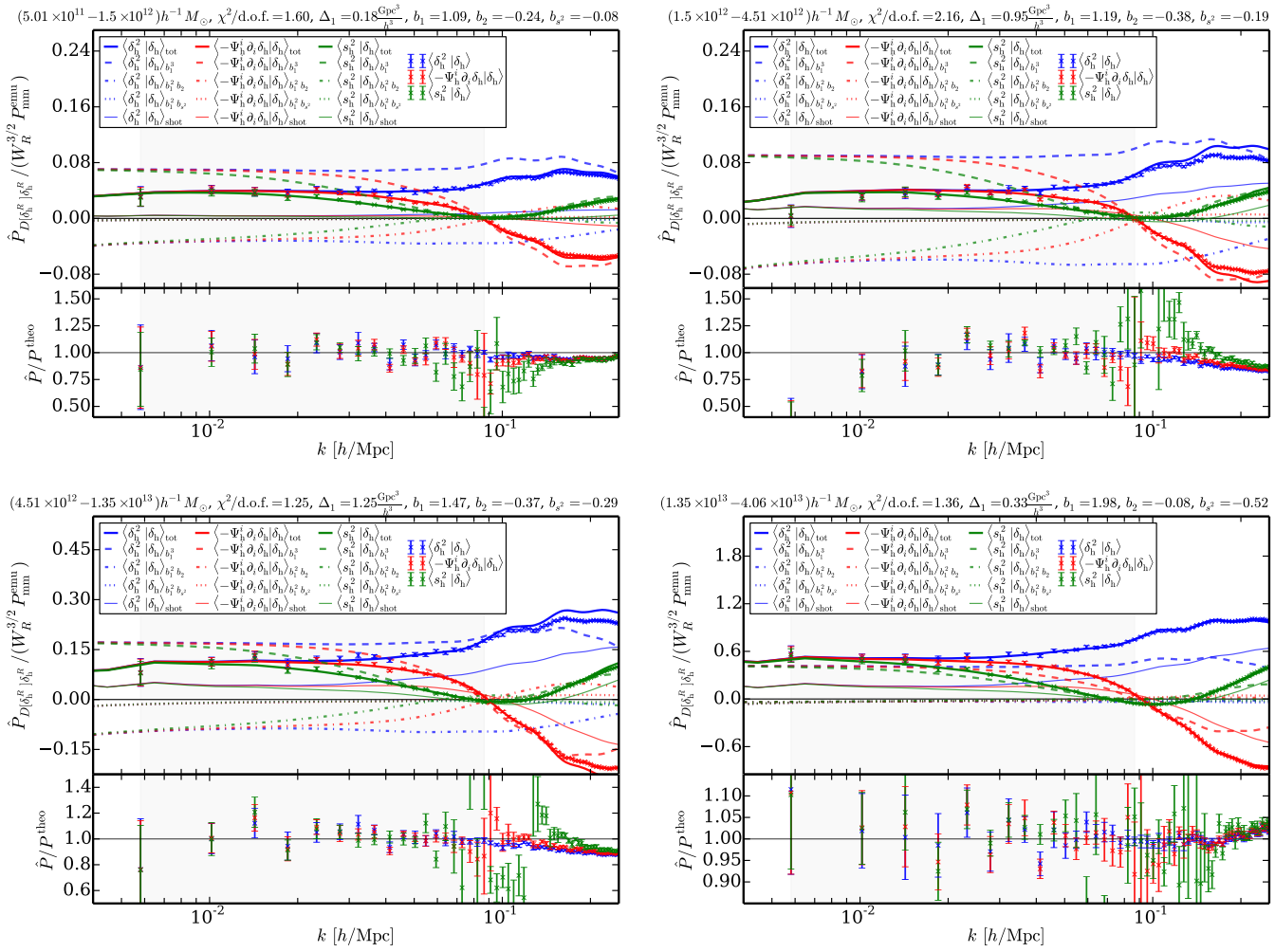

FIG. 5. Same as Fig. 4 if the shot noise correction Δ_1 in Eq. (77) is fitted to measured halo-halo-halo cross-spectra, fixing $\Delta_2=\Delta_1/\bar{n}_\mathrm{h}$ (still keeping b_1, b_2 and b_{s^2} fixed to the values obtained from matter-halo and matter-matter-halo measurements).

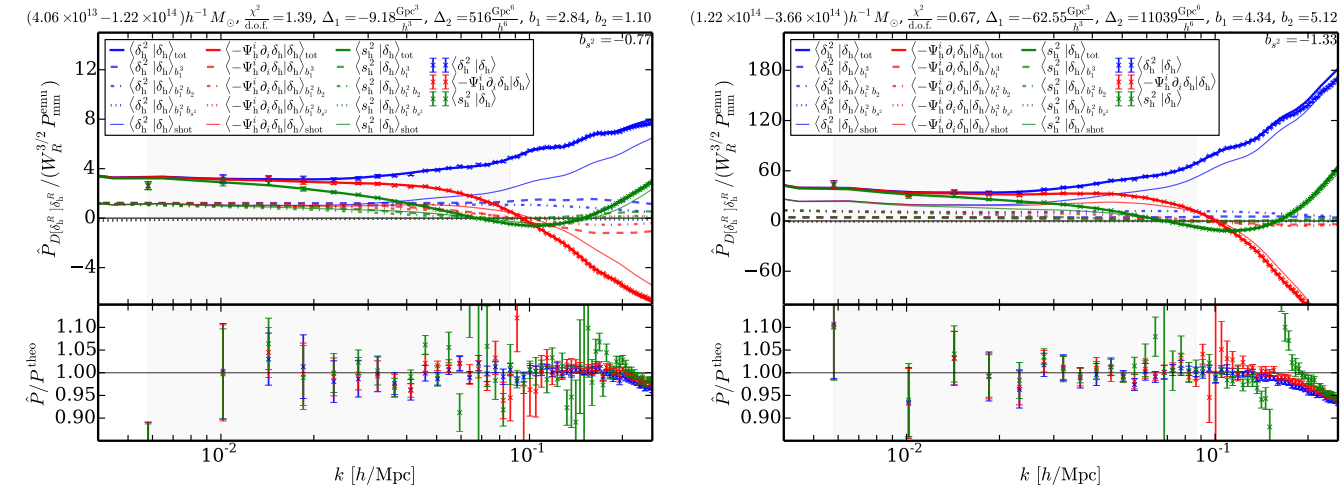

FIG. 6. Same as Fig. 5 for two higher mass bins and treating both shot noise corrections Δ_1 and Δ_2 in Eq. (77) as independent parameters.

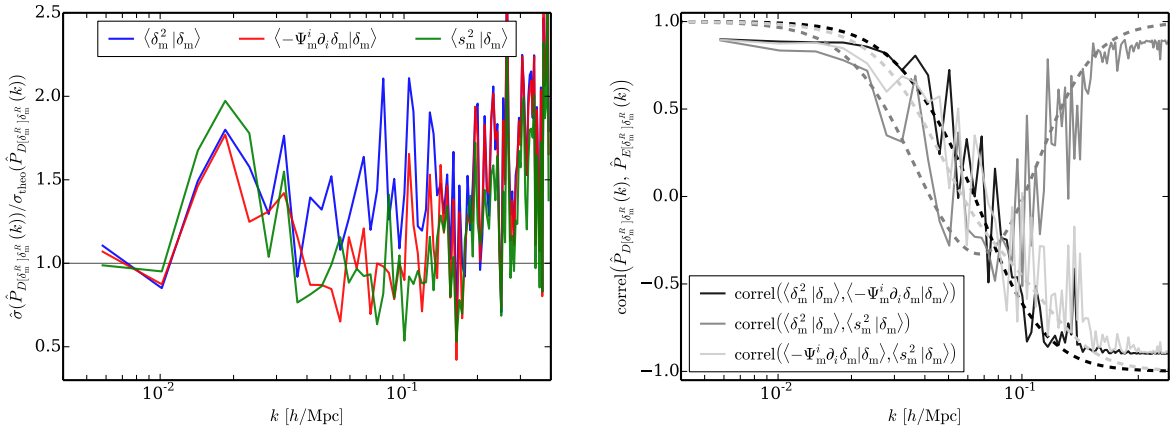

FIG. 7. *Left:* Ratio of estimated over theoretical standard deviation of matter-matter-matter cross-spectra (using $D = E$ and $a = b =$ m in Eq. (78)). *Right:* Correlations between matter-matter-matter cross-spectra at the same scale $k = k'$ predicted by theory (Eq. (79), dashed) and estimated from simulations (solid). Both panels assume Gaussian smoothing with $R_G = 20h^{-1}$Mpc and use the linear matter power spectrum in theory expressions. For smaller smoothing scale R, the zero crossing would move to higher k.

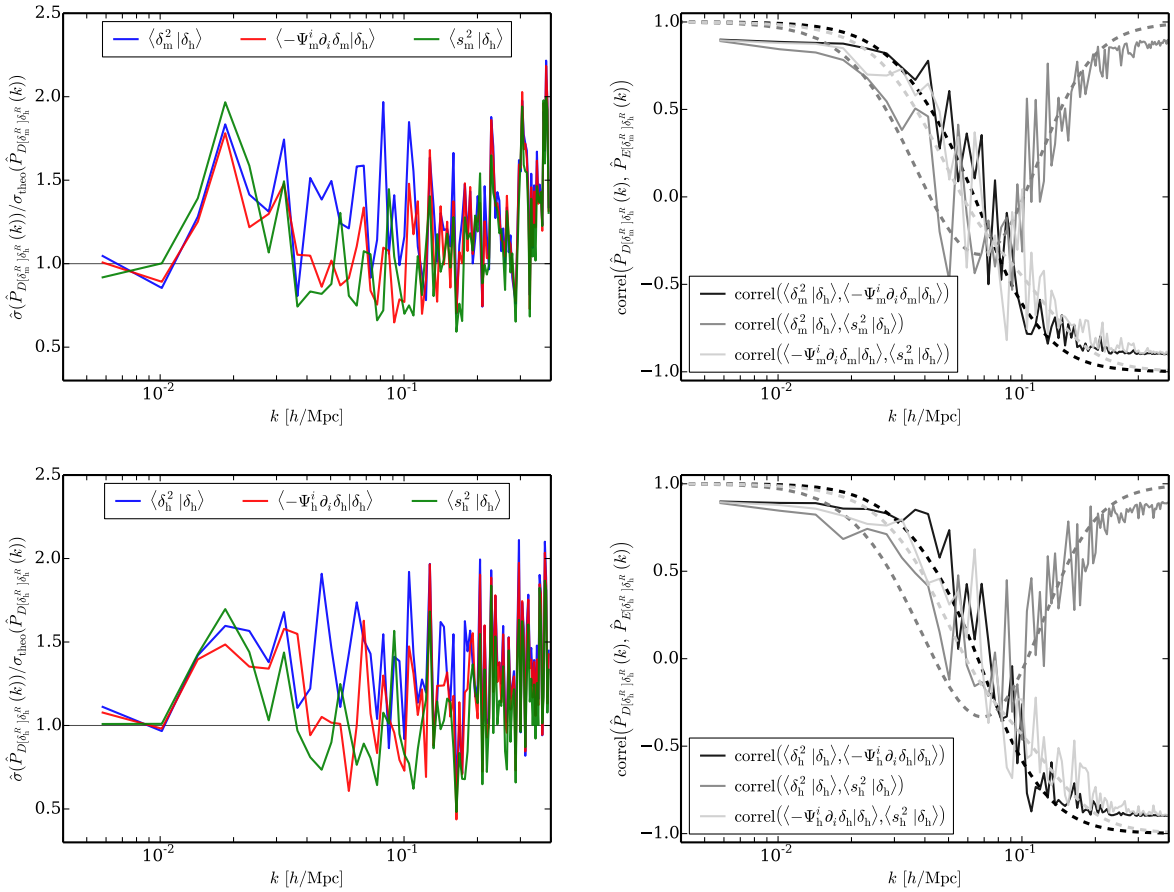

FIG. 8. Same as Fig. 7 but for matter-matter-halo cross-spectra (top) and halo-halo-halo cross-spectra (bottom), using the halo mass bin with $b_1 = 1.98$. Theoretical standard deviations (78) are evaluated using the estimated ensemble-averaged halo-halo power spectrum including shot noise for $P_{\rm hh}$ and the linear matter power spectrum for $P_{\rm mm}$. Theory correlations of Eq. (79) are the same as for the matter-matter-matter case. Results for other mass bins are similar (not shown).

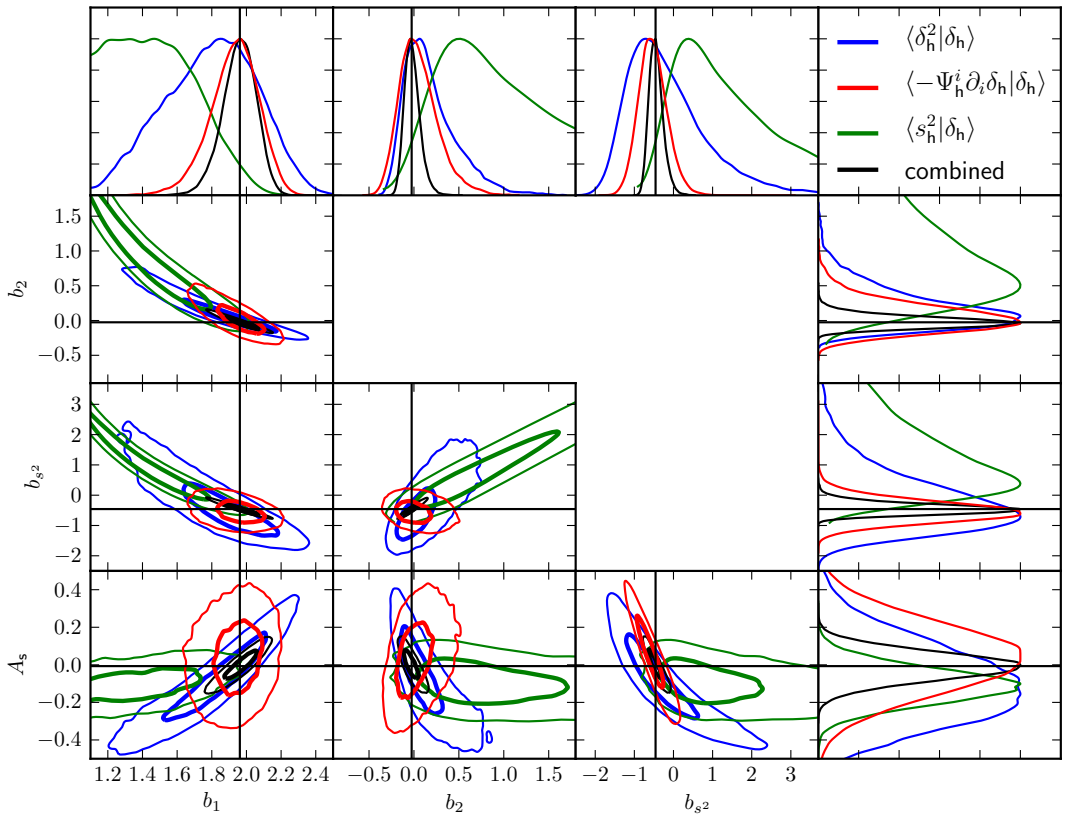

FIG. 9. Results of fitting b_1, b_2, b_{s^2} and $A_{\rm shot} \equiv -\bar{n}_{\rm h}\Delta_1$ (imposing $\Delta_2 = \Delta_1/\bar{n}_{\rm h}$) to halo-halo-halo cross-spectra involving the squared density (blue), shift term (red) or tidal term (green), or all three combined (black), choosing Gaussian $R = 20h^{-1}$Mpc smoothing and $k_{\rm max} = 0.09h$/Mpc for the halo mass bin with $b_1 = 1.98$. The 2d contours of the posterior show 68% and 95% confidence regions corresponding to the full volume of $V = 26.3h^{-3}{\rm Gpc}^3$ (i.e. errors in a single realization would be larger by a factor of $\sqrt{10}$). Thin black lines show the maximum-likelihood points corresponding to the black contours. The joint likelihood for the three halo-halo-halo cross-spectra is assumed to be Gaussian in the cross-spectra, with non-zero covariance between cross-spectra at $k = k'$ given by the theory expression of Eq. (78). The green contours are somewhat uncertain because it is not clear how well the MCMC chains sampled the elongated degeneracies.

[40, 41] need to be applied before it is possible to push to such small scales. This would also make the full advantage of the efficient cross-spectrum method over conventional brute-force methods more apparent because the latter become computationally unfeasible at high k.

The degeneracies of the combined contours in Fig. 9 correspond to the typical degeneracies between bias parameters estimated from the halo-halo-halo bispectrum if redshift-space distortions are neglected, e.g. higher b_1 can be compensated by lower b_2 or b_{s^2}. Since the shot noise correction parameterised by $A_{\rm shot}$ is also rather degenerate with the bias parameters, it would be interesting to model this analytically instead of treating it as a free parameter.

As a consistency check of the assumed likelihood, jack-knife error bars are obtained by fitting to each of the ten realizations individually and calculating the scatter of the best-fit values among realizations. These jack-knife error estimates are larger than the average uncertainty predicted by the likelihood width for a single realization by a factor of ~ 1.8 for the conservative $k_{\rm max} = 0.09h$/Mpc and by a factor of ~ 1.25 for the more ambitious $k_{\rm max} = 0.24h$/Mpc. This could be attributed to additional contributions to the true covariance that are neglected in our theoretical covariance (e.g. at $k' \neq k$), or uncertainty in the determination of the jack-knife errors due to the small number of realizations, or a departure of the true likelihood from a Gaussian pdf.

Summarizing the above, it seems possible to reach percent-level estimates of the linear bias b_1 from halo-halo-halo cross-spectra in real surveys, but we re-emphasize that we made a number of unrealistic assumptions, which could easily change this conclusion (to the better or worse).

VI. EXTENSION TO PRIMORDIAL NON-GAUSSIANITY

So far we have assumed Gaussian initial conditions. Multiple-field inflation models can generate local primordial non-Gaussianity that induces an additional contribution to the matter-matter-matter bispectrum of the local form

$$B^{\rm loc}_{\rm mmm}(k_1,k_2,k_3) = 2f^{\rm loc}_{\rm NL}\left[\frac{M(k_3)}{M(k_1)M(k_2)}P_{\rm mm}(k_1)P_{\rm mm}(k_2) + 2 \text{ perms}\right], \tag{82}$$

where $M(k) = M(k,z)$ is the linear Poisson conversion factor between the primordial potential Φ and the late-time matter density at redshift z,

$$M(k,z) \equiv \frac{2}{3}\frac{k^2T(k)D(z)}{\Omega_{\rm m}H_0^2}, \tag{83}$$

so that $\delta^{\rm lin}_{\rm m}(\mathbf{k},z) = M(k,z)\Phi(\mathbf{k})$. Here, $T(k)$ is the linear transfer function normalized to $T(k) = 1$ on large scales, and the linear growth factor $D(z)$ for $\Omega_{\rm rad} = 0$ is normalized to $D(z) = 1/(1+z)$ during matter domination. Note that $M(k) \propto k^2$ for $k \ll k_{\rm eq}$ and $M(k) \propto k^0$ for $k \gg k_{\rm eq}$. The bispectrum (82) is maximal in the squeezed limit (e.g. $k_1 \ll k_2 \approx k_3$).

Plugging the bispectrum (82) into Eq. (36), we get

$$\hat{f}^{\rm loc}_{\rm NL} = \frac{24\pi L^3}{N_{\rm loc}}\int {\rm d}k \frac{k^2M^2(k)}{P_{\rm mm}(k)}\hat{P}_{[\frac{\delta_{\rm m}}{M}]^2,\frac{\delta_{\rm m}}{M}}(k), \tag{84}$$

where we defined the quadratic field

$$\left[\frac{\delta_{\rm m}}{M}\right]^2(\mathbf{k}) \equiv \int \frac{{\rm d}^3q}{(2\pi)^3}\frac{\delta_{\rm m}(\mathbf{q})}{M(q)}\frac{\delta_{\rm m}(\mathbf{k}-\mathbf{q})}{M(|\mathbf{k}-\mathbf{q}|)} \tag{85}$$

and the filtered density

$$\frac{\delta_{\rm m}}{M}(\mathbf{k}) \equiv \frac{\delta_{\rm m}(\mathbf{k})}{M(k)}. \tag{86}$$

At leading order, this equals the primordial potential Φ reconstructed from the DM density $\delta_{\rm m}$. The cross-spectrum in Eq. (84) then probes the cross-spectrum of this reconstructed Φ with $\Phi^2(\mathbf{x})$, which corresponds to the mechanism that generates primordial non-Gaussianity of the local kind (adding $f_{\rm NL}\Phi^2(\mathbf{x})$ to $\Phi(\mathbf{x})$).

The cross-spectrum appearing in Eq. (84) could be used to estimate $f^{\rm loc}_{\rm NL}$ if the dark matter density was directly observable. The extension to observable halo densities is left for future work. It would also be straightforward to extend the cross-spectra to other separable types of primordial non-Gaussianity generated by other inflation models (e.g. equilateral or orthogonal).

VII. CONCLUSIONS

In this paper we explore methods to probe large-scale structure bispectrum parameters in a nearly optimal way. The tree level bispectrum receives contributions from gravity at second order, which can be Legendre decomposed into the squared density $\delta^2(\mathbf{x})$, the shift term $-\Psi^i(\mathbf{x})\partial_i\delta(\mathbf{x})$ and the tidal term $s^2(\mathbf{x}) = \frac{3}{2}s_{ij}(\mathbf{x})s_{ij}(\mathbf{x})$. When applied to galaxies or halos the gravity term is multiplied by the appropriate linear bias b_1 factor (e.g. b_1^3 when investigating the halo bispectrum). In addition, nonlinear biasing can introduce two additional terms that contribute at second order, $b_2\delta^2(\mathbf{x})$ and $b_{s^2}s^2(\mathbf{x})$. There is no nonlinear bias associated with the shift term in the absence of velocity bias. Since any velocity bias must vanish in the $k \to 0$ limit as a consequence of Galilean invariance, we do not include any such term.[19]

These terms correspond to individual components of the bispectrum in a separable form. In this case, in the limit where tree level theory is valid, one can write an optimal bispectrum estimator using these terms. Specifically, given a density $\delta(\mathbf{x})$, smoothed on the smallest scale where we still trust the theory predictions, the procedure we propose is as follows:

---

[19] At a k^2 level there could be a velocity bias, but we ignore this here since we work at the lowest order in k. Indeed, all the biasing terms can receive k^2 type corrections [42].

1. Compute the density gradient $\partial_i\delta(\mathbf{x})$, the displacement field $\Psi_i(\mathbf{x}) = -\partial_i\partial^{-2}\delta(\mathbf{x})$, and the tidal tensor $s_{ij}(\mathbf{x}) = \left[\partial_i\partial_j\partial^{-2} - \frac{1}{3}\delta_{ij}^{(K)}\right]\delta(\mathbf{x})$.
2. Compute the squared density $\delta^2(\mathbf{x})$, the shift term $-\Psi^i(\mathbf{x})\partial_i\delta(\mathbf{x})$ and the tidal term $s^2(\mathbf{x}) = \frac{3}{2}s_{ij}(\mathbf{x})s_{ij}(\mathbf{x})$.
3. Fourier transform the three quadratic fields to get $[\delta^2](\mathbf{k})$, $[-\Psi^i\partial_i\delta](\mathbf{k})$ and $[s^2](\mathbf{k})$.
4. Compute the cross-spectra between the quadratic fields and the density, i.e. (suppressing division by the number of modes)

$$\hat{P}_{\delta^2,\delta}(k) \sim \sum_{\mathbf{k},|\mathbf{k}|=k} [\delta^2](\mathbf{k})\delta(-\mathbf{k}), \tag{87}$$

$$\hat{P}_{-\Psi^i\partial_i\delta,\delta}(k) \sim \sum_{\mathbf{k},|\mathbf{k}|=k} [-\Psi^i\partial_i\delta](\mathbf{k})\delta(-\mathbf{k}) \tag{88}$$

$$\hat{P}_{s^2,\delta}(k) \sim \sum_{\mathbf{k},|\mathbf{k}|=k} [s^2](\mathbf{k})\delta(-\mathbf{k}). \tag{89}$$

As expected from the Legendre decomposition of the halo bispectrum, the $l = 1$ shift cross-spectrum (88) contains almost the entire bispectrum information on the linear bias b_1, while the $l = 0$ squared density cross-spectrum (87) and the $l = 2$ tidal cross-spectrum (89) mostly improve constraints on b_2 and b_{s^2} once b_1 is known. Measuring all three cross-spectra and comparing them to their theory predictions is equivalent to an optimal maximum-likelihood estimation of the amplitudes of contributions to dark matter or halo bispectra (under certain regularity conditions; see Section III). Therefore, these cross-spectra contain the same constraining power on bias parameters and σ_8 as a full optimal bispectrum analysis. Measuring cross-spectra is both simpler and computationally cheaper than performing direct bispectrum measurements for individual triangle configurations. Since they only depend on a single rather than three wavenumbers, modeling the covariance is also simpler.

We have derived leading-order perturbation theory predictions for the expectation values and covariances of the three cross-spectra, where both the quadratic and the single field can be dark matter or halo fields, and second order bias b_2 and tidal tensor bias b_{s^2} are included. The results are given by integrals over matter-matter-matter, matter-matter-halo or halo-halo-halo bispectra.

The proposed cross-spectra were measured on a set of ten large N-body simulations. The expectation values are consistent with perturbation theory at the few percent level for $k \lesssim 0.09h/\mathrm{Mpc}$ at $z = 0.55$ for matter-matter-matter and matter-matter-halo combinations, if all fields are smoothed by a Gaussian with smoothing scale $R = 20h^{-1}\mathrm{Mpc}$. For halo-halo-halo cross-spectra, one must include corrections to the Poisson stochasticity. While these corrections are qualitatively similar to corrections to the halo-halo power spectrum due to exclusion and nonlinear biasing [50], future work should investigate and model them in more detail. The predicted variance of the cross-spectra and the covariance between any two cross-spectra at the same wavenumber are found to be consistent with simulations (although the numerical noise is somewhat large given the small number of independent realizations).

The ultimate goal of this is to determine the three bias parameters and dark matter clustering power spectrum by combining these three statistics with the measured galaxy power spectrum. We have not performed this step in this paper: we plan to explore potential improvements of the modeling by including higher order perturbation theory terms as well as improved bias and stochasticity models to measure the new observables in galaxy surveys in future work. We also plan to include redshift space distortions by accordingly modifying and extending the cross-spectra. Given the simplicity of the method and the agreement with leading-order perturbation theory on large scales, we hope it will become a useful tool to break degeneracies of bias and cosmological parameters.

ACKNOWLEDGEMENTS

We thank Florian Beutler and Blake Sherwin for many helpful discussions, and Martin White and Beth Reid for providing the N-body simulations used in this paper. We are also very grateful to Anatoly Klypin for sharing a CIC and power spectrum code. We acknowledge useful discussions with James Fergusson, Airam Marcos-Caballero, Paul Shellard and Zvonimir Vlah. TB gratefully acknowledges support from the Institute for Advanced Study through the Corning Glass Works Foundation Fellowship. This research used resources of the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. It also used the COSMOS Shared Memory system at DAMTP, University of Cambridge operated on behalf of the STFC DiRAC HPC Facility and funded by BIS National E-infrastructure capital grant ST/J005673/1 and STFC grants ST/H008586/1, ST/K00333X/1.

Appendix A: Large-scale limits of integrals

The large-scale limits $k \ll q$ of the integrals I^R_{DE} and $I^{\mathrm{bare},R}_{DE}$ in Eqs. (62) and (63) can be obtained by expanding all integrands consistently in k/q (for similar calculations see e.g. [17, 33, 56–59]). We will only calculate the leading order terms which turn out to be of order $(k/q)^0$. The F_2 kernel contains a term scaling like q/k which diverges for $k/q \to 0$. Since this can combine with contributions of order k/q in other terms to give an overall term of order $(k/q)^0$, we generally need to keep all terms of order k/q. The orthogonality of Legendre polynomials will be needed for integrals $I^{\mathrm{bare},R}_{DE}$, but it is not used for integrals I^R_{DE} whose integrands are angle-independent on large scales.

In detail, to get the low-k limit of Eq. (62),

$$\lim_{k\to 0} I^R_{DE}(k) = \lim_{k\to 0} W_R(k) \int \frac{\mathrm{d}^3q}{(2\pi)^3} W_R(q) W_R(|\mathbf{k}-\mathbf{q}|) P_{\mathrm{mm}}(q) P_{\mathrm{mm}}(|\mathbf{k}-\mathbf{q}|) D(\mathbf{q},\mathbf{k}-\mathbf{q}) E(\mathbf{q},\mathbf{k}-\mathbf{q}), \tag{A1}$$

we expand all factors. We work with the cosine μ between $\mathbf{q}$ and $\mathbf{k}-\mathbf{q}$, see Eq. (1), and the cosine ν between $\mathbf{q}$ and $\mathbf{k}$

$$\nu \equiv \frac{\mathbf{q}\cdot\mathbf{k}}{qk}. \tag{A2}$$

From

$$\lim_{k\to 0} \frac{|\mathbf{k}-\mathbf{q}|}{q} = \lim_{k\to 0} \sqrt{1 - 2\frac{k\nu}{q} + \frac{k^2}{q^2}} = 1 - \frac{k\nu}{q} + \mathcal{O}\left(\frac{k^2}{q^2}\right) \tag{A3}$$

we get

$$\lim_{k\to 0} \mu = \lim_{k\to 0} \frac{-1 + \frac{k\nu}{q}}{1 - \frac{k\nu}{q}} = -1 + \mathcal{O}\left(\frac{k^2}{q^2}\right). \tag{A4}$$

Then

$$-\lim_{k\to 0} F_2^1(q,|\mathbf{k}-\mathbf{q}|)\mathsf{P}_1(\mu) = -\lim_{k\to 0} \frac{1}{2}\mu\left(\frac{|\mathbf{k}-\mathbf{q}|}{q} + \frac{q}{|\mathbf{k}-\mathbf{q}|}\right) = 1 + \mathcal{O}\left(\frac{k^2}{q^2}\right) \tag{A5}$$

and

$$\lim_{k\to 0} \mathsf{P}_2(\mu) = \lim_{k\to 0} \frac{3}{2}\left(\mu^2 - \frac{1}{3}\right) = 1 + \mathcal{O}\left(\frac{k^2}{q^2}\right). \tag{A6}$$

This can be summarized as

$$\lim_{k\to 0} D(\mathbf{q},\mathbf{k}-\mathbf{q}) = 1 + \mathcal{O}\left(\frac{k^2}{q^2}\right), \qquad D \in \left\{\mathsf{P}_0, -F_2^1\mathsf{P}_1, \mathsf{P}_2\right\}, \tag{A7}$$

which implies for the full F_2 kernel

$$\lim_{k\to 0} F_2(\mathbf{q},\mathbf{k}-\mathbf{q}) = 0 + \mathcal{O}\left(\frac{k^2}{q^2}\right). \tag{A8}$$

Some bare integrals (63) contain $E(\mathbf{q},-\mathbf{k})$. For $E = -F_2^1\mathsf{P}_1$ this becomes

$$\lim_{k\to 0}\left[-F_2^1(q,k)\mathsf{P}_1(-\nu)\right] = \frac{\nu}{2}\lim_{k\to 0}\left(\frac{q}{k} + \frac{k}{q}\right) = \frac{q\nu}{2k} + \mathcal{O}\left(\frac{k}{q}\right), \tag{A9}$$

which diverges for $k/q \to 0$. Note that for $E = \mathsf{P}_2$, $\mathsf{P}_2(-\nu) = \frac{3}{2}(\nu^2 - \frac{1}{3})$ is independent of k, so that

$$\lim_{k\to 0} F_2(\mathbf{q},-\mathbf{k}) = \frac{17}{21} - \frac{q\nu}{2k} + \frac{2}{7}\left(\nu^2 - \frac{1}{3}\right) + \mathcal{O}\left(\frac{k}{q}\right). \tag{A10}$$

We also expand

$$P(|\mathbf{k}-\mathbf{q}|) = P(q)\left[1 - \frac{k\nu}{q}\frac{\mathrm{d}\ln P(q)}{\mathrm{d}\ln q}\right] \tag{A11}$$

and

$$W_R(|\mathbf{k}-\mathbf{q}|) = W_R(q)\left[1 - \frac{k\nu}{q}\frac{\mathrm{d}\ln W_R(q)}{\mathrm{d}\ln q}\right]. \tag{A12}$$

For Gaussian smoothing this becomes

$$W_{R_G}(|\mathbf{k}-\mathbf{q}|) = W_{R_G}(q)\left[1 + \frac{k\nu}{q}(qR)^2\right]. \tag{A13}$$

Keeping only terms of order $(k/q)^0$ for the final results from now on we get from Eqs. (A7), (A8), (A11) and (A12)

$$\lim_{k\to 0} I^R_{DF_2}(k) = 0, \qquad D \in \left\{\mathsf{P}_0, -F_2^1\mathsf{P}_1, \mathsf{P}_2\right\}. \tag{A14}$$

and

$$\lim_{k\to 0} I^R_{DE}(k) = W_R(k)\tau_R^4, \qquad D \in \left\{\mathsf{P}_0, -F_2^1\mathsf{P}_1, \mathsf{P}_2\right\}, \;\; E \in \left\{\mathsf{P}_0, \mathsf{P}_2\right\}, \tag{A15}$$

where

$$\tau_R^4 \equiv \frac{1}{2\pi^2}\int \mathrm{d}q\, q^2 W_R^2(q) P^2_{\rm mm}(q). \tag{A16}$$

The limits of the bare integrals (63) for $E = F_2$ follow from Eqs. (A10) and (A12) by noting that only $\mathsf{P}_0^2(\nu)$ and $\mathsf{P}_1^2(\nu)$ survive the ν integration over the product of Eqs. (A10) and (A12) because of Legendre polynomial orthogonality:

$$\lim_{k\to 0} I^{{\rm bare},R_G}_{DF_2}(k) = \frac{1}{4}W_R(k)P_{\rm mm}(k)\left(\frac{68}{21}\sigma_R^2 - \frac{2}{3}\sigma^2_{R,W'_R}\right), \qquad D \in \left\{\mathsf{P}_0, -F_2^1\mathsf{P}_1, \mathsf{P}_2\right\}, \tag{A17}$$

where

$$\sigma_R^2 \equiv \frac{1}{2\pi^2}\int \mathrm{d}q\, q^2 W_R^2(q) P_{\rm mm}(q) \tag{A18}$$

and

$$\sigma^2_{R,W'_R} \equiv -\frac{1}{2\pi^2}\int \mathrm{d}q\, q^2 \frac{\mathrm{d}\ln W_R(q)}{\mathrm{d}\ln q} W_R^2(q) P_{\rm mm}(q). \tag{A19}$$

For Gaussian smoothing, $-\mathrm{d}\ln W_R(q)/\mathrm{d}\ln q = (qR)^2$, so no numerical derivatives are required to evaluate Eq. (A19). To gain intuition, Eq. (A19) can be integrated by parts to get $\sigma^2_{R,W'_R} = \frac{1}{2}\sigma^2_{R,P'}$, where

$$\sigma^2_{R,P'} \equiv \frac{1}{2\pi^2}\int \mathrm{d}q\, q^2 W_R^2(q) P_{\rm mm}(q) \frac{\mathrm{d}\ln q^3 P_{\rm mm}(q)}{\mathrm{d}\ln q}. \tag{A20}$$

Then, Eq. (A17) becomes for $D \in \{\mathsf{P}_0, -F_2^1\mathsf{P}_1, \mathsf{P}_2\}$

$$\lim_{k\to 0} I^{{\rm bare},R_G}_{DF_2}(k) = \frac{1}{4}W_R(k)P_{\rm mm}(k)\left(\frac{68}{21}\sigma_R^2 - \frac{1}{3}\sigma^2_{R,P'}\right) \tag{A21}$$

$$= \frac{1}{4}W_R(k)P_{\rm mm}(k)\,\frac{1}{2\pi^2}\int \mathrm{d}q\, q^2 W_R^2(q) P_{\rm mm}(q)\left(\frac{68}{21} - \frac{1}{3}\frac{\mathrm{d}\ln q^3 P_{\rm mm}(q)}{\mathrm{d}\ln q}\right). \tag{A22}$$

The integrand contains the linear response of the power spectrum to a long-wavelength overdensity (e.g. [17, 33, 58]). In contrast, the large-scale limit of the quantity proposed in [17] is (by construction) proportional to the response function on large scales rather than integrating over it.

For $E = \mathsf{P}_0$, only the $\mathsf{P}_0(\nu)$ part of Eq. (A12) survives the ν-integration so that

$$\lim_{k\to 0} I^{{\rm bare},R}_{D\mathsf{P}_0}(k) = W_R(k)P_{\rm mm}(k)\sigma_R^2, \qquad D \in \left\{\mathsf{P}_0, -F_2^1\mathsf{P}_1, \mathsf{P}_2\right\}. \tag{A23}$$

For $E = \mathsf{P}_2$, the integrand of Eq. (63) contains $\mathsf{P}_2(\nu)$ multiplied by $\mathsf{P}_0(\nu)$ and $\mathsf{P}_1(\nu)$ from Eq. (A12), which vanishes upon angular integration because of Legendre polynomial orthogonality, i.e.

$$\lim_{k\to 0} I^{{\rm bare},R}_{D\mathsf{P}_2}(k) = 0, \qquad D \in \left\{\mathsf{P}_0, -F_2^1\mathsf{P}_1, \mathsf{P}_2\right\}. \tag{A24}$$

Note that for our fiducial cosmology, at $z = 0.55$, for Gaussian smoothing with $R = 20h^{-1}$Mpc using the linear matter power spectrum in integrands, we get $\sigma_R^2 = 0.0215$, $\sigma^2_{R,P'} = 0.0464$ and $\tau_R^4 = 191.9h^{-3}\mathrm{Mpc}^3$. Using emulator instead of linear matter power gives $\sigma_R^2 = 0.0213$, $\sigma^2_{R,P'} = 0.0460$ and $\tau_R^4 = 189.2h^{-3}\mathrm{Mpc}^3$.

Appendix B: Analytical covariances

This appendix derives the covariance (78) between cross-spectra using leading-order perturbation theory. Since the cross-spectra are cubic in the density, their covariance depends on the density 3-point and 6-point functions,

$$\begin{aligned}\mathrm{cov}(\hat{P}_{D[\delta_a^R],\delta_b^R}(k),\hat{P}_{E[\delta_a^R],\delta_b^R}(k')) = \frac{1}{(4\pi L^3)^2}\int \mathrm{d}\Omega_{\hat{\mathbf{k}}}\int \mathrm{d}\Omega_{\hat{\mathbf{k}}'}\int\frac{\mathrm{d}^3q}{(2\pi)^3}\int\frac{\mathrm{d}^3q'}{(2\pi)^3}D(\mathbf{q},\mathbf{k}-\mathbf{q})E(\mathbf{q}',\mathbf{k}'-\mathbf{q}')\\ \times\Big[\langle\delta_a^R(\mathbf{q})\delta_a^R(\mathbf{k}-\mathbf{q})\delta_b^R(-\mathbf{k})\delta_a^R(\mathbf{q}')\delta_a^R(\mathbf{k}'-\mathbf{q}')\delta_b^R(-\mathbf{k}')\rangle\\ -\langle\delta_a^R(\mathbf{q})\delta_a^R(\mathbf{k}-\mathbf{q})\delta_b^R(-\mathbf{k})\rangle\langle\delta_a^R(\mathbf{q}')\delta_a^R(\mathbf{k}'-\mathbf{q}')\delta_b^R(-\mathbf{k}')\rangle\Big],\end{aligned} \tag{B1}$$

where the kernels D and E are symmetric in their arguments and $a,b\in\{\mathrm{h},\mathrm{m}\}$ denote halo or dark matter densities. The last term in the square brackets of Eq. (B1) can be neglected at leading order. Then, the fully disconnected part of the 6-point function results in[20]

$$\begin{aligned}\int\frac{\mathrm{d}^3q}{(2\pi)^3}\int\frac{\mathrm{d}^3q'}{(2\pi)^3}D(\mathbf{q},\mathbf{k}-\mathbf{q})E(\mathbf{q}',\mathbf{k}'-\mathbf{q}')\langle\delta_a^R(\mathbf{q})\delta_b^R(\mathbf{k}-\mathbf{q})\delta_b^R(-\mathbf{k})\delta_a^R(\mathbf{q}')\delta_b^R(\mathbf{k}'-\mathbf{q}')\delta_b^R(-\mathbf{k}')\rangle\\ = 2(2\pi)^6\delta_D^2(\mathbf{k}+\mathbf{k}')P_{bb}^R(k)I_{DE}^{P_{aa}^RP_{aa}^R}(k)\\ +4(2\pi)^3\delta_D(\mathbf{0})D(\mathbf{k}',\mathbf{k}-\mathbf{k}')E(\mathbf{k},\mathbf{k}'-\mathbf{k})P_{ab}^R(k)P_{ab}^R(k')P_{aa}^R(|\mathbf{k}-\mathbf{k}'|),\end{aligned} \tag{B2}$$

where

$$I_{DE}^{P_{aa}^RP_{aa}^R}(k)\equiv\int\frac{\mathrm{d}^3q}{(2\pi)^3}P_{aa}^R(q)P_{aa}^R(|\mathbf{k}-\mathbf{q}|)D(\mathbf{q},\mathbf{k}-\mathbf{q})E(\mathbf{q},\mathbf{k}-\mathbf{q}). \tag{B3}$$

All power spectra should be the full power spectra obtained by contracting two density fields, including shot noise. The covariance becomes

$$\begin{aligned}\mathrm{cov}(\hat{P}_{D[\delta_a^R],\delta_b^R}(k),\hat{P}_{E[\delta_a^R],\delta_b^R}(k')) = \frac{2L^6}{(4\pi L^3)^2}\delta_{k,k'}^{(\mathrm{K})}P_{bb}^R(k)I_{DE}^{P_{aa}^RP_{aa}^R}(k)\int\mathrm{d}\Omega_{\hat{\mathbf{k}}}\\ +\frac{4L^3}{(4\pi L^3)^2}P_{ab}^R(k)P_{ab}^R(k')\int\mathrm{d}\Omega_{\hat{\mathbf{k}}}\int\mathrm{d}\Omega_{\hat{\mathbf{k}}'}D(\mathbf{k}',\mathbf{k}-\mathbf{k}')E(\mathbf{k},\mathbf{k}'-\mathbf{k})P_{aa}^R(|\mathbf{k}-\mathbf{k}'|),\end{aligned} \tag{B4}$$

where we used $(2\pi)^3\delta_D(\mathbf{0})=L^3$. We neglect the $k\neq k'$ covariance in the main text because it is hard to test its validity with only ten realizations. This should be improved in future analyses. On a discrete grid, the angular integral is replaced by the sum over modes as in Eq. (15). For $k=k'$ this gives

$$\frac{1}{(4\pi)^2}\int\mathrm{d}\Omega_{\hat{\mathbf{k}}}\quad\to\quad\frac{1}{N_{\mathrm{modes}}^2(k)}\sum_{\mathbf{k},[k-\Delta k/2\le|\mathbf{k}|\le k+\Delta k/2]}=\frac{1}{N_{\mathrm{modes}}(k)}, \tag{B5}$$

which leads to Eq. (78) in the main text.

Appendix C: Reduced smoothing scale

Plots corresponding to Figures 1 and 3 for less aggressive $R_G=10h^{-1}$Mpc smoothing are shown in Figures 10 and 11. They are discussed in the main body of the paper.

[1] L. Anderson, É. Aubourg, S. Bailey, F. Beutler, V. Bhardwaj, M. Blanton, A. S. Bolton, J. Brinkmann, J. R. Brownstein, A. Burden, C.-H. Chuang, A. J. Cuesta, K. S. Dawson, D. J. Eisenstein, S. Escoffier, J. E. Gunn, H. Guo, S. Ho,

[20] To arrive at Eq. (B2) we used that any contraction of the 6-point function that connects two of the three leftmost densities with themselves, or two of the three rightmost densities with themselves, vanishes, because $\delta_{\mathbf{k}=\mathbf{0}}=0$ (i.e. $P(0)=0$). We also used $E(\mathbf{k}_1,\mathbf{k}_2)=E(-\mathbf{k}_1,-\mathbf{k}_2)$, which holds for all kernels considered in this paper.

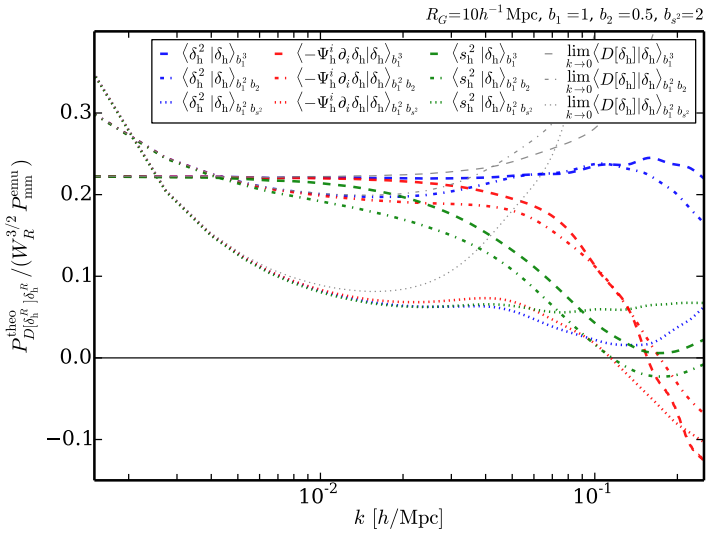

FIG. 10. Same as Fig. 1 but for $R_G = 10h^{-1}$Mpc.

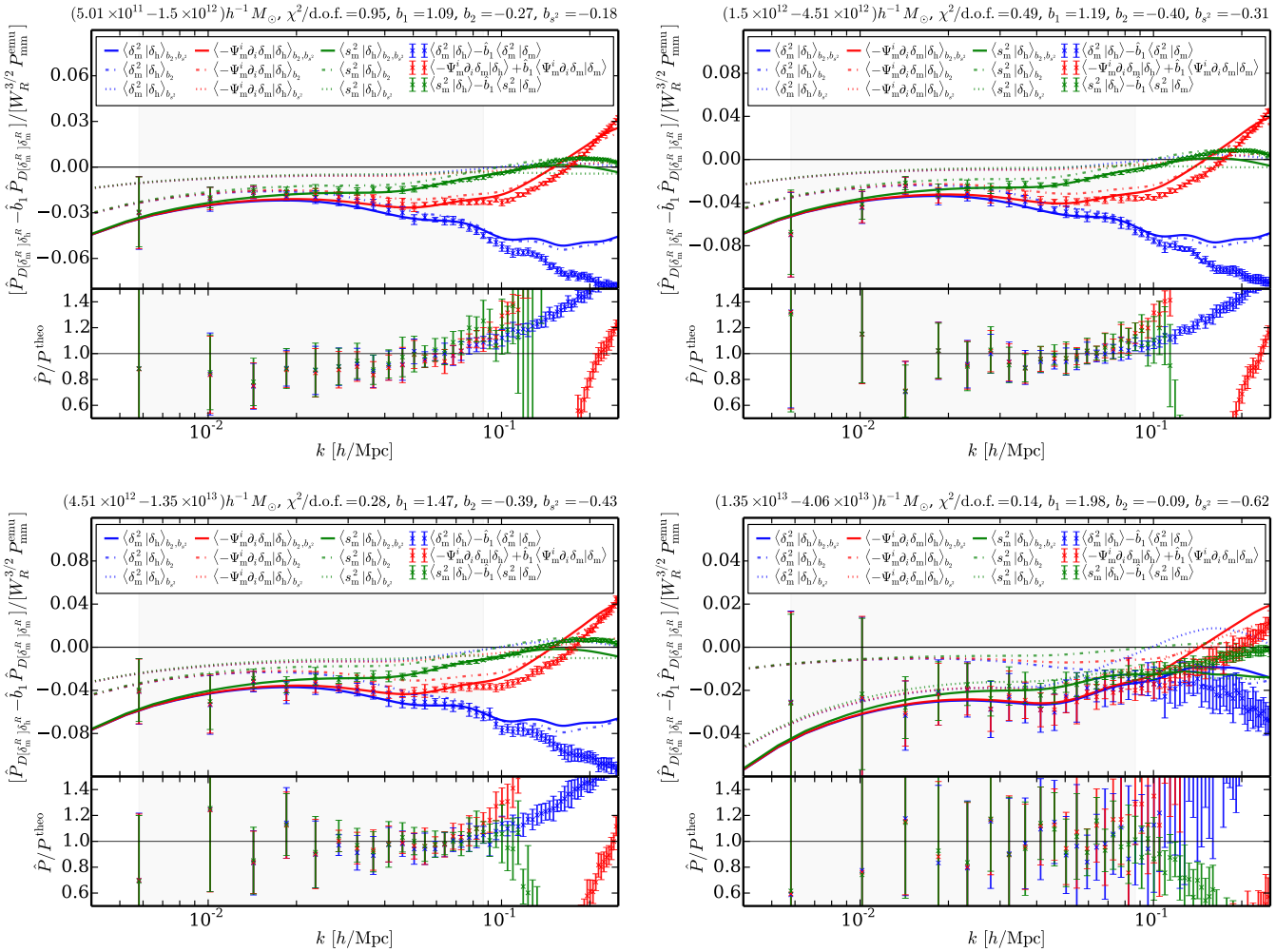

FIG. 11. Same as Fig. 3 but for $R_G = 10h^{-1}$Mpc smoothing.

K. Honscheid, C. Howlett, D. Kirkby, R. H. Lupton, M. Manera, C. Maraston, C. K. McBride, O. Mena, F. Montesano, R. C. Nichol, S. E. Nuza, M. D. Olmstead, N. Padmanabhan, N. Palanque-Delabrouille, J. Parejko, W. J. Percival, P. Petitjean, F. Prada, A. M. Price-Whelan, B. Reid, N. A. Roe, A. J. Ross, N. P. Ross, C. G. Sabiu, S. Saito, L. Samushia, A. G. Sánchez, D. J. Schlegel, D. P. Schneider, C. G. Scoccola, H.-J. Seo, R. A. Skibba, M. A. Strauss, M. E. C. Swanson, D. Thomas, J. L. Tinker, R. Tojeiro, M. V. Magaña, L. Verde, D. A. Wake, B. A. Weaver, D. H. Weinberg, M. White, X. Xu, C. Yèche, I. Zehavi, and G.-B. Zhao, Mon. Not. R. Astron. Soc. **441**, 24 (2014), arXiv:1312.4877.

[2] R. Scoccimarro, H. A. Feldman, J. N. Fry, and J. A. Frieman, Astrophys. J. **546**, 652 (2001), astro-ph/0004087.

[3] H. A. Feldman, J. A. Frieman, J. N. Fry, and R. Scoccimarro, Physical Review Letters **86**, 1434 (2001), astro-ph/0010205.

[4] L. Verde, A. F. Heavens, W. J. Percival, S. Matarrese, C. M. Baugh, J. Bland-Hawthorn, T. Bridges, R. Cannon, S. Cole, M. Colless, C. Collins, W. Couch, G. Dalton, R. De Propris, S. P. Driver, G. Efstathiou, R. S. Ellis, C. S. Frenk, K. Glazebrook, C. Jackson, O. Lahav, I. Lewis, S. Lumsden, S. Maddox, D. Madgwick, P. Norberg, J. A. Peacock, B. A. Peterson, W. Sutherland, and K. Taylor, Mon. Not. R. Astron. Soc. **335**, 432 (2002), astro-ph/0112161.

[5] Y. P. Jing and G. Börner, Astrophys. J. **607**, 140 (2004), astro-ph/0311585.

[6] Y. Wang, X. Yang, H. J. Mo, F. C. van den Bosch, and Y. Chu, Mon. Not. R. Astron. Soc. **353**, 287 (2004), astro-ph/0404143.

[7] F. A. Marín, C. Blake, G. B. Poole, C. K. McBride, S. Brough, M. Colless, C. Contreras, W. Couch, D. J. Croton, S. Croom, T. Davis, M. J. Drinkwater, K. Forster, D. Gilbank, M. Gladders, K. Glazebrook, B. Jelliffe, R. J. Jurek, I.-h. Li, B. Madore, D. C. Martin, K. Pimbblet, M. Pracy, R. Sharp, E. Wisnioski, D. Woods, T. K. Wyder, and H. K. C. Yee, Mon. Not. R. Astron. Soc. **432**, 2654 (2013), arXiv:1303.6644 [astro-ph.CO].

[8] C. K. McBride, A. J. Connolly, J. P. Gardner, R. Scranton, R. Scoccimarro, A. A. Berlind, F. Marín, and D. P. Schneider, Astrophys. J. **739**, 85 (2011), arXiv:1012.3462 [astro-ph.CO].

[9] C. K. McBride, A. J. Connolly, J. P. Gardner, R. Scranton, J. A. Newman, R. Scoccimarro, I. Zehavi, and D. P. Schneider, Astrophys. J. **726**, 13 (2011), arXiv:1007.2414 [astro-ph.CO].

[10] H. Gil-Marín, J. Noreña, L. Verde, W. J. Percival, C. Wagner, M. Manera, and D. P. Schneider, ArXiv e-prints (2014), arXiv:1407.5668.

[11] H. Gil-Marín, L. Verde, J. Noreña, A. J. Cuesta, L. Samushia, W. J. Percival, C. Wagner, M. Manera, and D. P. Schneider, ArXiv e-prints (2014), arXiv:1408.0027.

[12] H. Gil-Marín, C. Wagner, J. Noreña, L. Verde, and W. Percival, ArXiv e-prints (2014), arXiv:1407.1836.

[13] F. Bernardeau, Astronomy and Astrophysics **312**, 11 (1996), astro-ph/9602072.

[14] J. Bel and C. Marinoni, Mon. Not. R. Astron. Soc. **424**, 971 (2012), arXiv:1205.3200 [astro-ph.CO].

[15] K. Hoffmann, J. Bel, E. Gaztanaga, M. Crocce, P. Fosalba, and F. J. Castander, ArXiv e-prints (2014), arXiv:1403.1259 [astro-ph.CO].

[16] J. E. Pollack, R. E. Smith, and C. Porciani, Mon. Not. R. Astron. Soc. **440**, 555 (2014), arXiv:1309.0504 [astro-ph.CO].

[17] C.-T. Chiang, C. Wagner, F. Schmidt, and E. Komatsu, JCAP **5**, 048 (2014), arXiv:1403.3411.

[18] J. R. Fergusson, D. M. Regan, and E. P. S. Shellard, Phys. Rev. D **86**, 063511 (2012), arXiv:1008.1730 [astro-ph.CO].

[19] D. M. Regan, M. M. Schmittfull, E. P. S. Shellard, and J. R. Fergusson, Phys. Rev. D **86**, 123524 (2012), arXiv:1108.3813 [astro-ph.CO].

[20] M. M. Schmittfull, D. M. Regan, and E. P. S. Shellard, Phys. Rev. D **88**, 063512 (2013), arXiv:1207.5678 [astro-ph.CO].

[21] E. Komatsu, D. N. Spergel, and B. D. Wandelt, Astrophys. J. **634**, 14 (2005), astro-ph/0305189.

[22] A. Lewis, A. Challinor, and D. Hanson, JCAP **3**, 018 (2011), arXiv:1101.2234 [astro-ph.CO].

[23] Planck Collaboration, P. A. R. Ade, N. Aghanim, C. Armitage-Caplan, M. Arnaud, M. Ashdown, F. Atrio-Barandela, J. Aumont, C. Baccigalupi, A. J. Banday, and et al., ArXiv e-prints (2013), arXiv:1303.5079 [astro-ph.CO].

[24] J. R. Fergusson, M. Liguori, and E. P. S. Shellard, Phys. Rev. D **82**, 023502 (2010), arXiv:0912.5516 [astro-ph.CO].

[25] J. R. Fergusson, M. Liguori, and E. P. S. Shellard, JCAP **12**, 032 (2012), arXiv:1006.1642 [astro-ph.CO].

[26] Planck Collaboration, P. A. R. Ade, N. Aghanim, C. Armitage-Caplan, M. Arnaud, M. Ashdown, F. Atrio-Barandela, J. Aumont, C. Baccigalupi, A. J. Banday, and et al., ArXiv e-prints (2013), arXiv:1303.5084 [astro-ph.CO].

[27] M. Zaldarriaga and U. Seljak, Phys. Rev. D **59**, 123507 (1999), astro-ph/9810257.

[28] T. Okamoto and W. Hu, Phys. Rev. D **67**, 083002 (2003), astro-ph/0301031.

[29] D. Hanson, A. Challinor, G. Efstathiou, and P. Bielewicz, Phys. Rev. D **83**, 043005 (2011), arXiv:1008.4403 [astro-ph.CO].

[30] T. Baldauf, U. Seljak, V. Desjacques, and P. McDonald, Phys. Rev. D **86**, 083540 (2012), arXiv:1201.4827 [astro-ph.CO].

[31] F. Bernardeau, S. Colombi, E. Gaztañaga, and R. Scoccimarro, Physics Reports **367**, 1 (2002), astro-ph/0112551.

[32] F. R. Bouchet, R. Juszkiewicz, S. Colombi, and R. Pellat, APJL **394**, L5 (1992).

[33] B. D. Sherwin and M. Zaldarriaga, Phys. Rev. D **85**, 103523 (2012), arXiv:1202.3998 [astro-ph.CO].

[34] P. McDonald and A. Roy, JCAP **8**, 020 (2009), arXiv:0902.0991 [astro-ph.CO].

[35] K. C. Chan, R. Scoccimarro, and R. K. Sheth, Phys. Rev. D **85**, 083509 (2012), arXiv:1201.3614 [astro-ph.CO].

[36] R. Scoccimarro and H. M. P. Couchman, Mon. Not. R. Astron. Soc. **325**, 1312 (2001), astro-ph/0009427.

[37] H. Gil-Marín, C. Wagner, F. Fragkoudi, R. Jimenez, and L. Verde, JCAP **2**, 047 (2012), arXiv:1111.4477 [astro-ph.CO].

[38] T. Baldauf, L. Mercolli, M. Mirbabayi, and E. Pajer, ArXiv e-prints (2014), arXiv:1406.4135.

[39] R. E. Angulo, S. Foreman, M. Schmittfull, and L. Senatore, ArXiv e-prints (2014), arXiv:1406.4143.

[40] M. C. Neyrinck, I. Szapudi, and A. S. Szalay, APJL **698**, L90 (2009), arXiv:0903.4693 [astro-ph.CO].

[41] F. Simpson, J. B. James, A. F. Heavens, and C. Heymans, Physical Review Letters **107**, 271301 (2011), arXiv:1107.5169 [astro-ph.CO].

[42] T. Baldauf, V. Desjacques, and U. Seljak, ArXiv e-prints (2014), arXiv:1405.5885.

[43] M. Crocce, R. Scoccimarro, and F. Bernardeau, Mon. Not. R. Astron. Soc. **427**, 2537 (2012), arXiv:1207.1465 [astro-ph.CO].
[44] A. Taruya, F. Bernardeau, T. Nishimichi, and S. Codis, Phys. Rev. D **86**, 103528 (2012), arXiv:1208.1191 [astro-ph.CO].
[45] K. Heitmann, E. Lawrence, J. Kwan, S. Habib, and D. Higdon, Astrophys. J. **780**, 111 (2014), arXiv:1304.7849 [astro-ph.CO].
[46] K. Heitmann, M. White, C. Wagner, S. Habib, and D. Higdon, Astrophys. J. **715**, 104 (2010), arXiv:0812.1052.
[47] K. Heitmann, D. Higdon, M. White, S. Habib, B. J. Williams, E. Lawrence, and C. Wagner, Astrophys. J. **705**, 156 (2009), arXiv:0902.0429 [astro-ph.CO].
[48] E. Lawrence, K. Heitmann, M. White, D. Higdon, C. Wagner, S. Habib, and B. Williams, Astrophys. J. **713**, 1322 (2010), arXiv:0912.4490 [astro-ph.CO].
[49] D. Jeong, PhD dissertation, `http://www.personal.psu.edu/duj13/dissertation/djeong_diss.pdf` (2010).
[50] T. Baldauf, U. Seljak, R. E. Smith, N. Hamaus, and V. Desjacques, Phys. Rev. D **88**, 083507 (2013), arXiv:1305.2917 [astro-ph.CO].
[51] B. A. Reid, H.-J. Seo, A. Leauthaud, J. L. Tinker, and M. White, Mon. Not. R. Astron. Soc. **444**, 476 (2014), arXiv:1404.3742.
[52] M. White, B. Reid, C.-H. Chuang, J. L. Tinker, C. K. McBride, F. Prada, and L. Samushia, ArXiv e-prints (2014), arXiv:1408.5435.
[53] M. White, Astrophys. J. Supp. **143**, 241 (2002), astro-ph/0207185.
[54] A. Lewis, A. Challinor, and A. Lasenby, Astrophys. J. **538**, 473 (2000), astro-ph/9911177.
[55] D. Foreman-Mackey, D. W. Hogg, D. Lang, and J. Goodman, Publications of the Astronomical Society of the Pacific **125**, 306 (2013), arXiv:1202.3665 [astro-ph.IM].
[56] P. Valageas, Phys. Rev. D **89**, 123522 (2014), arXiv:1311.4286.
[57] A. Kehagias, H. Perrier, and A. Riotto, ArXiv e-prints (2013), arXiv:1311.5524 [astro-ph.CO].
[58] Y. Li, W. Hu, and M. Takada, Phys. Rev. D **89**, 083519 (2014), arXiv:1401.0385.
[59] D. G. Figueroa, E. Sefusatti, A. Riotto, and F. Vernizzi, JCAP **8**, 036 (2012), arXiv:1205.2015 [astro-ph.CO].