

# An Assessment of the Information Content of the Power Spectrum and Bispectrum

Kwan Chuen Chan\* and Linda Blot

*Institute of Space Sciences, IEEC-CSIC, Campus UAB,  
Carrer de Can Magrans, s/n, 08193 Bellaterra, Barcelona, Spain*

(Dated: October 24, 2016)

We study the covariance matrix of the power spectrum and bispectrum for dark matter and halos both numerically and analytically. Using a large suite of simulations from the DEUS-PUR project, we find that the non-Gaussian contribution to the covariance of the power spectrum and bispectrum is significant for both dark matter and halos already at mildly non-linear scales. We compute the leading disconnected non-Gaussian correction to the matter bispectrum covariance using perturbation theory, and find that the predictions agree with the covariance of the equilateral configurations up to  $k \sim 0.16 \text{ Mpc}^{-1} h$  at  $z = 0$ . The shot noise contributions to the covariance of the halo power spectrum and bispectrum are computed. When the ensemble averaged number density is used, the Poisson model covariances are in decent agreement with the measurements. When the number density is estimated and subtracted from each realization, the covariances are significantly reduced and get close to the Gaussian ones for both halo power spectrum and bispectrum. The information content of the power spectrum and bispectrum is assessed using the signal-to-noise ratio, S/N. Similar to what was found for the dark matter power spectrum, the S/N of the halo power spectrum also levels off in the mildly nonlinear regime,  $k \sim 0.1\text{--}0.2 \text{ Mpc}^{-1} h$ , depending on the number density of the samples. In the nonlinear regime the S/N of the matter and halo bispectrum increases but much slower than the Gaussian results suggest. We find that both the S/N for power spectrum and bispectrum are overestimated by the Gaussian covariances, but the problem being much more serious for the bispectrum. While the Gaussian covariance suggests that the S/N of the matter bispectrum should surpass that of the matter power spectrum at  $k \sim 0.2 \text{ Mpc}^{-1} h$ , the full non-Gaussian results show that this only occurs at  $k \sim 1 \text{ Mpc}^{-1} h$ . Because the bispectrum is affected strongly by nonlinearity and shot noise, inclusion of the bispectrum only adds modest amount of S/N compared to that of the power spectrum.

## I. INTRODUCTION

Current large scale structure surveys such as those aiming at measuring the baryonic acoustic oscillations are giving important insights into the physics of the universe [1–6]. Future surveys such as Euclid [7] and LSST [8] will measure the large scale structure over a large volume and in deep redshift ranges with unprecedented precision. To correctly interpret the data, the challenge is not only to accurately model the observables, such as the power spectrum and bispectrum, but the covariance matrix of these quantities must also be known with sufficient precision.

The covariance is often measured by running large numbers of mocks, where the survey is modelled starting from a large scale structure simulation. In this approach the mocks not only can take into account the intrinsic correlation, but also the survey geometry and various systematic effects. To get good control of the covariance, often hundreds or even thousands of simulations are required. As it is computationally expensive to run the full  $N$ -body simulations, various cheaper approximate methods are often used (see Ref. [9] and references therein for a review of some of the methods that have been proposed).

However, to reach a better understanding and modelling of the covariance it is easier to start with the relatively simpler case of dark matter and halos in a periodic

box. This approach has been used to study the covariance of the power spectrum for dark matter [10–15] and halos [12]. For the bispectrum PTHalos [16] has been used to investigate the covariance systematically [17, 18].

Analytical methods also prove fruitful, as they give interesting insights in what are the contributions that are relevant for the estimation of the covariance. The dark matter power spectrum has been modelled using perturbation theory [11, 19, 20] and halo model [21, 22]. The Poisson model has been invoked to model the shot noise contribution to the covariance [12, 23]. In particular, by combining the analytical and numerical approaches, it has been realized that beat coupling or super-sample covariance can be a significant contribution to the covariance in real surveys [24–28].

In this paper we study the power spectrum and bispectrum covariance numerically and analytically. Taking advantage of the large suite of simulations available in the DEUS-PUR project, we study the covariance of the dark matter and halo power spectrum and bispectrum. It is worth stressing that this is the first systematic study of the bispectrum covariance using such large number of  $N$ -body simulations. We also model the covariance and compare the predictions with the numerical results.

As the power spectrum has been well explored and the bispectrum becomes the next frontier in large scale structure, it is crucial to address how much information one can gain by going beyond the 2-point level. Previous work [29] suggested that there is substantial information content in the high-point statistics such as the bispec-

---

\*Electronic address: chan@ice.cat

trum based on the Gaussian covariance approximation. Armed with the accurate covariance measured from a large suite of simulations, we assess the information loss in the power spectrum and bispectrum due to the correlations that arise in the nonlinear regime.

This paper is organized as follows. In Sec. II, after reviewing the basic theory of the power spectrum covariance for dark matter and halos, we compare the model prediction against the numerical covariance. Sec. III is devoted to the bispectrum covariance. We shall first lay down the theory of the dark matter and halo bispectrum covariance. We then present the numerical bispectrum covariance and compare it with the theory predictions. The information content of the power spectrum and the bispectrum is assessed by means of the signal-to-noise ratio in Sec. IV. We conclude in Sec. V. We check the probability distribution of the bispectrum estimator in Appendix A. In Appendix B, we show the derivation of some of formulas used in the main text. We derive the shot noise contribution to the halo power spectrum and bispectrum covariance using the Poisson model in Appendix C. In Appendix D, we check the dependence of the signal-to-noise ratio on the binning width.

## II. COVARIANCE OF POWER SPECTRUM

The covariance of the dark matter power spectrum has been studied quite extensively both numerically [10, 11, 13, 15] and theoretically [10, 11, 19, 20]. On the other hand the covariance of the halo power spectrum is relatively less explored, but see Refs. [12, 30]. Although the focus of this section is on the covariance of the halo power spectrum, we will also present the results for dark matter for comparison. We will first review the basic theory of the power spectrum covariance, which will pave the way for the bispectrum covariance that we will discuss later on.

### A. Theory of the power spectrum covariance

Here we first review the theory of the covariance matrix of the matter power spectrum, and then we discuss the covariance of the halo power spectrum.

Suppose that the Fourier modes of the density contrast  $\delta$  are binned into bands of width  $\Delta k$  in Fourier space. The power spectrum of  $\delta$ ,  $P$ , is defined as

$$\langle \delta(\mathbf{k})\delta(\mathbf{k}') \rangle = P(k)\delta_{\text{D}}(\mathbf{k} + \mathbf{k}') \quad (1)$$

where  $\delta_{\text{D}}$  is the Dirac delta function. From the definition, we can construct a power spectrum estimator  $\hat{P}$  as (e.g. [11, 31])

$$\hat{P}(k) = k_{\text{F}}^3 \int_k \frac{d^3 p}{V_{\text{s}}(k)} \delta(\mathbf{p})\delta(-\mathbf{p}), \quad (2)$$

where  $k_{\text{F}}$  is the fundamental mode of the box. Note that the integral is done over all the modes that fall into the

band of width  $[k - \Delta k/2, k + \Delta k/2)$ .  $V_{\text{s}}$  is the volume of the spherical shell

$$V_{\text{s}}(k) = \int_k d^3 p = 4\pi k^2 \Delta k + \frac{\pi}{3} \Delta k^3. \quad (3)$$

The covariance matrix of  $\hat{P}$  is defined as

$$\begin{aligned} C^P(k, k') &\equiv \text{cov}(\hat{P}(k), \hat{P}(k')) \\ &= \langle [\hat{P}(k) - \langle \hat{P}(k) \rangle][\hat{P}(k') - \langle \hat{P}(k') \rangle] \rangle \\ &= \langle \hat{P}(k)\hat{P}(k') \rangle - \langle \hat{P}(k) \rangle \langle \hat{P}(k') \rangle. \end{aligned} \quad (4)$$

#### 1. Dark matter

The covariance matrix of the matter power spectrum was first investigated using perturbation theory in Ref. [11], and it has been extended to include loop corrections in Refs. [19, 20]. Furthermore, it has also been shown that the large scale mode can modulate the small scale modes to cause the so-called beat-coupling effect or supersample covariance [24–28]. This can be a significant source of covariance at small scales in simulations of large box size or real surveys. For small periodic box simulations, this effect is negligible due to the lack of large scale modes. We will not consider beat coupling in details in this work, except that we will compare the results obtained from simulations of different box sizes. Here we will limit ourselves to the theory laid down in Ref. [11] as it is not the focus of this paper to model the matter power spectrum covariance as accurately as possible.

Plugging Eq. 2 into Eq. 4, for Gaussian  $\delta$ , we get the Gaussian covariance of the power spectrum estimator [31]

$$C_{\text{G}}^P(k, k') = \frac{2k_{\text{F}}^3}{V_{\text{s}}(k)} P^2(k) \delta_{k, k'}, \quad (5)$$

where  $\delta_{k, k'}$  is the Kronecker delta function that ensures that the Gaussian covariance is diagonal. For Gaussian field, the power spectrum  $P$  in Eq. 5 should be the linear one. As we will include non-Gaussian corrections below, one of the power spectrum should be the 1-loop one so that it is of the same order as the non-Gaussian results. However, since the 1-loop power spectrum over-predicts the numerical power spectrum already at  $k \sim 0.1 \text{ Mpc}^{-1} h$  at  $z = 0$ , we will use the nonlinear power spectrum measured from simulations in place of the 1-loop result.

The non-Gaussian contribution to the power spectrum covariance comes from the trispectrum [10, 11]

$$\begin{aligned} C_{\text{NG}}^P(k, k') &= k_{\text{F}}^3 \int_k \frac{d^3 p}{V_{\text{s}}(k)} \int_{k'} \frac{d^3 p'}{V_{\text{s}}(k')} T(\mathbf{p}, -\mathbf{p}, \mathbf{p}', -\mathbf{p}'). \end{aligned} \quad (6)$$

Note that both the Gaussian and non-Gaussian covariance are inversely proportional to the volume of the box (through  $k_{\text{F}}^3$ ). We will see that the same is true for the

bispectrum covariance. We can trace back the factor  $k_F^3$  to the Dirac delta function in the definition of the polyspectrum, which arises from the statistical translational invariance of the field. Because of this invariance, the amount of statistics is simply proportional to the volume. On the other hand, when a window function is imposed, the statistical translational invariance is broken. Indeed the beat-coupling term scales differently with volume [27].

Nonlinearity induces mode coupling and hence non-Gaussianity. The tree-level dark matter trispectrum  $T$  has two distinct contributions,  $T_1$  and  $T_2$  [11, 32]

$$T_1(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) = 6F_3(\mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4)P_L(k_2)P_L(k_3)P_L(k_4) + 3 \text{cyc.}, \quad (7)$$

$$T_2(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) = [4F_2(-\mathbf{k}_3, \mathbf{k}_{23})F_2(\mathbf{k}_4, \mathbf{k}_{23})P_L(k_{23})P_L(k_3)P_L(k_4) + (\mathbf{k}_1 \leftrightarrow \mathbf{k}_2)] + 5 \text{cyc.}, \quad (8)$$

where *cyc.* denotes cyclic permutations.  $P_L$  is the linear power spectrum, and  $F_2$  and  $F_3$  are the coupling kernels in standard perturbation theory, see [33, 34].

The  $T_1$  and  $T_2$  contributions to the covariance matrix are then given by

$$\begin{aligned} & T_1(\mathbf{k}_1, -\mathbf{k}_1, \mathbf{k}_2, -\mathbf{k}_2) \\ &= 12F_3(\mathbf{k}_1, \mathbf{k}_2, -\mathbf{k}_2)P_L^2(k_2)P_L(k_1) + (\mathbf{k}_1 \leftrightarrow \mathbf{k}_2), \\ & T_2(\mathbf{k}_1, -\mathbf{k}_1, \mathbf{k}_2, -\mathbf{k}_2) \\ &= 4P_L(|\mathbf{k}_1 - \mathbf{k}_2|)[F_2(-\mathbf{k}_1, \mathbf{k}_1 - \mathbf{k}_2)P_L(k_1) \\ &+ F_2(-\mathbf{k}_1, \mathbf{k}_1 - \mathbf{k}_2)P_L(k_2)]^2 + (\mathbf{k}_2 \rightarrow -\mathbf{k}_2). \end{aligned} \quad (9)$$

The resultant covariance matrix contribution reads [11]

$$\begin{aligned} \bar{T}(k, k') &= \int_k \frac{d^3p}{V_s(k)} \int_{k'} \frac{d^3p'}{V_s(k')} \\ &\times \left\{ 12F_3(\mathbf{p}, \mathbf{p}', -\mathbf{p}')P_L(p)P_L^2(p') + (\mathbf{p} \leftrightarrow \mathbf{p}') \right. \\ &\left. + 8P_L(|\mathbf{p} - \mathbf{p}'|)[F_2(-\mathbf{p}, \mathbf{p} - \mathbf{p}')P_L(p) + (\mathbf{p} \leftrightarrow \mathbf{p}')]^2 \right\}. \end{aligned} \quad (10)$$

The contribution to the covariance of the power spectrum can be represented graphically, and the results are shown in Fig. 1. We will apply similar techniques to the bispectrum shortly, and it will be more useful as there are more contributions to the bispectrum covariance. In each diagram two black dots on both sides represent the two  $\delta$ 's in the power spectrum estimator. The legs branching from each dot represents the perturbation theory kernels. We use the wavy line to represent the linear power spectrum, while the additional dot on top means that the 1-loop power spectrum should be used instead. The term in the top-left corner is the Gaussian term, while the others are the non-Gaussian terms. The term  $C_{F_3}$  gives the  $T_1$  contribution, and both  $C_{F_2I}$  and  $C_{F_2II}$  combine to give  $T_2$ .

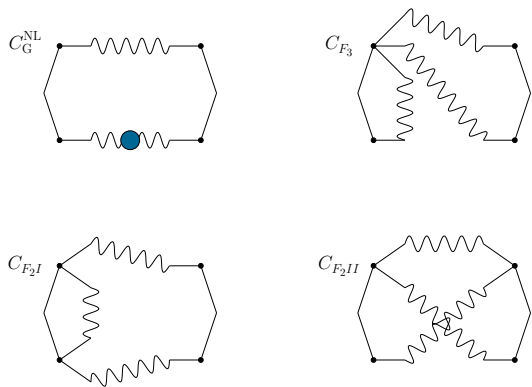


FIG. 1: A diagrammatic representation of the leading contributions to the covariance of the matter power spectrum.  $C_G^{\text{NL}}$  is the Gaussian term, while the rest are non-Gaussian contributions. The set of black dots on the left and right of each diagram represent the two  $\delta$ 's in each power spectrum estimator. The legs branching from each dot represent the perturbation theory kernel,  $F_1$ ,  $F_2$ , and  $F_3$  respectively. Each wavy line represents the linear power spectrum. In  $C_G^{\text{NL}}$ , the filled circle indicates that the linear power spectrum is replaced by the 1-loop power spectrum.

## 2. Halos

We now turn to the halo power spectrum covariance. In this case, besides the complications due to dark matter nonlinearity and halo biasing, the discrete nature of halos contributes further to stochastic fluctuations. In fact, shot noise is the major source of halo covariance as we will see below. In Appendix C 1 we derive the Poisson shot noise contribution to the covariance matrix of the power spectrum using the Poisson model. We refer the readers to the Appendix C 1 for details on the derivation. Here we summarize the key results.

There are both Gaussian and non-Gaussian contributions to the covariance due to Poisson shot noise. The Gaussian shot noise contribution can be combined with the smooth Gaussian covariance Eq. 5 to be written in a compact form [12, 31]

$$C_G^P(k, k') = \frac{2k_F^3}{V_s(k)} \left[ P_h(k) + \frac{1}{\bar{n}} \right]^2 \delta_{k, k'}, \quad (11)$$

where  $P_h$  is the smooth halo power spectrum and  $\bar{n} = (2\pi)^3 \bar{n}$  with  $\bar{n}$  being the mean number density of halos.

<sup>1</sup> The non-Gaussian shot noise contribution is given by

$$\begin{aligned} C_{\text{NG}}^P(k, k') &= k_F^3 \left[ \frac{1}{\bar{n}^3} + \frac{2}{\bar{n}^2} (P_h(k) + P_h(k')) \right. \\ &\left. + \frac{2}{\bar{n}^2} \int_k \frac{d^3p}{V_s(k)} \int_{k'} \frac{d^3p'}{V_s(k')} P_h(|\mathbf{p} + \mathbf{p}'|) \right] \end{aligned}$$

<sup>1</sup> The presence of the  $(2\pi)^3$  factor is due to the Fourier convention used in this paper. We can convert the formula to the perhaps more popular convention by replacing  $\bar{n}$  by  $\bar{n}$ .

$$+ \dots \quad (12)$$

where the dots denote the contribution associated with the halo bispectrum and also the smooth trispectrum contributions. See Eq. C19 for the full expressions. For the halo power spectrum we will not include the shot noise contribution associated with the bispectrum for simplicity because we find that the prediction without bispectrum works reasonably well. A simple estimate suggests that this is small compared to the dominant Gaussian term but not negligible.

In Fig. 2, we plot the components of the Gaussian and non-Gaussian contributions to the diagonal of the covariance. Note that the covariance depends on the volume of the simulation and the bin width  $\Delta k$ . For the theoretical computations, unless otherwise stated, we use box size  $656.25 \text{ Mpc } h^{-1}$ , which corresponds to the Small or Hires set shown in Table I. For the power spectrum, we use the binning width of  $9.6 \times 10^{-3} \text{ Mpc}^{-1} h$ , which is equal to the fundamental mode of the Small set. We show the results for two representative halo groups, which correspond to the Large group 4 at  $z = 0.5$  and the Hires group 2 at  $z = 0$  shown in Table II. For the low number density group in the range  $k \lesssim 0.2 \text{ Mpc}^{-1} h$ , the two Gaussian terms associated with shot noise dominate, while for higher  $k$ , the non-Gaussian term  $1/\bar{n}^3$  is the only significant term. The smooth Gaussian contribution is a few orders of magnitude smaller. On the other hand, for the low bias and high number density sample, the Gaussian terms are dominant up to  $k \sim 1 \text{ Mpc}^{-1} h$ . The smooth Gaussian term is important in this case although it is still smaller than the Gaussian term  $\propto P/\bar{n}$  by more than one order of magnitude.

## B. Numerical results

In this paper we use the simulations from the DEUS-PUR project [15, 35]. We will consider three sets of simulations labelled as Large, Small, and Hires respectively. The detailed properties of these simulations are shown in Table I. A flat  $\Lambda$ CDM model with the WMAP7 cosmological parameters [36] is adopted for these simulations. In particular,  $h = 0.72$ ,  $\Omega_m = 0.257$ ,  $n_s = 0.963$ , and  $\sigma_8 = 0.801$ . Zel'dovich approximation is used to generate the Gaussian initial conditions at  $z_i = 105$  for the Large and Small sets, and  $z_i = 190$  for the Hires set. The transfer function is computed with CAMB [37]. The simulations are evolved using the adaptive mesh refinement solver RAMSES [38]. We will consider simulation snapshots at  $z = 1, 0.5$ , and  $0$  respectively. The Large and Small sets have the same mass resolution although the box size of the Large set is twice that of the Small set. The Hires has higher mass resolution than the Small set while their box sizes are the same. For more details on the descriptions of the simulations, see Ref. [15].

The halos used in this work are obtained using the friends-of-friends algorithm with linking length set to  $0.2$

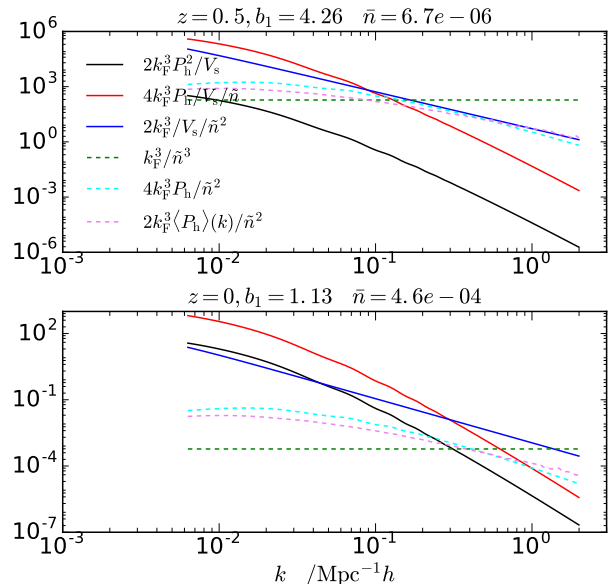


FIG. 2: Various contributions to the diagonal elements of the covariance of the halo power spectrum for two selected halo groups. The terms in the Gaussian covariance are plotted separately (solid lines). The three non-Gaussian terms in Eq. 12 are also plotted (dashed lines).

times the mean inter-particle separation. Only halos with at least 100 particles are used. We divide the halos into four mass groups. The details of these mass groups are shown in Table II. The simulations of the Large and Small sets only have the highest mass group, Gr. 4, while the Hires set has groups 1, 2, 3, and 4. Moreover, the results at  $z = 1, 0.5$  and  $0$  are available for the Large and Small sets, while only  $z = 0$  ones are available for the Hires set.

### 1. Dark matter

The covariance of the dark matter power spectrum has been shown in Ref. [15] using the same dataset. Here we only show the dark matter power spectrum results for completeness. We will also compare the results between the Small and Large set, while in Ref. [15] only results from the Small set were shown.

As we mentioned the covariance depends on the volume and the binning, thus comparing simulations of different box sizes is non-trivial. For the power spectrum measurements, we choose the same band width  $\Delta k = 9.6 \times 10^{-3} \text{ Mpc}^{-1} h$  for all the simulations. This means that even the Gaussian parts are different for the simulations of different box sizes. Nonetheless we will often show the results as ratios with respect to the respective Gaussian covariance so that the results from different boxes agree with each other at least at large scales.

TABLE I: Details of the simulations.

Box label	Box size (Mpc $h^{-1}$ )	Number of particles	redshift snapshots	Number of realizations
Large	1312.5	512 <sup>3</sup>	1, 0.5, 0	512
Small	656.25	256 <sup>3</sup>	1, 0.5, 0	4096
Hires	656.25	1024 <sup>3</sup>	1, 0.5, 0	96

TABLE II: Properties of the halo groups.

Box label	Mass group	redshift snapshots	mass range ( $10^{12} M_{\odot} h^{-1}$ )	linear bias	number density (Mpc $h^{-1}$ ) <sup>-3</sup>
Large, Small	4	1	> 120	6.44	$1.94 \times 10^{-6}$
Large, Small	4	0.5	> 120	4.26	$6.74 \times 10^{-6}$
Large, Small	4	0	> 120	2.90	$1.57 \times 10^{-5}$
Hires	1	0	1.88 – 5.63	0.94	$9.77 \times 10^{-4}$
Hires	2	0	5.63 – 18.8	1.13	$4.60 \times 10^{-4}$
Hires	3	0	18.8 – 120	1.57	$1.77 \times 10^{-4}$
Hires	4	0	> 120	2.78	$1.80 \times 10^{-5}$

We estimate the covariance of the power spectrum as

$$C^P(k, k') = \frac{1}{N-1} \sum_{i=1}^N [P_i(k) - \bar{P}(k)][P_i(k') - \bar{P}(k')], \quad (13)$$

where  $N$  is the number of realizations used and  $\bar{P}$  is the mean of the power spectrum measured from the simulations.

In Fig. 3, we plot the diagonal elements of  $C^P$  for dark matter. We show the results using the three sets of simulations. The results are normalized with respect to the Gaussian covariance, in which the power spectrum is the linear one. We note that the covariance from the Small set is lower than that from the Large set, and this is particularly evident at higher redshift. This is likely due to the beat-coupling effects as the Large simulation set has more large scale modes than the Small set. At low  $z$  the results from the Small and Hires are very similar, thus the mass resolution effects are small at low  $z$ , consistent with that reported in Ref. [15]. We also plot the perturbation theory results, which include the nonlinear power spectrum correction to the Gaussian covariance and the trispectrum Eq. 10. The agreement between perturbation theory and simulation results seems to be worse than what shown in Ref. [13]. A possible reason is that we have used the nonlinear power spectrum from simulations, which is lower than the 1-loop one in the nonlinear regime.

We now show the correlation coefficient  $r_P$  defined as

$$r_P(k_i, k_j) \equiv \frac{C^P(k_i, k_j)}{\sqrt{C^P(k_i, k_i)C^P(k_j, k_j)}}. \quad (14)$$

In Fig. 4 we plot  $r_P(k_i, k_j)$  as a function of  $k_j$ , with  $k_i$  fixed at the values of  $k_i = 0.076, 0.19, \text{ and } 0.40 \text{ Mpc}^{-1} h$

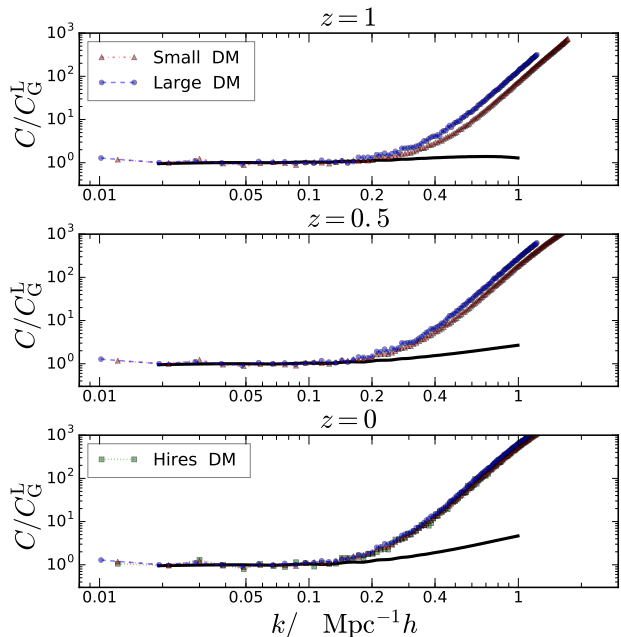


FIG. 3: The diagonal elements of the dark matter power spectrum covariance at  $z = 1, 0.5,$  and  $0$  (from top to bottom). The results are normalized with respect to the Gaussian covariance. The results from the Large (blue circles), Small (red triangles), and Hires (green squares) are shown. The perturbation theory predictions are overplotted (black solid line).

respectively. When both  $k_i$  and  $k_j$  are small, the agreement between the tree-level perturbation theory and the numerical results is reasonable, but it deteriorates for large  $k_i$  or  $k_j$ . See Ref.s [19, 20] for the improvement of the agreement by including the 1-loop corrections to the

trispectrum.

## 2. Halos

We now look at the the covariance of the halo power spectrum. We are usually interested only in the continuous halo power spectrum signal, and the Poisson shot noise is subtracted using Eq. C6. Caution must be taken for the  $\bar{n}$  appearing in the shot noise formulas. The mean density of halos  $\bar{n}$  is obtained by ensemble average, but is usually estimated using the volume averaged density measured in a particular simulation/survey. We will see that using the volume average number density results in a substantially smaller covariance.

We first look at the case when the ensemble averaged number density is used. This corresponds directly to the Poisson model prediction given in Appendix C 1. The ensemble averaged number density is obtained by further averaging the volume averaged ones over the realizations of the simulations. It is clear that in this case the covariance of the power spectrum with the Poisson shot noise subtracted is the same as that of the raw power spectrum. Here we use raw power spectrum to refer to the one measured directly from simulation without any shot noise subtraction.

In Fig. 5, we show the diagonal elements of the covariance of the halo power spectrum. We normalize the results using the Gaussian covariance Eq. 11. Halo groups from Large, Small, and Hires sets are used. Overall, we find that the Poisson model including the non-Gaussian corrections gives a reasonably good agreement with the numerical covariance up to  $k \sim 1 \text{ Mpc}^{-1} h$ . Similar to the dark matter case, we find that the Large simulation set yields a slightly larger covariance than the Small set in the nonlinear regime, while the Small and Hires set at  $z = 0$  agree with each other. Again, the difference between the Large and Small set could be caused by the beat coupling effect [24, 25, 27, 28] because the Large box has additional large scale modes compared to the Small set that source additional covariance. However, it is beyond the scope of this paper to investigate this in detail.

Because of the lower number of realizations available, the trends for the low bias groups from the Hires set are noisy. However, the model appears to agree well with the data up to  $k \sim 0.5 \text{ Mpc}^{-1} h$  thanks to the smallness of the non-Gaussian corrections for these kinds of halo group, as expected from Fig. 2. For larger  $k$ , the model underpredicts the covariance. This is also expected because the non-Gaussian corrections are more significant in this regime as can be seen from Fig. 2. Moreover, the other terms left out in the prediction may play a non-negligible role in this regime.

To compute the prediction we use a simple linear bias model

$$P_h(k) = b_1^2 P(k). \quad (15)$$

The linear bias  $b_1$  is obtained by fitting the model to the mean of the simulations up to  $k = 0.05 \text{ Mpc}^{-1} h$ . The best-fit values are shown in Table II. To compute the prediction in the Poisson model we should use the fully nonlinear polyspectra, but we find that the results are similar even if we use the nonlinear power spectrum measured from simulations instead.

In Fig. 6, we plot the correlation coefficient  $r_P$  for the halo power spectrum. We choose the same sets of  $k_i$  as in Fig. 4. We note that for dark matter,  $r_P(k_i, k_j)$  generally increases as  $k_j$  goes beyond the pivot scale  $k_i$ , while for halos,  $r_P(k_i, k_j)$  tends to level off or increases very mildly beyond the pivot scale. When  $k_i$  is small, the model can predict  $r_P(k_i, k_j)$  near the pivot scale, but tends to overpredict it when the separation from the pivot is large. We also note that the model performs worse at low  $z$  than at high  $z$ . This is again because at low  $z$  higher order correlators are more important.

We now consider the case when the Poisson shot noise obtained with the volume average number density is subtracted from each realizations. The covariance can then be expressed as

$$\begin{aligned} & \text{cov}(P_{\text{raw}}(k) - P_{\text{Pois}}(k), P_{\text{raw}}(k') - P_{\text{Pois}}(k')) \\ &= \text{cov}(P_{\text{raw}}(k), P_{\text{raw}}(k')) - \text{cov}(P_{\text{raw}}(k), P_{\text{Pois}}(k')) \\ & \quad - \text{cov}(P_{\text{Pois}}(k), P_{\text{raw}}(k')) + \text{cov}(P_{\text{Pois}}(k), P_{\text{Pois}}(k')), \end{aligned} \quad (16)$$

where  $P_{\text{Pois}}$  is defined as

$$P_{\text{Pois}} = \frac{1}{(2\pi)^3 n_{\text{va}}}, \quad (17)$$

with  $n_{\text{va}}$  being the volume-averaged number density obtained in a particular realization. When  $n_{\text{va}}$  fluctuates over realizations, we find that the second, third, and fourth terms in Eq. 16 result in negative correlation, and hence the Poisson subtracted correlation is lower than the raw results. For the diagonal part, if we assume that  $P_{\text{raw}}$  consists of  $P_{\text{Pois}}$  and the rest is not correlated with  $P_{\text{Pois}}$ , it is clear that the sum of the last three terms are negative. We leave it for future work to model the fluctuations of  $n_{\text{va}}$ , except noting that the fluctuations of the halo number density could be modeled along the line of Ref. [39]. In the context of the Poisson model, the number density is strictly constant since it is the ensemble averaged one, so it is not clear if we can self-consistently incorporate the shot noise fluctuations by simply using Eq. 17. For example, from Eq. 16, we do not get the non-Gaussian terms that we derived from the Poisson model.

In Fig. 7 we plot the diagonal elements of the halo power spectrum covariance when the Poisson shot noise from the individual realization is subtracted. Note that we still normalize the covariance using Eq. 11. Indeed, we find that the resultant covariance becomes much closer to the Gaussian one. This is good news because for the more realistic scenario, where the mean density is computed as a volume average over the survey/simulation, the halo

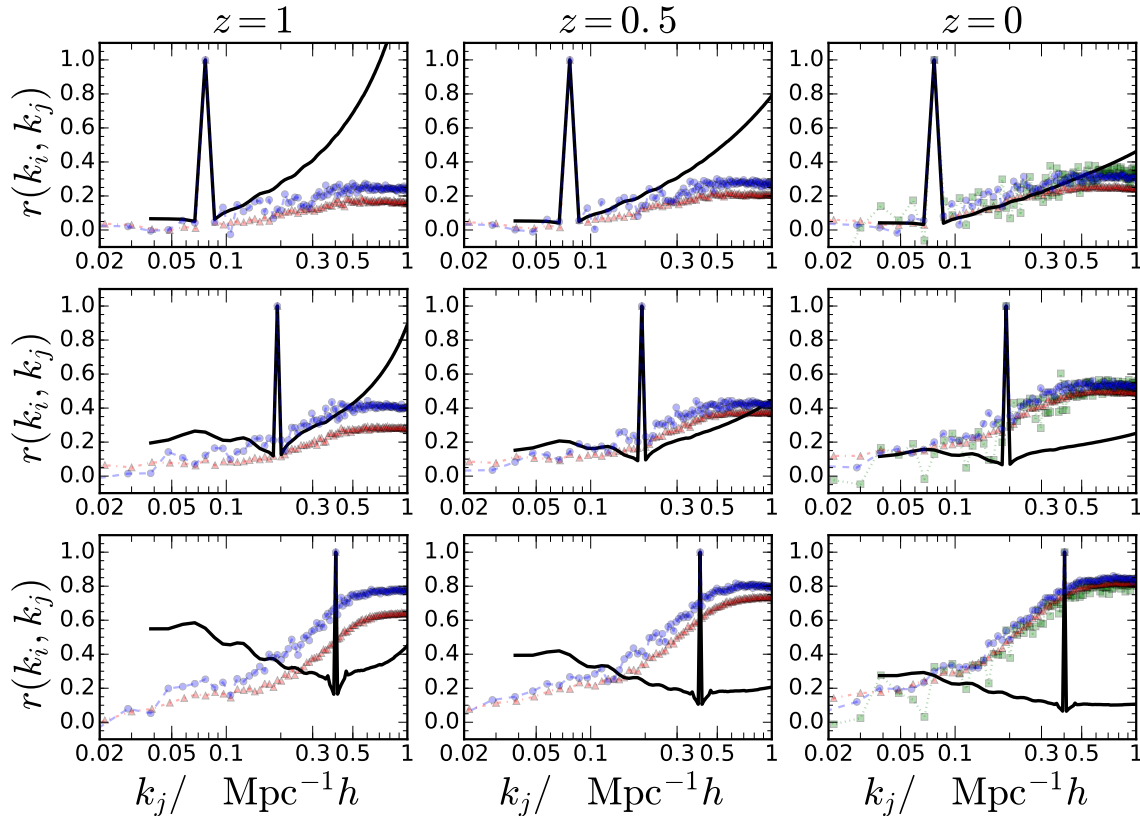


FIG. 4: The correlation coefficients of the dark matter power spectrum covariance at  $z = 1, 0.5,$  and  $0$  (from left to right). The results from the Large (blue circles), Small (red triangles), and Hires (green squares) are shown. The perturbation theory predictions are overplotted (black solid line).

power spectrum covariance is reduced and is easier to predict. This makes sense intuitively since when the number density is estimated from the realization, part of the fluctuations are absorbed in the shot noise term. The distinction between the local volume averaged density and the global ensemble one is analogous to the effects arising from defining the density contrast with the local or global average density found in Ref. [26]. We speculate that in this case the higher order correlators due to shot noise are partially destroyed. However, from results in Fig. 7 alone, we cannot unambiguously conclude that only the non-Gaussian contribution is reduced. It can still be possible that both the Gaussian and non-Gaussian parts of the shot noise contribution are reduced so that the total results are closer to the Gaussian predictions.

In Fig. 8 we show the correlation coefficient when the Poisson shot noise is subtracted from the individual realizations. Similarly to the diagonal elements, the off-diagonal elements are also reduced.

In the plots of this section the covariances are always normalized with respect to the Gaussian covariance. Here we would like to comment on the magnitudes of the matter and halo power spectrum covariances. For the massive halos in group 4, the power spectrum co-

variance is 3 to 5 order of magnitudes higher than the matter one, depending on the redshift in question. Thus for these kinds of halos, the shot noise contribution to the covariance completely dwarfs the matter power spectrum covariance. This also explains why the simple linear bias model works well. However, as the number density of halos increases, the shot noise contribution to the power spectrum decreases. For the Hires group 2, we still find that the halo power spectrum covariance is one order of magnitude higher than that of the dark matter. When the number density is as high as that of the Hires group 1, we find that the halo covariance is comparable in magnitude to that of the dark matter. In Euclid, the number density of galaxies is expected to reach  $\gtrsim 10^{-3}(\text{Mpc } h^{-1})^{-3}$  [7], hence the covariance is expected to get non-negligible contributions from dark matter non-linearity and galaxy biasing.

Finally we point out that Ref. [12] also studied the covariance of halo power spectrum using  $N$ -body simulations although using only 30 simulations of box size  $1500 \text{ Mpc } h^{-1}$ . Our results are similar to those in Ref. [12] regarding the effects of different shot noise subtraction procedures to the covariance. In particular, our finding that the Gaussian shot noise model works reasonably well

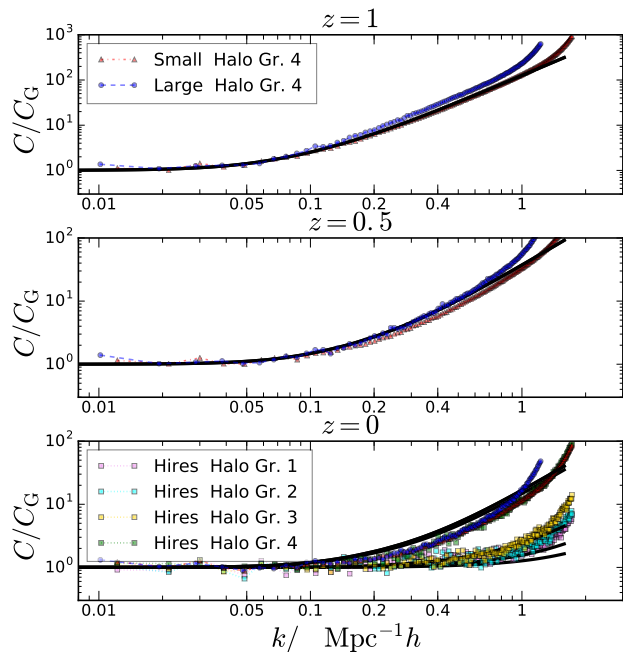


FIG. 5: The diagonal elements of the halo power spectrum covariance. The mean Poisson shot noise is subtracted. The results from Large Gr. 4 (blue circles), Small Gr. 4 (red triangles), and Hires (squares) Gr. 1 (violet), Gr. 2 (cyan), Gr. 3 (yellow), and Gr. 4 (green) are shown. The predictions are computed using Eq. 12 (black lines).

for the samples at  $z = 0$  with the shot noise subtracted from each realization agrees with Ref. [12].

### III. COVARIANCE OF BISPECTRUM

In this section, we will first discuss the theory of the bispectrum covariance for dark matter and halo. Then we will present the numerical covariance results and the comparison with theory.

#### A. Theory of the bispectrum covariance

The theoretical discussion on the bispectrum covariance in previous works has been mainly limited to the Gaussian contribution to the covariance [18, 40]. This is partly because the bispectrum covariance has a relatively large number of elements and the number of available realizations in most of the existing simulation sets is not large enough to get good signal to noise. Here we will consider the non-Gaussian contribution as well. We will see shortly that this is crucial to get good agreement with the simulation results.

Given the definition of the bispectrum

$$\langle \delta(\mathbf{k}_1)\delta(\mathbf{k}_2)\delta(\mathbf{k}_3) \rangle = B(k_1, k_2, k_3)\delta_D(\mathbf{k}_{123}), \quad (18)$$

we can construct an estimator as [18, 40]

$$\hat{B}(k_1, k_2, k_3) = \frac{k_F^3}{V_{123}} \int_{k_1} d^3p \int_{k_2} d^3q \int_{k_3} d^3r \times \delta_D(\mathbf{p} + \mathbf{q} + \mathbf{r})\delta(\mathbf{p})\delta(\mathbf{q})\delta(\mathbf{r}), \quad (19)$$

where  $k_i$  indicates the integration is over a spherical shell of width  $[k_i - \Delta k/2, k_i + \Delta k/2)$ , with  $\Delta k$  being the width of the bin in Fourier space. The term  $V_{123}$  counts the number of modes satisfying the triangle constraint:

$$V_{123} = \int_{k_1} d^3p \int_{k_2} d^3q \int_{k_3} d^3r \delta_D(\mathbf{p} + \mathbf{q} + \mathbf{r}). \quad (20)$$

In fact, we can compute  $V_{123}$  analytically to get [18]

$$V_{123} = 8\pi^2 k_1 k_2 k_3 (\Delta k)^3 \beta(\Delta), \quad (21)$$

where  $\Delta$  is defined as

$$\Delta = \hat{\mathbf{k}}_1 \cdot \hat{\mathbf{k}}_2 = \frac{k_3^2 - k_1^2 - k_2^2}{2k_1 k_2}, \quad (22)$$

and  $\beta(\Delta)$  is given by

$$\beta(\Delta) = \begin{cases} \frac{1}{2} & \text{if } \Delta = \pm 1 \\ 1 & \text{if } 0 < \Delta < 1 \\ 0 & \text{otherwise} \end{cases} \quad (23)$$

For more details on the derivation of Eq. 21, see Appendix B. In Appendix A, we check the probability distribution of  $\hat{B}$ . We find that it is close to Gaussianly distributed but with non-negligible skewness and kurtosis as well.

The covariance matrix of  $\hat{B}$ ,  $C^B$  is defined as:

$$\begin{aligned} C^B(k_1, k_2, k_3, k'_1, k'_2, k'_3) \\ \equiv \text{cov}[\hat{B}(k_1, k_2, k_3), \hat{B}(k'_1, k'_2, k'_3)] \\ = \langle \hat{B}(k_1, k_2, k_3)\hat{B}(k'_1, k'_2, k'_3) \rangle \\ - \langle \hat{B}(k_1, k_2, k_3) \rangle \langle \hat{B}(k'_1, k'_2, k'_3) \rangle. \end{aligned} \quad (24)$$

In the following we will skip the superscript  $B$  in the notation for the sake of simplicity. The contribution to  $C$  arises from both the disconnected and connected 6-point function in  $\langle \hat{B}\hat{B}' \rangle$ .

#### 1. Dark matter

When  $\delta$  is Gaussian,  $\langle \hat{B} \rangle$  is zero, however  $C$  does not vanish. The Gaussian covariance can be written as [18]

$$C_G^L = \frac{k_F^3}{V_{123}} \delta_{k_1 k_2 k_3, k'_1 k'_2 k'_3} s_{123} P_L(k_1) P_L(k_2) P_L(k_3), \quad (25)$$

where  $\delta_{k_1 k_2 k_3, k'_1 k'_2 k'_3}$  is non-vanishing only if the shape of the triangle  $k_1 k_2 k_3$  is the same as that of  $k'_1 k'_2 k'_3$ . If none of the sides of the triangle are equal to each other,  $s_{123} = 1$ . If the triangles are isocetes,  $s_{123} = 2$ . For



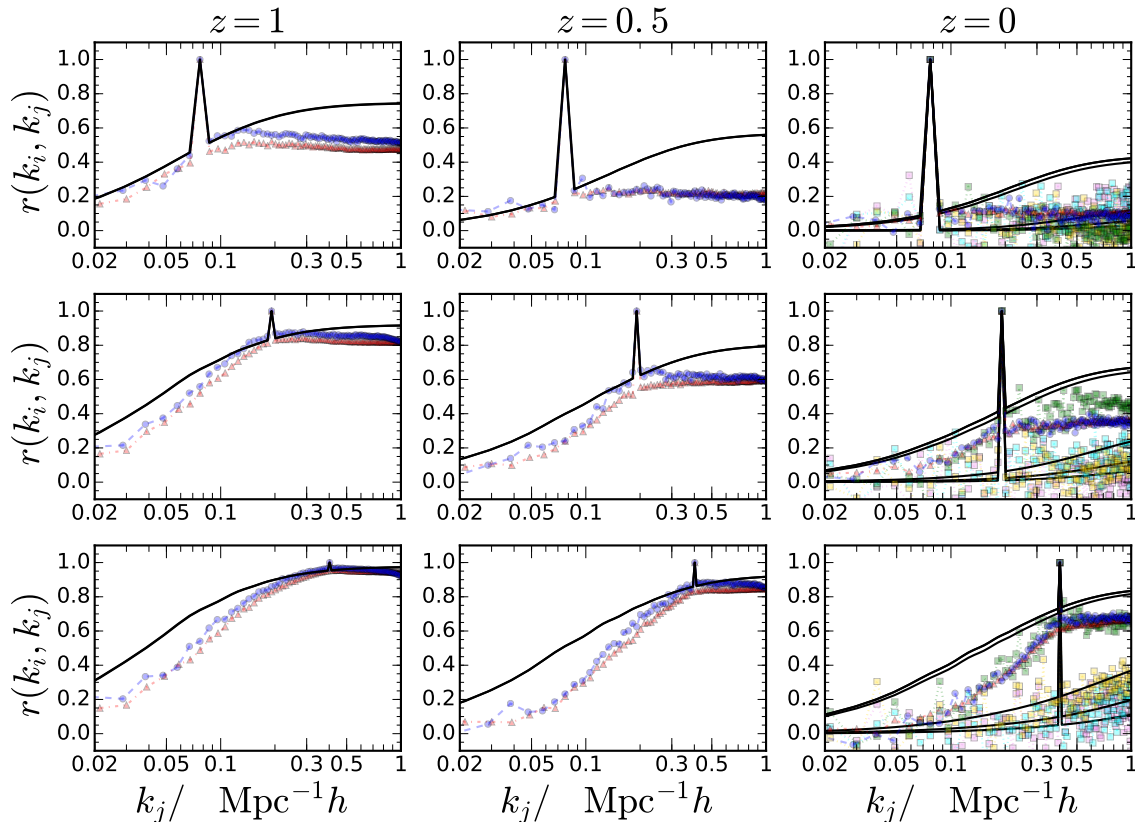


FIG. 6: Similar to Fig. 4, except for the halo power spectrum covariance. The mean Poisson shot noise is subtracted. The results from Large Gr. 4 (blue circles), Small Gr. 4 (red triangles), and Hires (squares) Gr. 1 (violet), Gr. 2 (cyan), Gr. 3 (yellow), and Gr. 4 (green) are shown. The predictions are computed using Eq. 12 (black lines).

equilateral triangles, we have  $s_{123} = 6$ . The derivation of Eq. 25 is reviewed in Appendix B. In Eq. 25, for Gaussian  $\delta$ ,  $P_L(k)$  is the linear power spectrum. As we consider the non-Gaussian contribution below, we find that part of the contribution can be resummed if we use the 1-loop matter power spectrum instead of the linear one for one of the power spectra. However, the 1-loop power spectrum overestimates the matter power spectrum from simulation in the weakly nonlinear regime already. Similar to the case of power spectrum, we shall use the nonlinear power spectrum measured from simulations in place of the 1-loop results. We will use the notation  $C_G^{\text{NL}}$  to distinguish the case when the nonlinear power spectrum is used, that is

$$C_G^{\text{NL}} = C_G^L + \frac{k_F^3}{V_{123}} \delta_{k_1 k_2 k_3, k'_1 k'_2 k'_3} s_{123} \\ \times [P_L(k_1)P_L(k_2)(P_{\text{NL}}(k_3) - P_L(k_3)) + 2 \text{cyc.}], \quad (26)$$

where cyc. denotes cyclic permutations and  $P_{\text{NL}}$  denotes the nonlinear power spectrum.

In the top-left corner of Fig. 9, we show a diagrammatic representation of  $C_G^{\text{NL}}$ . The rules are the same as that in Fig. 1. The black dots on the left and right hand

side represent the three  $\delta$ 's in each of the bispectrum estimator  $\hat{B}$ . The curly line represents the linear power spectrum. We put a filled circle on one of the curly line to indicate that the nonlinear power spectrum is used instead of the linear one.

Nonlinear evolution causes mode coupling and the departure from Gaussianity. We will also consider the disconnected non-Gaussian contributions to the 6-point function. At the tree-level order, there are both connected and disconnected contributions to the 6-point function. In this paper we shall take the ansatz that the disconnected part is more important than the connected ones in the weakly nonlinear regime. This ansatz is partly motivated by the fact that the Gaussian contribution is also disconnected. The disconnected non-Gaussian tree level contributions can arise either from  $F_3$  or  $F_2$  kernel. In Fig. 9, we show a graphical representation of these non-Gaussian contributions. There is one contribution due to  $F_3$ ,  $C_{F_3}$ , in which  $F_3$  is represented as three legs branching from a black dot. The non-Gaussian tree level can also arise from two  $F_2$  kernels. We further classify these diagrams into type I if both  $F_2$  kernels are on the same triangle, and type II if they are distributed on dif-

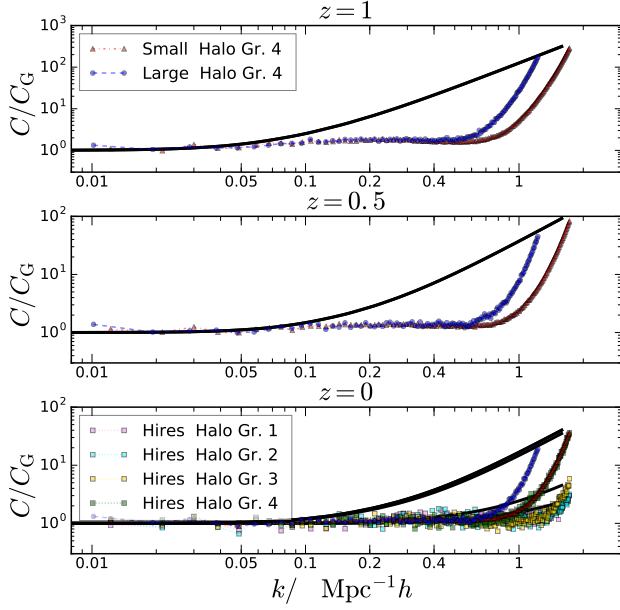


FIG. 7: Similar to Fig. 5, except that the Poisson shot noise estimated from individual realization is subtracted. By so doing the resultant covariance is closer to the Gaussian ones. The predictions are the same as those shown in Fig. 5.

First for convenience we define the notation

$$\int \mathcal{D} \equiv \frac{k_F^3}{V_{123} V'_{123}} \int_{k_1} d^3 p \int_{k_2} d^3 q \int_{k_3} d^3 r \delta_D(\mathbf{p} + \mathbf{q} + \mathbf{r}) \int_{k'_1} d^3 p' \int_{k'_2} d^3 q' \int_{k'_3} d^3 r' \delta_D(\mathbf{p}' + \mathbf{q}' + \mathbf{r}'). \quad (27)$$

The contribution due to  $F_3$  reads

$$C_{F_3} = 6 \int \mathcal{D} \delta_D(\mathbf{p} + \mathbf{p}') F_3(\mathbf{q}, \mathbf{q}', \mathbf{r}') P(p) P(q) P(q') P(r') + [(3! \times 3 + k_{123} \leftrightarrow k'_{123}) - 1] \text{cyc.} \quad (28)$$

By inspecting the diagrammatic representation of  $C_{F_3}$  in Fig. 9, it is easy to see that there are  $3!$  permutations for  $k_1 k_2 k_3$  and 3 permutations for  $k'_1 k'_2 k'_3$  as two of the legs are symmetric. There are also additional contributions from interchanging  $k_1 k_2 k_3$  and  $k'_1 k'_2 k'_3$ . Thus there are altogether 36 permutations. For other diagrams the number of permutations can be worked out in a similar manner.

The type I contributions due to  $F_2$  are

$$C_{F_2I}^1 = 2 \int \mathcal{D} \delta_D(\mathbf{p} - \mathbf{p}') F_2(\mathbf{q}, \mathbf{r}') F_2(\mathbf{p}', \mathbf{r}) P_L(q) P_L(r') P_L(p') P_L(r) + [(3! \times 3 + k_{123} \leftrightarrow k'_{123}) - 1] \text{cyc.}, \quad (29)$$

$$C_{F_2I}^2 = 8 \int \mathcal{D} \delta_D(\mathbf{p} + \mathbf{p}') F_2(-\mathbf{q}', \mathbf{q} + \mathbf{q}') F_2(\mathbf{q} + \mathbf{q}', \mathbf{r}') P_L(p) P_L(q') P_L(r') P_L(|\mathbf{q} + \mathbf{q}'|) + [(3 \times 3 + k_{123} \leftrightarrow k'_{123}) - 1] \text{cyc.} \quad (30)$$

The type II contribution reads

$$C_{F_2II}^1 = 4 \int \mathcal{D} \delta_D(\mathbf{p} - \mathbf{p}') F_2(\mathbf{q}, \mathbf{p}) F_2(\mathbf{q}', \mathbf{p}) P_L(q) P_L(p') P_L(p) P_L(q') + [(3! \times 3!) - 1] \text{cyc.}, \quad (31)$$

$$C_{F_2II}^2 = 4 \int \mathcal{D} \delta_D(\mathbf{p} + \mathbf{p}') F_2(\mathbf{r}, \mathbf{p}) F_2(-\mathbf{p}, \mathbf{r}') P_L(p) P_L(r) P_L(|\mathbf{q} + \mathbf{r}|) P_L(r') + [(3! \times 3!) - 1] \text{cyc.} \quad (32)$$

$$C_{F_2II}^3 = 4 \int \mathcal{D} \delta_D(\mathbf{p} - \mathbf{p}') F_2(\mathbf{q}, \mathbf{r}) F_2(\mathbf{q}', \mathbf{r}') P_L(q) P_L(q') P_L(r) P_L(r') + [(3 \times 3) - 1] \text{cyc.} \quad (33)$$

$$C_{F_2II}^4 = 4 \int \mathcal{D} \delta_D(\mathbf{p} + \mathbf{p}') F_2(-\mathbf{r}', \mathbf{q} + \mathbf{r}') F_2(\mathbf{q} + \mathbf{r}', \mathbf{r}) P_L(p) P_L(r') P_L(r) P_L(|\mathbf{q} + \mathbf{r}'|) + [(3! \times 3!) - 1] \text{cyc.} \quad (34)$$

ferent triangles. There are two type I diagrams and four type II.

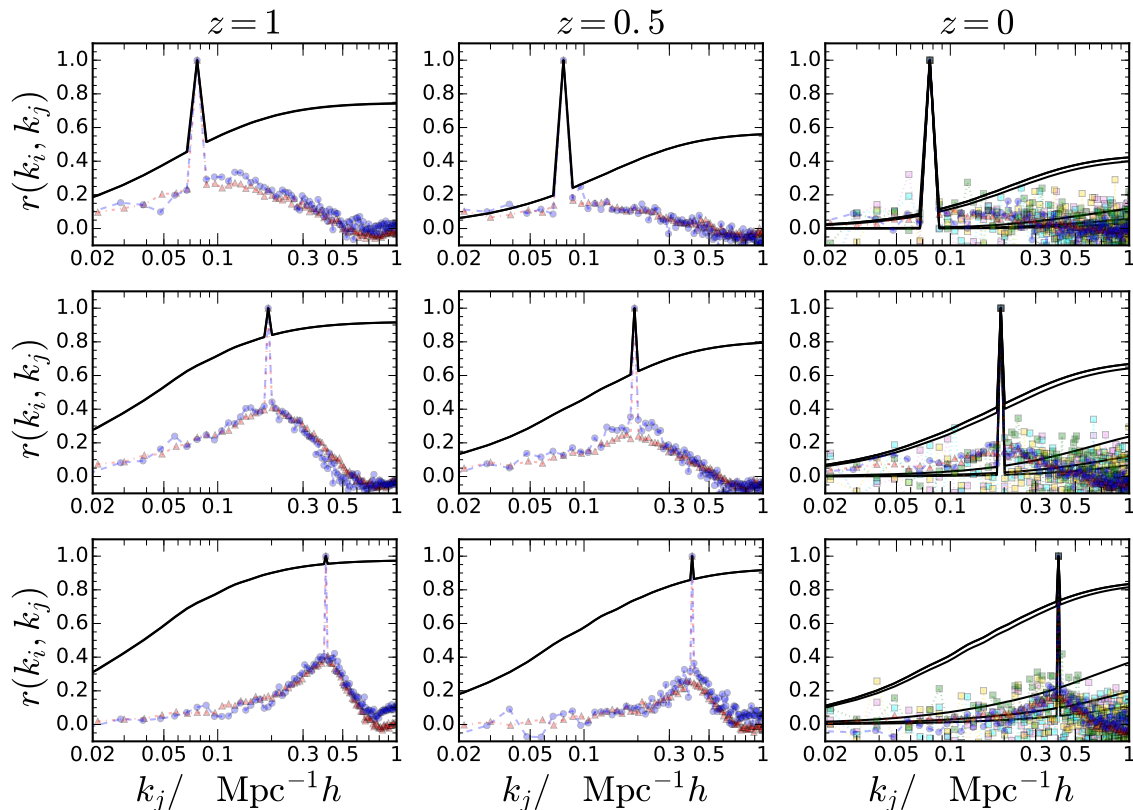


FIG. 8: Similar to Fig. 6, except that the Poisson shot noise estimated from individual realization is subtracted. The predictions are the same as those shown in Fig. 6.

In each of the non-Gaussian terms there are four Dirac delta functions: two due to the triangle constraints  $\delta_{\mathbf{D}}(\mathbf{p}+\mathbf{q}+\mathbf{r})$  and  $\delta_{\mathbf{D}}(\mathbf{p}'+\mathbf{q}'+\mathbf{r}')$ , and another two imposed on two disjoint sets of  $\mathbf{p}$ ,  $\mathbf{q}$ ,  $\mathbf{r}$ ,  $\mathbf{p}'$ ,  $\mathbf{q}'$ , and  $\mathbf{r}'$ . These can also be seen by inspecting the non-Gaussian diagrams in Fig. 9. Hence it is clear that one of these latter two Dirac delta functions is redundant, and it will give  $\delta_{\mathbf{D}}(\mathbf{0})$ . That is why in  $\int \mathcal{D}$  there is a factor of  $k_{\mathbf{F}}^3$  only. The other one relates two vectors, each coming from one of the bispectrum estimator, which is  $\delta_{\mathbf{D}}(\mathbf{p} \pm \mathbf{p}')$  in Eq. 28-34. This Dirac delta function is analytically integrable. We will make use of the pattern of the Dirac delta functions to construct an efficient Monte Carlo integration method.

Except for the Dirac delta  $\delta_{\mathbf{D}}(\mathbf{p} \pm \mathbf{p}')$ , the other two remaining Dirac delta functions in general cannot be integrated analytically. Thus the resultant integral is of quite high dimension (15). High dimensionality causes problem for numerical integrators in general. The presence of the remaining two Dirac delta functions make the integrand non-vanishing only in narrow regions. Although high dimensional integrals can often be attacked by Monte Carlo integration method, generic Monte Carlo integration would fail as they would miss the narrow peaks in the high dimensional space.

Here we present a Monte Carlo method that can efficiently sample the points that satisfy the Dirac delta function constraints. We first note that the vectors that fulfill the triangle constraint must be some small perturbations of the triangle  $k_1 k_2 k_3$  and  $k'_1 k'_2 k'_3$ . Thus instead of sampling all the points in the full integration domain, we can proceed as follows. We first generate a vector  $\mathbf{p}$  in the  $k_1$  shell randomly. For  $\mathbf{q}$  we must have  $\hat{\mathbf{p}} \cdot \hat{\mathbf{q}} \equiv \mu \approx \Delta$ . To determine the allowed variation of  $\mu$ , we consider

$$d\Delta = \frac{k_1^2 - k_3^2 - k_2^2}{2k_1 k_2^2} dk_2 + \frac{k_3}{k_1 k_2} dk_3. \quad (35)$$

Hence we sample  $\mu$  uniformly in the range  $\pm \left[ \left( \frac{k_1^2 - k_3^2 - k_2^2}{2k_1 k_2^2} \Delta k \right)^2 + \left( \frac{k_3}{k_1 k_2} \Delta k \right)^2 \right]^{1/2}$ . To further fix  $\mathbf{q}$ , we sample the polar angle of  $\mathbf{q}$ ,  $\theta_2$ , uniformly in the interval  $[0, \pi]$ . The azimuthal angle of  $\mathbf{q}$ ,  $\phi_2$  is fixed by the relation

$$\cos(\phi_1 - \phi_2) = \frac{\mu - \cos \theta_1 \cos \theta_2}{\sin \theta_1 \sin \theta_2}, \quad (36)$$

where  $\theta_1$  and  $\phi_1$  are the spherical coordinates of  $\mathbf{p}$ . If the length  $|\mathbf{p}+\mathbf{q}|$  falls within the interval  $[k_3 - \Delta/2, k_3 + \Delta/2]$ , the vector  $\mathbf{r}$  is assigned to be  $-(\mathbf{p} + \mathbf{q})$ , and the three

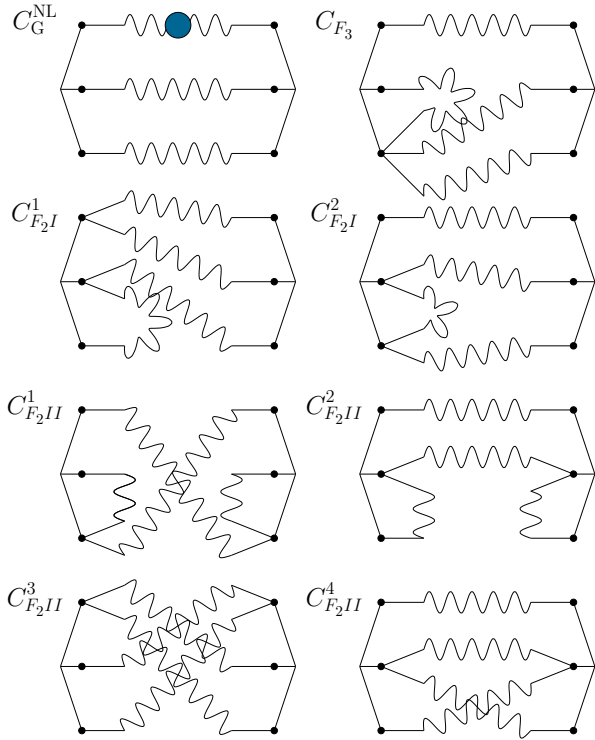


FIG. 9: A diagrammatic representation of the covariance of the matter bispectrum to the tree level order.  $C_G^{\text{NL}}$  is the Gaussian term, while the rest are non-Gaussian contributions. The set of black dots on the left and right of each diagram represent the three  $\delta$ 's in each bispectrum estimator. The legs branching from each dot represent the perturbation theory kernel,  $F_1$ ,  $F_2$ , and  $F_3$  respectively. Each wavy line represents the linear power spectrum. In  $C_G^{\text{NL}}$ , the filled circle indicates that the linear power spectrum is replaced by the nonlinear power spectrum.

vectors  $\mathbf{p}$ ,  $\mathbf{q}$ , and  $\mathbf{r}$  are accepted, otherwise the procedure is repeated until the proposed vectors are accepted. We find that the acceptance rate can reach about 20% and it does not vary much with the triangle configuration considered. For the triads  $\mathbf{p}'$ ,  $\mathbf{q}'$ , and  $\mathbf{r}'$ , we make use of the Dirac delta function  $\delta_{\text{D}}(\mathbf{p} \mp \mathbf{p}')$  and assign  $\mathbf{p}' = \pm \mathbf{p}$  accordingly. The construction of  $\mathbf{q}'$  and  $\mathbf{r}'$  are then similar to those for  $\mathbf{q}$  and  $\mathbf{r}$ .

After developing an efficient algorithm to sample the vectors satisfying the constraints imposed by the three Dirac delta functions, we can attack the non-Gaussian integrals using the Monte Carlo method (see e.g. [41] for a review). The integrals can be schematically written as

$$I_{\text{NG}} = \int \mathcal{D}\delta_{\text{D}}(\mathbf{p} \mp \mathbf{p}') f(\mathbf{p}, \mathbf{q}, \mathbf{r}, \mathbf{p}', \mathbf{q}', \mathbf{r}') \\ = \langle f(\mathbf{p}, \mathbf{q}, \mathbf{r}, \mathbf{p}', \mathbf{q}', \mathbf{r}') \rangle \int \mathcal{D}\delta_{\text{D}}(\mathbf{p} \mp \mathbf{p}'), \quad (37)$$

where the integrand  $f$  is averaged over the points in the integration domain defined by the Dirac delta functions. The success of this method relies on the fact that the integration volume can be computed analytically and it

reads

$$\int \mathcal{D}\delta_{\text{D}}(\mathbf{p} \mp \mathbf{p}') = \frac{k_{\text{F}}^3}{V_{123} V'_{123}} U(k_1, k'_1), \quad (38)$$

where  $U$  is given by

$$U(k_1, k'_1; k_2, k_3, k'_2, k'_3) \\ = 2^4 \pi^3 k_2 k_3 k'_2 k'_3 (\Delta k)^5 \beta(\Delta) \beta(\Delta') \delta_{k_1, k'_1}. \quad (39)$$

In the Appendix B, we show the derivation of  $U$ . From Eq. 37 and 38, we find that the non-Gaussian integral corrections scales with the bin width  $\Delta k$  as  $1/\Delta k$ . In contrast the Gaussian term  $C_G$  exhibits a stronger scaling  $1/(\Delta k)^3$ .

From Eq. 39, we deduce that the non-Gaussian terms scale with the volume of the simulation as  $k_{\text{F}}^3$ , the same as the Gaussian term. The volume dependence is the same as that for the power spectrum covariance. As we mentioned, this is a consequence of the statistical translational invariance.

In Fig. 10, we show the various contributions to the diagonal elements of the covariance matrix of the dark matter bispectrum at  $z = 0$ . Unless otherwise stated, for the bispectrum theory computations, the Small box size  $656.25 \text{ Mpc } h^{-1}$  and the binning  $\Delta k = 0.019 \text{ Mpc}^{-1} h$  are adopted. In order not to proliferate the paper with too many figures, we shall only show the results for the equilateral shape. In the lower panel of Fig. 10, we plot the ratio of various terms with respect to  $C_G^{\text{L}}$ . We have split  $C_G^{\text{NL}}$  into  $C_G^{\text{L}}$  and  $C_G^{\text{Looop}}$ , which denotes the corrections beyond the tree level in Eq. 26. The Gaussian covariance contributions are dominant over the non-Gaussian contributions in the scales shown. The linear Gaussian term  $C_G^{\text{L}}$  is the leading contribution, although it is overtaken by nonlinear correction  $C_G^{\text{Looop}}$  for  $k \sim 0.2 \text{ Mpc}^{-1} h$ . All the non-Gaussian contributions are positive except  $C_{F_2II}^4$ , but its magnitude is negligible compared to others. At  $k \lesssim 0.04 \text{ Mpc}^{-1} h$ ,  $C_{F_2II}^1$  is largest among all the non-Gaussian contributions. Beyond that,  $C_{F_2I}^2$  is the only dominant non-Gaussian contribution. Overall for  $k \sim 0.15 \text{ Mpc}^{-1} h$  the non-Gaussian contribution can be as high as 10 to 20 % of the Gaussian covariance.

In this paper, we only evaluate the disconnected tree-level 6-point function, assuming that the connected tree-level contributions are of smaller magnitude in the weakly nonlinear regime. While the numerical results appear to support this ansatz, we will leave it to future work to explicitly compute the connected tree-level contributions, and perhaps higher order corrections as well.

## 2. Halos

The discrete nature of halos causes stochastic fluctuations. Similar to the case of power spectrum, the shot noise is the main source of halo bispectrum covariance as we will discuss in next section. In Appendix C 2, we

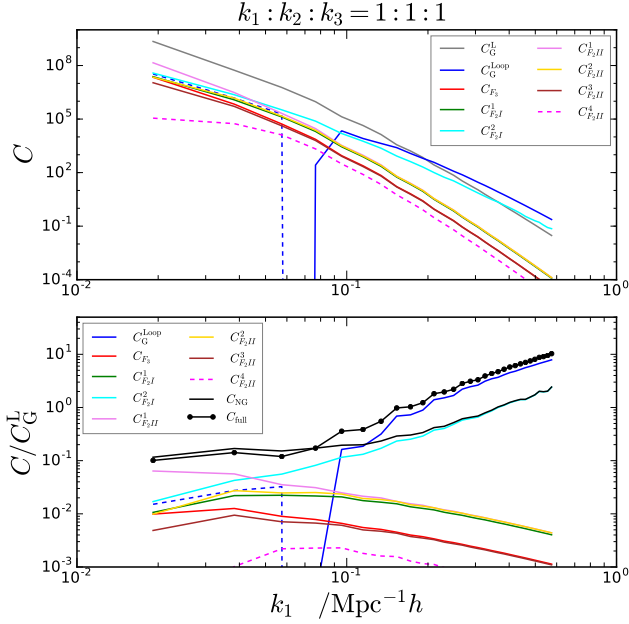


FIG. 10: The leading tree-level contributions to the diagonal elements of the dark matter bispectrum covariance matrix. The equilateral triangle configuration with the binning  $\Delta k = 0.019 \text{ Mpc}^{-1} h$  is used here and all the plots for bispectrum. The results are evaluated at  $z = 0$ . Upper panel: the Gaussian contributions ( $C_G^L$  and  $C_G^{L,oop}$ ) and various non-Gaussian contributions. Lower panel: the ratio of various terms normalized with respect to  $C_G^L$ . The positive parts are represented by solid lines, while dashed for the negative. At low  $k$  the dominant contribution is  $C_G^L$  (grey), while it is overtaken by  $C_G^{L,oop}$  (blue) at  $k \sim 0.2 \text{ Mpc}^{-1} h$ . The major non-Gaussian contributions are  $C_{F_{2II}}^1$  (violet) at low  $k$  and  $C_{F_{2I}}^2$  (cyan) for  $k \gtrsim 0.1 \text{ Mpc}^{-1} h$ . In the lower panel the full result is shown as a black line with markers.

use the Poisson model to derive the contribution to the bispectrum covariance due to Poisson fluctuations. We refer the readers to the appendix for the details. Here we summarize the main results.

We find that there are terms in the Poisson model that couple triangles of the same shape, as the Gaussian term Eq. 25. These terms contains two Dirac delta functions in the 6-point function Eq. C21. In fact these terms can be combined with Eq. 25 as

$$C_G = \frac{k_F^3}{V_{123}} \delta_{k_1 k_2 k_3, k'_1 k'_2 k'_3} s_{123} \left[ P_h(k_1) + \frac{1}{\bar{n}} \right] \times \left[ P_h(k_2) + \frac{1}{\bar{n}} \right] \left[ P_h(k_3) + \frac{1}{\bar{n}} \right]. \quad (40)$$

Eq. 40 agree with [40]. Although the Poisson contributions do not arise from Gaussian fluctuations, they are on the same footing as the smooth Gaussian term, we call them Gaussian tems as well. It is worth stressing that for both the power spectrum and bispectrum, the Gaussian Poisson contribution can be obtained by replacing

the dark matter power spectrum by the halo power spectrum plus the Poisson shot noise.

There are also non-Gaussian contributions due to Poisson fluctuations. We remind the readers that all the continuous correlators in Appendix C 2 are the nonlinear correlators of the tracer, such as the halos, although they are not labeled explicitly in the formulas there for simplicity. We further classify the terms containing correlator up to the power spectrum and having no Dirac delta function in Eq. C21 as  $C_{NG}^{(0)}$  (Eq. C27). Those containing correlator up to the power spectrum and having one Dirac delta function in Eq. C21 are classified as  $C_{NG}^{(1)}$  (Eq. C26). Both  $C_{NG}^{(0)}$  and  $C_{NG}^{(1)}$  are relatively easy to compute. We group the terms with the halo bispectrum as  $C_{NG}^{(2)}$  (Eq. C28). Some of the terms in  $C_{NG}^{(2)}$  can be analytically integrable and we will only consider these terms in this paper. That is we consider

$$C_{NG}^{(2)'} = \frac{k_F^3 U(k_1, k'_1)}{VV'} \left\{ \frac{1}{\bar{n}^2} [B_h(k_1, k'_2, k'_3) + B_h(k'_1, k_2, k_3)] + 8 \text{ cyc.} \right. \\ \left. + \frac{1}{\bar{n}} [(P_h(k'_1) + P_h(k_2) + P_h(k_3)) B_h(k_1, k'_2, k'_3) \right. \\ \left. + (P_h(k_1) + P_h(k'_2) + P_h(k'_3)) B_h(k'_1, k_2, k_3)] + 8 \text{ cyc.} \right\} \quad (41)$$

The remaining part of  $C_{NG}^{(2)}$  involves an integral of halo bispectrum. Because of various nonlinear effects have to be properly taken into account, modelling the halo bispectrum accurately for  $k \gtrsim 0.15 \text{ Mpc}^{-1} h$  at  $z = 0$  is still difficult. We can instead use the mean of the bispectrum measured from simulation. For  $C_{NG}^{(2)'}$  this is relatively easy to do as it only depends on bispectrum evaluated at the external wave number ranges. But for the integrated terms, we need the bispectrum values over a large range outside the external wave numbers. To do so we need a set of dedicated measurements, which we will leave it for future works. We will not consider the remaining terms with even higher order correlators in the discrete 6-point function (see Eq. C21).

We plot the shot noise contributions to the covariance of the halo bispectrum in Fig. 11 for two selected halo groups, which correspond to Gr. 4 of the Large/Small simulation set at  $z = 0.5$  and Gr. 2 of the Hires set at  $z = 0$ . We have split the Gaussian covariance into the smooth and the Poisson parts. For the high bias group, the Gaussian Poisson component is dominant over the smooth Gaussian part, while for the low bias group they are comparable. We find that for the high mass group, the non-Gaussian components are non-negligible, especially  $C_{NG}^{(1)}$ . The analytic part  $C_{NG}^{(3)'}$  also suggests that other terms in the bispectrum contribution  $C_{NG}^{(3)}$  are small but not negligible in comparison with the Gaussian contribution. On the other hand, for the low bias group the non-Gaussian contributions can be neglected.

Although in the Poisson model the continuous correlators are the nonlinear ones, we have compared the results

obtained using the measured  $P_h$  and Eq. 15 and found that the differences are negligible for our purpose here. For  $C_{\text{NG}}^{(3)'}$ , we have used the measured  $B_h$  and that is the reason why the results are sometimes noisy in Fig. 11.

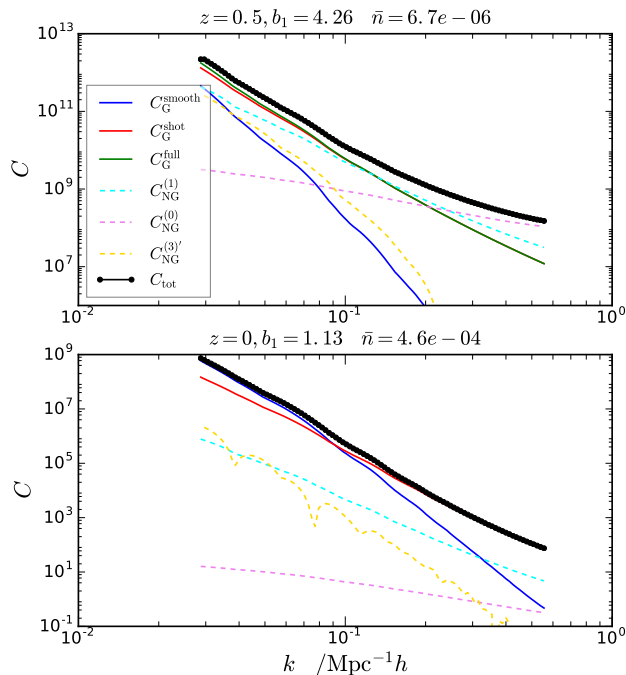


FIG. 11: Various shot noise contributions to the covariance of the halo bispectrum. The diagonal elements for the equilateral triangle configuration are shown. We have plotted the results for two selected groups with the parameters used written on the plot. The solid lines are the smooth Gaussian contribution (blue), the Gaussian Poisson contribution (red). The dashed lines are the non-Gaussian Poisson contributions  $C_{\text{NG}}^{(0)}$  (violet),  $C_{\text{NG}}^{(1)}$  (cyan), and  $C_{\text{NG}}^{(3)}$  (yellow). The sum of the Gaussian and non-Gaussian contributions are represented as a black line with markers.

## B. Numerical results

### 1. Dark matter

We first look into the covariance of the dark matter bispectrum. The covariance is estimated from the available realizations as

$$\begin{aligned}
 & C(k_1, k_2, k_3, k'_1, k'_2, k'_3) \\
 &= \frac{1}{N-1} \sum_{i=1}^N [B_i(k_1, k_2, k_3) - \bar{B}(k_1, k_2, k_3)] \\
 & \quad \times [B_i(k'_1, k'_2, k'_3) - \bar{B}(k'_1, k'_2, k'_3)], \quad (42)
 \end{aligned}$$

where  $N$  is the number of realizations used and  $\bar{B}$  is the mean of the bispectrum measured from the simulations.

We show the diagonal elements of the bispectrum covariance matrix in Fig. 12. In this plot we use the equilateral triangle configurations. The results are normalized with respect to the Gaussian covariance  $C_G$ , Eq. 25. The results from the Large, Small, and Hires simulation sets agree with each other well. We find that the non-Gaussian contribution to the covariance increases rapidly as  $z$  decreases from 1 to 0 in the mildly nonlinear regime  $k \sim 0.4 \text{ Mpc}^{-1} h$ . The covariance is within 20% from  $C_{\text{NG}}^{\text{L}}$  up to  $k \sim 0.2 \text{ Mpc}^{-1} h$  at  $z = 1$  and  $k \sim 0.1 \text{ Mpc}^{-1} h$  at  $z = 0$ . We also plotted the non-Gaussian prediction described in Sec. III A 1. The non-Gaussian correction gives the predictions agree with the simulation results up to  $k \sim 0.3 \text{ Mpc}^{-1} h$  at  $z = 1$  and  $0.16 \text{ Mpc}^{-1} h$  at  $z = 0$ .

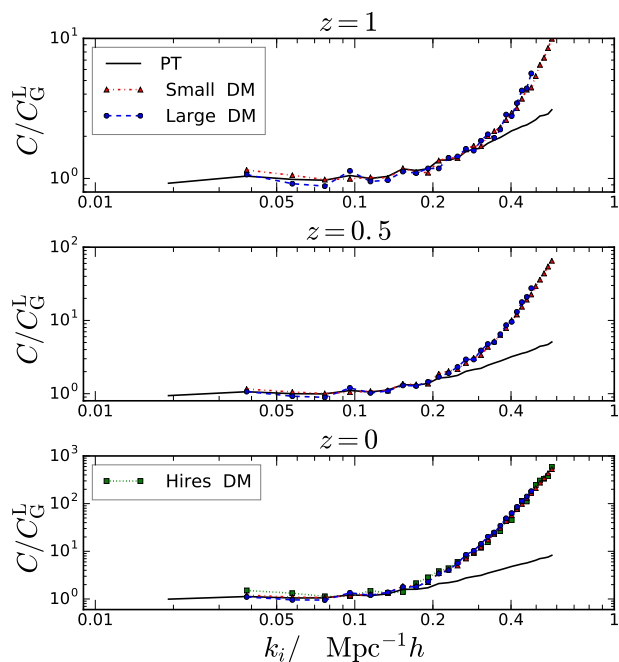


FIG. 12: The diagonal elements of the dark matter bispectrum covariance matrix for equilateral triangle configuration at  $z = 1, 0.5$ , and  $0$  (from top to bottom). The results from the simulation set Small (red triangles), Large (blue circles), and Hires (green squares) are shown. The perturbation theory predictions (solid black lines) are also overplotted.

Similar to the case of the power spectrum, the cross correlation coefficient  $r$  is defined as

$$\begin{aligned}
 & r(k_1, k_2, k_3, k'_1, k'_2, k'_3) \\
 &= \frac{C(k_1, k_2, k_3, k'_1, k'_2, k'_3)}{\sqrt{C(k_1, k_2, k_3, k_1, k_2, k_3)C(k'_1, k'_2, k'_3, k'_1, k'_2, k'_3)}}. \quad (43)
 \end{aligned}$$

We plot  $r$  for equilateral triangle configuration in Fig. 13. In each subplot, we fix one of the length of the triangle

$k_i$  and vary the length of the other one,  $k_j$ . In these plots, we have fixed  $k_i$  to be 0.076, 0.19, and 0.40  $\text{Mpc}^{-1} h$  respectively.

As we noted previously, the Gaussian covariance couples only the triangle of the same shape, while the leading non-Gaussian corrections couple triangles with at least one side equal to each other. Thus both the Gaussian and disconnected non-Gaussian tree-level corrections cannot give rise to cross correlations between equilateral triangles of different sizes. We see that indeed when both  $k_i$  and  $k_j$  are small, the correlation coefficients are consistent with being zero. As one of the wavenumbers increases  $r$  also starts to increase. The larger the value of the wavenumbers the larger the covariance. This is expected from the fact that nonlinearity increases the coupling of different wave modes. The accuracy of the perturbation theory prediction is qualitatively similar to that for power spectrum covariance shown in Fig. 4.

## 2. Halos

We now move to the halo bispectrum covariance. The Poisson shot noise for the halo bispectrum is given by Eq. C10. As we are interested only in the smooth correlation function signal, the contribution to the halo bispectrum due to Poisson fluctuations is usually subtracted. Similar to the case of power spectrum, we will distinguish between the case when the mean Poisson shot noise is subtracted and the Poisson shot noise is estimated using  $n_{\text{va}}$  and subtracted from each realization.

We first show the results when the mean Poisson shot noise is subtracted. This case corresponds precisely to the derivation done in Appendix C 2 and briefly summarized in Sec. III A 2. In Fig. 14, we plot the diagonal elements of the covariance for the equilateral triangle configurations. The results are normalized with respect to the Gaussian covariance Eq. 40. We have also plotted the theory prediction, which include the non-Gaussian corrections  $C_{\text{NG}}^{(0)}$ ,  $C_{\text{NG}}^{(1)}$ , and  $C_{\text{NG}}^{(3)'}$ . We find that the deviation of the numerical results from the Gaussian covariance is significant even at low  $k$  for the rare mass groups Large/Small Gr. 4 at  $z = 0.5$  and 1. The predictions underestimate the covariance in comparison with the numerical results. This may be attributed to the fact that we have not included the other terms in the discrete 6-point function  $\mathcal{Y}_d$ , especially the integrated bispectrum terms in  $C_{\text{NG}}^{(3)}$ . On the other hand, for the relatively more abundant groups, the deviation from the Gaussian result is mild. We also find that the predictions agree with the data well, mainly because for these low bias groups the non-Gaussian corrections are small as we have seen in Sec. III A 2.

In Fig. 15, we plot the correlation coefficient for the halo bispectrum with the mean Poisson shot noise subtracted. The triangle configurations are equilateral. One of the wavenumbers,  $k_i$  is fixed to be 0.076, 0.19, and 0.40  $\text{Mpc}^{-1} h$  respectively. The theory predictions are

also overplotted for comparison. By inspecting Eq. C26, it is clear that  $C_{\text{NG}}^{(1)}$  gives no off-diagonal contributions for the equilateral triangle configurations. Similarly,  $C_{\text{NG}}^{(3)'}$  also vanishes for the off-diagonal elements. Thus the only term that contributes to the off-diagonal elements is  $C_{\text{NG}}^{(0)}$  in this case. Similar to the diagonal case, for the rare halo groups at  $z = 1$  and 0.5 the model underestimates the covariance compared to the simulation results, while the model agrees reasonably well with the data for the more abundant groups at  $z = 0$ .

We now consider the scenario when the Poisson shot noise is subtracted from individual realizations. Unlike the previous case, which is equivalent to no shot noise subtraction at all, we have to rely on the accuracy of the Poisson model here. We find that the shot noise subtracted equilateral bispectrum goes negative for  $k \gtrsim 0.35 \text{Mpc}^{-1} h$  at  $z = 1$ . At  $z = 0$ , for Gr. 4, this occurs at  $k \sim 0.21 \text{Mpc}^{-1} h$ . Although there seems to be no fundamental reason that the smooth halo bispectrum must be positive, this may well indicate that the Poisson model is not accurate enough. Therefore we shall not show the results for  $k$  beyond  $0.35 \text{Mpc}^{-1} h$ . In Fig. 16, the diagonal elements of the covariance are plotted. We still normalize the covariance by the Gaussian covariance Eq. 40. Although the prediction does not correspond to this case, we still plot the theory prediction for reference. Similar to the power spectrum case, the subtraction of shot noise from the individual realizations significantly reduce the covariance, that gets much closer to the Gaussian one. For the rare halo groups at  $z = 1$  and 0.5, the covariance is roughly within a factor of 2 and 80% from the Gaussian limit respectively up to  $k \sim 0.35 \text{Mpc}^{-1} h$ . However, we also note that there is still a significant non-Gaussianity contribution even at low  $k \sim 0.05 \text{Mpc}^{-1} h$ . For the groups at  $z = 0$ , the covariance of all the groups is roughly consistent with Gaussianity within 50% or so, although the results are noisy for the Hires groups.

As we can see from Fig. 17, the off-diagonal elements also exhibit substantially lower level of correlation, especially for the groups at  $z = 1$  and 0.5. The correlations are roughly consistent with zero for the groups at  $z = 0.5$  and 0 within the scatter.

Therefore similar to the power spectrum case, we find that when the individual Poisson shot noise is subtracted, the resultant bispectrum covariance is significantly reduced and gets closer to the Gaussian one. This is good news because this means that the modes are less correlated and the halo covariance could be easier to model. However, we note that there is still appreciable amount of non-Gaussianity left. We will leave it to future work to model the halo covariance when the Poisson shot noise is subtracted from each individual realization.

We now comment on the magnitudes of the dark matter and halo bispectrum covariances. For the rare group Gr. 4, depending on the redshift, the covariance of the halo bispectrum is nine to five orders of magnitude larger than that of the dark matter bispectrum covariance. The

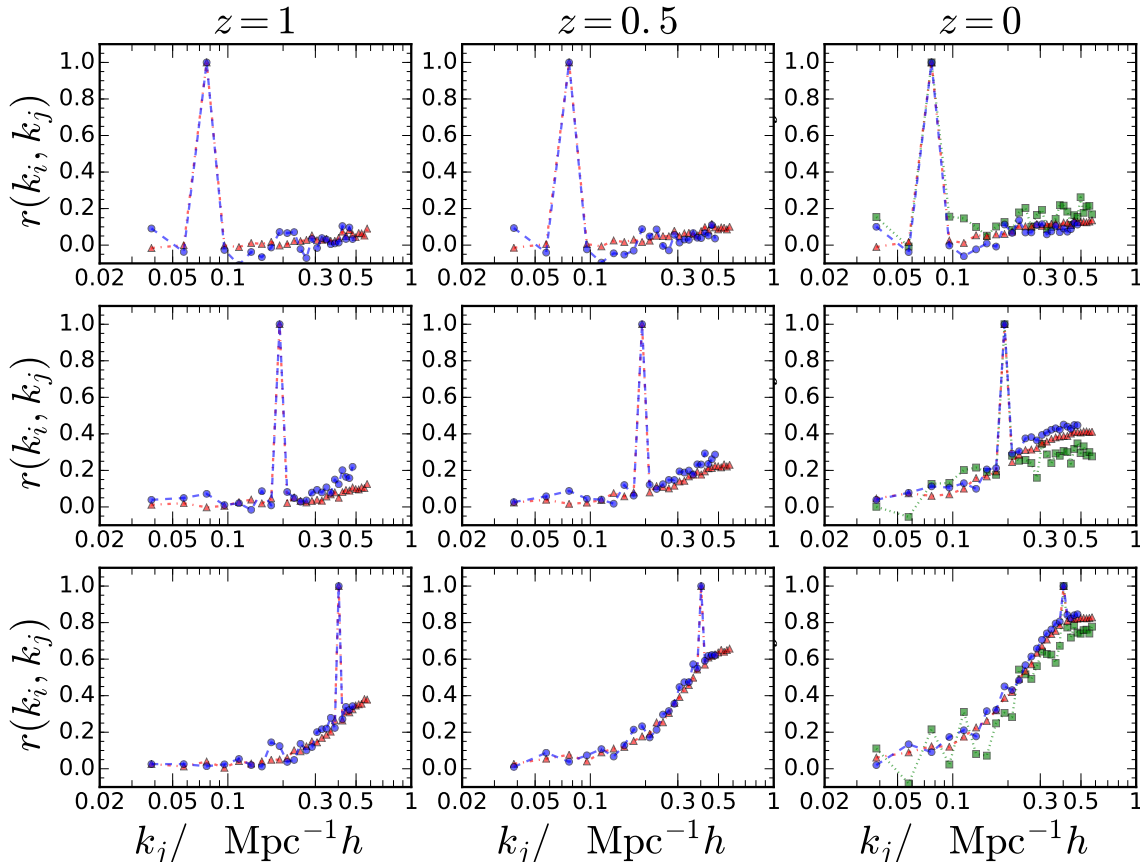


FIG. 13: The correlation coefficient for the dark matter bispectrum. The equilateral triangle configurations are used. Results at  $z = 1, 0.5$ , and  $0$  (left to right columns) are shown. The three rows correspond to the results obtained with  $k_i$  fixed to be  $0.076, 0.19$ , and  $0.40 \text{ Mpc}^{-1} h$  respectively. The results from the simulation set Small (red triangles), Large (blue circles), and Hires (green squares) are shown. Both the Gaussian and the leading disconnected non-Gaussian terms vanish for the off-diagonal elements in this case, hence the predictions are not plotted.

differences decrease when the number density of the sample increases. Even for the abundant group, Hires Gr. 2, the covariance of the halo bispectrum is still an order of magnitude higher than that of the dark matter. When the number density reaches as high as that of the Hires Gr. 1 [ $\bar{n} \sim 10^{-3} (\text{Mpc } h^{-1})^{-3}$ ], the magnitude of the halo covariance is comparable to that of the dark matter. Hence the relative differences are quite similar to the power spectrum covariance. This suggests that shot noise is the predominant contributions to the halo bispectrum covariance.

Gaussian covariance for the bispectrum is often used in forecast e.g. [42–44], our results demonstrate that Gaussian approximation for the bispectrum covariance is a poor approximation in the mildly nonlinear regime at low redshifts. Use of the Gaussian covariance significantly overestimate the signal-to-noise, which we will study in more details in Sec. IV.

#### IV. THE INFORMATION CONTENT OF THE POWER SPECTRUM AND BISPECTRUM

As the power spectrum has been well explored, and the bispectrum becomes the next frontier in large scale structure, it is crucial to address how much information one can gain by going to higher order correlators. The signal-to-noise ratio,  $S/N$ , is often used to quantify the information content in the power spectrum and bispectrum, e.g. [13, 15, 18, 29]. The Fisher information [45] is an alternative way to characterize the information content. However, we shall not consider it here as it involves the derivatives of the polyspectra with respect to the cosmological parameters, that would require a dedicated set of simulations with varying cosmological parameters.

The Gaussian covariance is usually used to make forecast and quantify the information content. In particular, the forecast based on Gaussian covariance suggests that there is a lot of information in the bispectrum ([29], see also below). However, since we have seen in the pre-



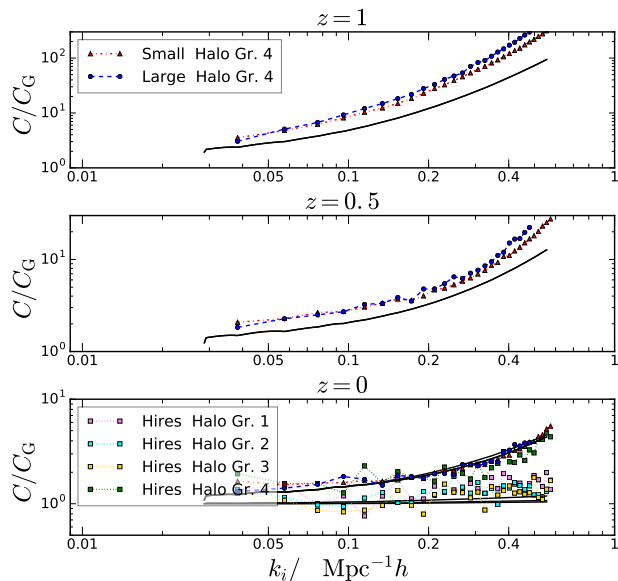


FIG. 14: Similar to Fig. 12, except for the halo bispectrum. The mean Poisson shot noise is subtracted. The halo groups from simulation set Large (blue circles), Small (red triangles), and Hires (squares: violet (group 1), cyan (group 2), yellow (group 3), and green (group 4)). The predictions (solid black lines) include  $C_{\text{NG}}^{(0)}$ ,  $C_{\text{NG}}^{(1)}$ , and  $C_{\text{NG}}^{(3)'}$ .

vious sections that the non-Gaussian contributions significantly increase the covariance already in the weakly nonlinear regime, it is important to ask how the results are modified when the non-Gaussian covariance is taken into account.

The signal-to-noise ratio is defined as

$$(S/N)^2 = S_i C^{-1}_{ij} S_j, \quad (44)$$

where  $C^{-1}$  is the inverse of the covariance matrix, also called the precision matrix, and  $S$  is the signal. Here we use the non-Gaussian covariance measured from simulations to quantify the information content in the power spectrum and bispectrum. For dark matter, the signal  $S$  is simply the mean of the measured power spectrum and bispectrum respectively. For halos, we use the power spectrum and bispectrum with the Poisson shot noise subtracted, using Eq. C6 and C10 respectively, with the Poisson shot noise estimated and subtracted from each realization. At high  $k$  the Poisson shot noise subtracted power spectrum and bispectrum can go negative, which is deemed unphysical, hence we shall only show the results that are reliable. We also show the results obtained using the Gaussian covariance for comparison.

There is one more complication because we require the precision matrix rather than the covariance matrix. Ref. [46] pointed out that for a  $p \times p$  covariance matrix,  $p$  must be smaller than the number of realizations,  $n$ , for the covariance matrix estimated from the realiza-

tions to be invertible. The basic reason is that we can regard each realization as an independent random vector in  $p$ -dimensional space. When  $n$  is larger than  $p$ , these vectors are sufficient to span the space of dimension  $p$ , otherwise the rank of the covariance matrix is less than  $p$ . See [46, 47] for a formal proof. Moreover, even when the covariance matrix is invertible, because inversion is a nonlinear operation, the inverse of an unbiased estimator is in general biased. If the distribution of the estimator is Gaussian, the bias-corrected estimator for the precision matrix reads [46, 47]

$$C_{\text{unbiased}}^{-1} = \frac{n-p-2}{n-1} C_{\text{sample}}^{-1}, \quad (45)$$

where  $C_{\text{sample}}$  is the unbiased sample covariance matrix.

For the power spectrum the number of bins,  $p$ , is typically smaller than the number of realizations available in our case, e.g. we use 90 bins for power spectrum below. Ref. [48] checked that Eq. 45 works very well for the power spectrum precision matrix when  $n/p \gtrsim 2$ . On the other hand, because of the large number of configurations, the number of bins of the measured bispectrum can be comparable to or even larger than the number of realizations that we use. For example, for the binning of the Small simulation set that we used in the previous sections, there are 2825 bins.<sup>2</sup> As Eq. 45 assumes that the distribution of the bispectrum estimator is Gaussian, a priori it is not clear how good it works for the bispectrum covariance. In Appendix A, we check that the distribution of the bispectrum estimator follows the Gaussian distribution reasonably well, although there are also some non-negligible deviations. To test Eq. 45, we shall use the Small simulation set as it has the largest number of realizations. Because the key parameter in Eq. 45 is the ratio  $n/p$ , we rebin the bispectrum into wider bins so that the sides of the bispectrum are in unit of  $8k_F$  instead of  $2k_F$  that we have been using so far. After rebinning, there are altogether  $p = 429$  bins of bispectrum configurations. As the correction is only an overall factor, following [46], we plot the trace of the precision matrix against  $n/p$  in Fig. 18. We show the dark matter and halo data from the Small set at  $z = 0$ . Each data point in this figure corresponds to the results obtained with a precision matrix estimated by randomly choosing  $n$  realizations from the total 4096 ones. We compare the results from the naive estimate  $C_{\text{sample}}^{-1}$  and the bias-corrected estimate  $C_{\text{unbiased}}^{-1}$ . As  $n$  increases, the naive estimate decreases and approaches the bias corrected estimate. For  $n/p \gtrsim 2$ , the correction works very well already. Note that for  $n/p \lesssim 1$  we encounter difficulties in

<sup>2</sup> We can compute the total number of distinct triangle configurations (including the folded ones) with the formula  $\sum_{j=1}^n \left[ \frac{j+1}{2} \right] \left[ \frac{j+1}{2} \right] + \left[ \frac{j}{2} \right]$ , where  $\lceil \cdot \rceil$  and  $\lfloor \cdot \rfloor$  are the ceiling and floor functions. This formula is derived from the integer sequence A002620 in <https://oeis.org/A002620> (see also [49]). In this case we use  $n = 30$  bins, and we get 2825 configurations.

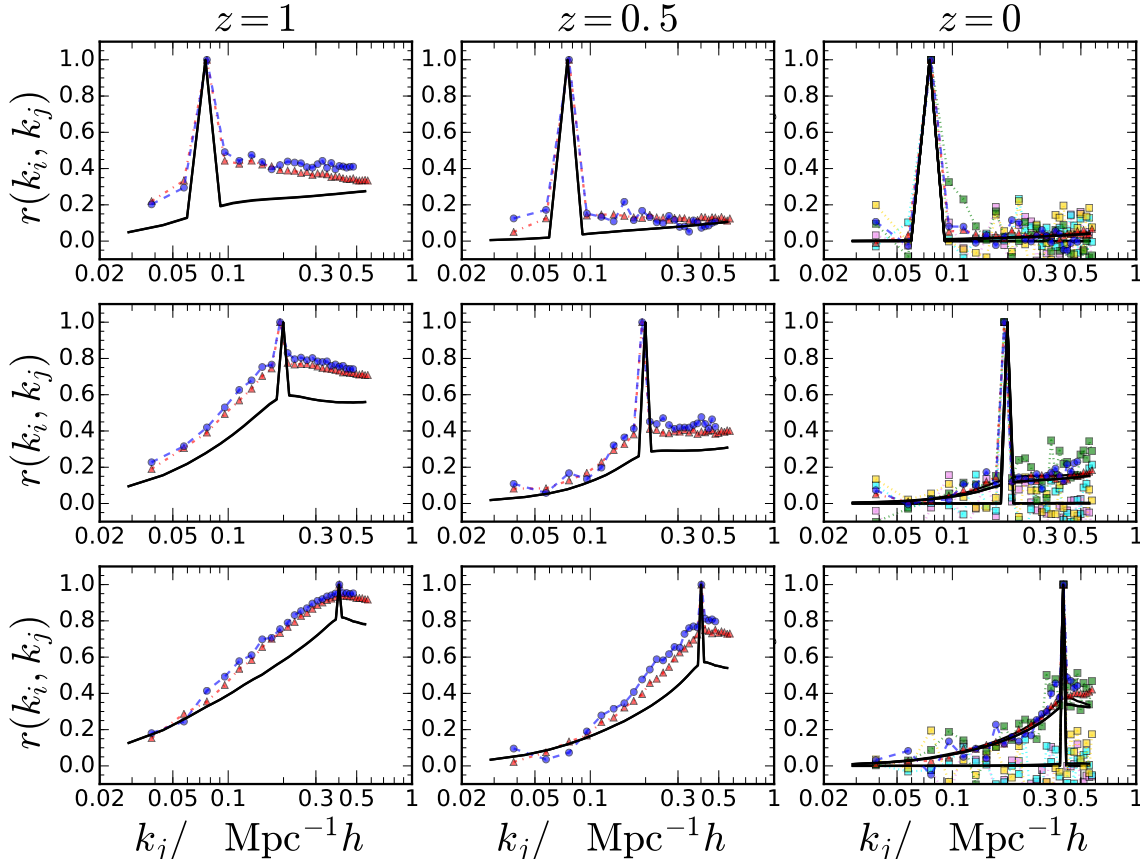


FIG. 15: Similar to Fig. 13, except for the halo bispectrum. The mean Poisson shot noise is subtracted. The halo groups from simulation set Large (blue circles), Small (red triangles), and Hires (squares; violet (group 1), cyan (group 2), yellow (group 3), and green (group 4)). The predictions (solid black lines) include  $C_{\text{NG}}^{(0)}$ ,  $C_{\text{NG}}^{(1)}$ , and  $C_{\text{NG}}^{(3)}$ .

inverting the covariance matrix due to the reason mentioned above. As Eq. 45 works very well, we shall use the bias corrected precision matrix for both power spectrum and bispectrum.

Finally, we are ready to present the S/N of the power spectrum and bispectrum. We will only show the results from the Small set due to its large number of realizations available. However, we will also comment on the results from Hires. In Fig. 19, we plot the S/N for both dark matter and halo power spectrum against the maximum  $k$  used to compute S/N,  $k_{\text{max}}$ . For the power spectrum, the S/N does not depend on the binning width  $\Delta k$  used to the lowest order. Of course there is some binning dependence if the field varies appreciably across the bin, but this dependence is of higher order in the binning width. Similar statement was also stated in [13]. We verify it in Appendix D.

We find that for the power spectrum of dark matter, the information content increases as  $k_{\text{max}}$  increases in the linear regime. The S/N then starts to level off at the weakly nonlinear regime  $k_{\text{max}} \sim 0.2 \text{Mpc}^{-1} h$ . The

flattening of the S/N becomes more and more serious as the redshift decreases. In particular at  $z = 0$ , there is almost no increase in S/N beyond  $k_{\text{max}} \sim 0.2 \text{Mpc}^{-1} h$ . In contrast, the S/N obtained with Gaussian covariance keeps on increasing with the phase volume. The saturation of the information content of the dark matter  $P(k)$  in the weakly nonlinear regime has been observed by many authors [13, 15, 50–52]. This casts doubts on the efforts to model the nonlinear power spectrum accurately beyond the BAO scales and has motivated using alternative statistics to extract information from the large scale structure, notably the log transform [53–55]. Nevertheless, Fisher analysis seems to suggest that the information content on cosmological parameters is not completely erased in the nonlinear regime [48].

We find that the Poisson shot noise subtracted halo power spectrum goes negative for  $k$  in between 0.5 and  $0.8 \text{Mpc}^{-1} h$  for Gr. 4. Because the power spectrum must be non-negative, this is mathematically inconsistent. This typically happens when the signal is so low that the theoretical uncertainty of the Poisson model is

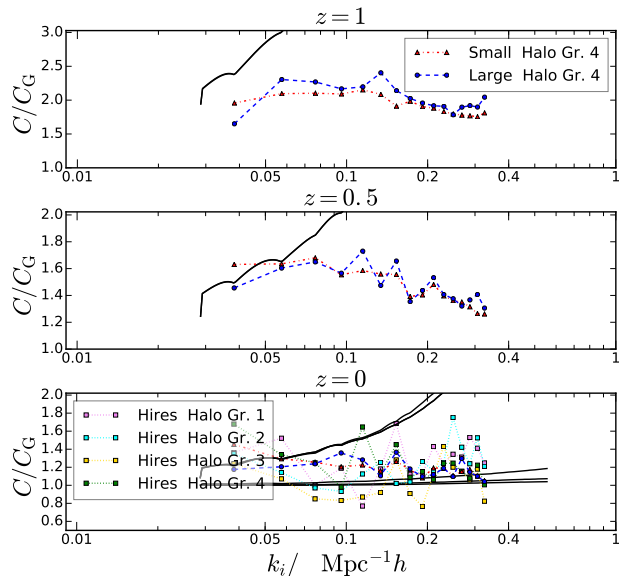


FIG. 16: Similar to Fig. 14, except the shot noise estimated from individual realization is subtracted. The predictions (solid black lines) are the same as those in Fig. 14.

larger than or comparable to the signal. As we believe the contribution to the cumulative signal-to-noise from this range of data is negligible, we will show it as well. For the halo power spectrum, the trends are qualitatively similar to that of the dark matter and they reach roughly constant at  $k_{\max} \sim 0.1 - 0.2 \text{ Mpc}^{-1} h$  depending on the number density of the sample. The saturation of the information content of the halo power spectrum is also hinted in [12, 30]. We find that the signal-to-noise is in between a factor of a few to one order of magnitude lower than that of the dark matter for the rare halo group Gr. 4. Reassuringly for the more abundant groups from Hires, the S/N is comparable to that of the dark matter. We find that the Gaussian approximation overestimates the S/N by a factor two to a few at  $k_{\max} \sim 0.4 \text{ Mpc}^{-1} h$ .

We now turn to the S/N for the bispectrum. We consider all the triangle configurations with the sides of the triangle less than certain  $k_{\max}$ , which is plotted against in Fig. 19. In order to sample the low  $k$  modes well, and also to be able to probe the S/N to high  $k$ , we combine the results from two different binning for the bispectrum measurements. For the low  $k$  results we use  $\Delta k = 2k_F = 0.019 \text{ Mpc}^{-1} h$ , while for high  $k$  we bin the bispectrum with  $\Delta k = 8k_F = 0.077 \text{ Mpc}^{-1} h$ . In Appendix D, we verify that the S/N is invariant to rescalings of the binning width to the lowest order. We check this for the contributions to the covariance computed in this work.

For the dark matter bispectrum, already at  $k_{\max} \sim 0.1 \text{ Mpc}^{-1} h$  the S/N increases significantly slower than what the Gaussian predictions suggest. The departure from Gaussian prediction sets in at lower and lower  $k$

as the redshift decreases. We also find that the deviation generally occur at lower  $k$  than the case of power spectrum, thus suggesting that the Gaussian approximation is worse for the bispectrum. At first sight the differences between the Gaussian and non-Gaussian results are much bigger than we expect from previous figures. However from the discussion on the inverse of the covariance matrix, the issue is likely to be more subtle. We note that there is some mismatch between the results from fine binning and coarse binning measurements. As mismatches exist between both the Gaussian and non-Gaussian results, these should partly arise from the poor sampling of the field using a coarse grid. On the other hand, we also find that the mismatch of non-Gaussian one is more serious. This could be because the non-Gaussian covariance matrix is more affected by poor sampling as the off-diagonal elements of the Gaussian one are zeros. A less likely possibility is that there can be some non-Gaussian binning dependence that results in an explicit  $\Delta k$  dependence of the S/N coming from terms that we did not include in the analysis.

The non-Gaussian contribution significantly degrades the S/N for the bispectrum. For example, at  $z = 0$ , according to the Gaussian covariance predictions, the S/N of the dark matter bispectrum should surpass that of the power spectrum at  $k_{\max} = 0.14 \text{ Mpc}^{-1} h$ , while the full non-Gaussian case shows that the S/N of the bispectrum is only 30% of the power spectrum value. However, it is encouraging to find that relative to the dark matter power spectrum, whose S/N already saturates at  $k_{\max} \sim 0.2 \text{ Mpc}^{-1} h$ , the S/N of the bispectrum still keeps on increasing mildly up to  $k_{\max} \sim 1 \text{ Mpc}^{-1} h$ . Interestingly, at  $k_{\max} \sim 1 \text{ Mpc}^{-1} h$ , the S/N increases sharply and overshoots the S/N of the power spectrum. Thus by delving into the nonlinear regime, the information gain of the bispectrum is higher than that from the power spectrum. We note that the information content of the dark matter power spectrum was found to increase sharply at  $k_{\max} \sim 1.5 \text{ Mpc}^{-1} h$  beyond the plateau [50, 51]. We suspect that this sharp rise in bispectrum S/N also due to the same reason. However, we shall not investigate this further here.

After the Poisson shot noise subtraction, the halo bispectrum of Gr. 4 at  $z = 1$  starts to go negative for  $k_{\max} \gtrsim 0.35 \text{ Mpc}^{-1} h$ . For Gr. 4 at  $z = 0$ , it happens at even lower  $k$ ,  $k \gtrsim 0.22 \text{ Mpc}^{-1} h$ . This occurs at lower  $k$  than that for the halo power spectrum. This first happens for the triangles of shape close to the equilateral. As  $k_{\max}$  increases, we find that more squeezed shapes also become negative. Although there appears to be no fundamental reason that the bispectrum must be positive, we believe these negative values could well indicate that the Poisson model is not reliable as the bispectrum signal gets too small. To be conservative, we shall not show the halo bispectrum results beyond  $k_{\max} = 0.35 \text{ Mpc}^{-1} h$ . For the halo bispectrum case, the overall trends are also similar to that of the power spectrum. Again for these rare groups shown, the Gaussian approximation significantly

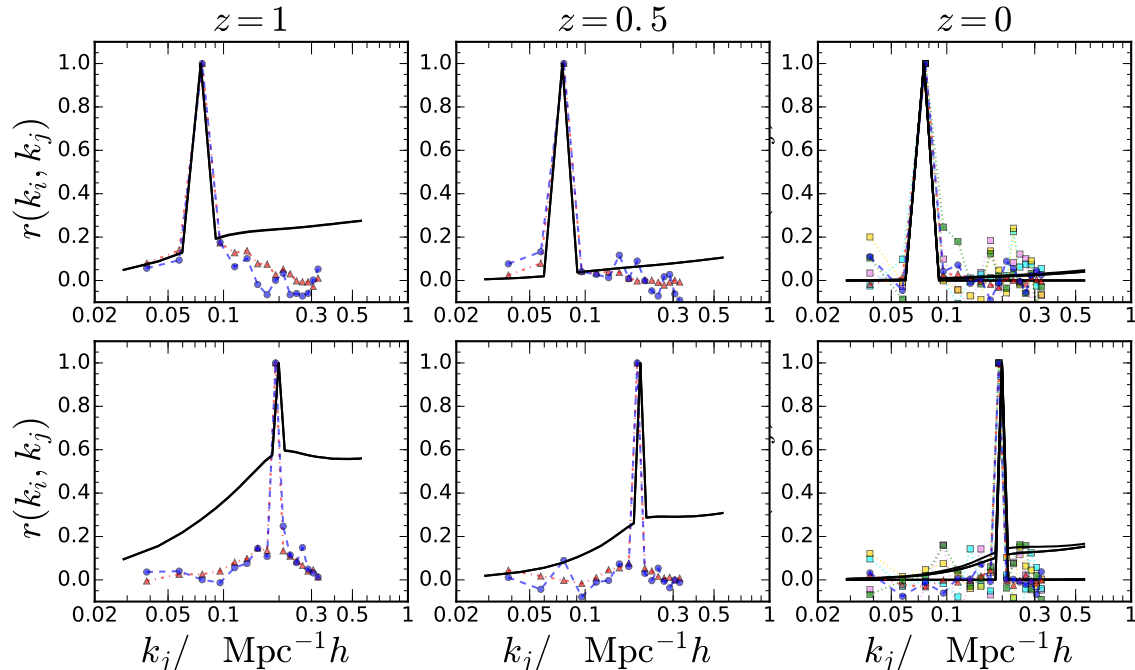


FIG. 17: Similar to Fig. 15, except the shot noise estimated from individual realization is subtracted. The predictions (solid black lines) are the same as those in Fig. 15. Only the results for  $k_i = 0.076$  and  $0.19 \text{ Mpc}^{-1} h$  are shown.

overestimates the S/N, even more seriously than for the dark matter bispectrum. At  $k_{\text{max}} \sim 0.35 \text{ Mpc}^{-1} h$ , the S/N can be overestimated by as much as an order of magnitude. However, the Gaussian approximation gets better for the more abundant groups, and the overestimation is narrowed to a factor of a few at  $k_{\text{max}} \sim 0.35 \text{ Mpc}^{-1} h$ .

We also show the S/N obtained by combining power spectrum and bispectrum measurements. Here we simply add up the S/N for the power spectrum and bispectrum together ignoring the covariance between them. This may not be a good approximation as both the power spectrum and bispectrum are modulated by the same large scale modes, hence leading to beat-coupling effects between the power spectrum and bispectrum [17]. However, the effect is expected to be a small correction to our main results here. With this caveat in mind, we find that the total S/N is only larger than the S/N of  $P(k)$  by a small amount, and the amount is much less than that suggested by Gaussian covariance approximation. Our results suggest that adding the bispectrum information to the power spectrum only improves the total information content mildly.

In cosmology, we are ultimately interested in how good the measurements of the polyspectra can put constraints on the cosmological parameters, and this can be estimated using the Fisher analysis. Furthermore, it has previously been shown that the Fisher analysis results may not be easily interpreted from a simpler signal-to-noise analysis. For example, ref. [56] found that the power

spectrum can still place strong constraint on the parameters which are not sensitive to the amplitude of the power spectrum in the nonlinear regime (see also [48]). In the context of weak lensing, based on Fisher analysis ref. [57] obtained stronger constraints on the cosmological parameters using bispectrum than the power spectrum, even though the signal-to-noise ratio of the bispectrum is lower than that of the power spectrum. We will leave a Fisher analysis for combined power spectrum and bispectrum for future work.

## V. CONCLUSIONS

Gaussian covariance is often assumed in making forecast. In this paper, we have used a large suite of simulations from DEUS-PUR project to study the covariance of the power spectrum and bispectrum, paying special attention to quantifying the effects of the non-Gaussian contributions to the covariance.

This work is the first to use such a large number of simulations (altogether 4704) to estimate the covariance of the bispectrum. We find that the diagonal elements of the covariance of the dark matter bispectrum already deviate from the Gaussian covariance at  $k \sim 0.1 \text{ Mpc}^{-1} h$  by 10 % at  $z = 0$ . The correlation increases as the redshift decreases and we find that at  $z = 0$  the correlation coefficient  $r(k_i, k_j)$  is within 20% if  $k_i$  and  $k_j$  are less than  $0.2 \text{ Mpc}^{-1} h$  for the equilateral triangle config-

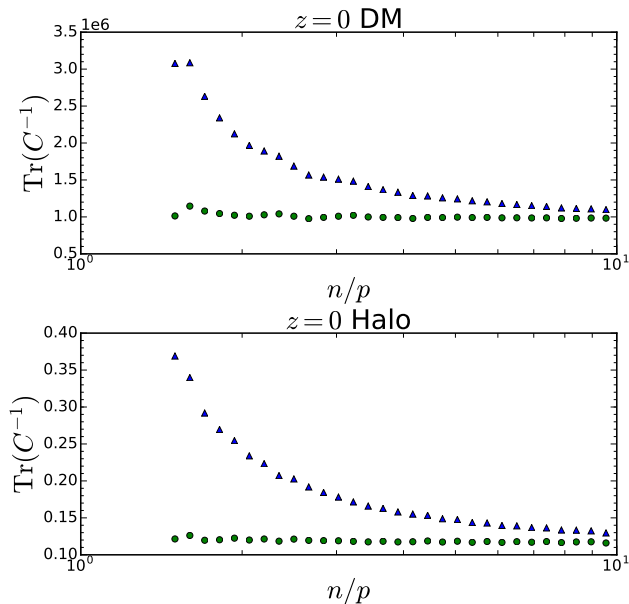


FIG. 18: The trace of the precision matrix as a function of  $n/p$ , where  $n$  is the number realizations used to estimate the covariance matrix and  $p$  is the number of the bispectrum bins, which is equal to 429 here. Both the results from dark matter and halo (Gr. 4) at  $z = 0$  are shown. The results from the naive estimate  $C_{\text{sample}}^{-1}$  (blue triangles) and bias-corrected estimate  $C_{\text{unbiased}}^{-1}$  (green circles) are compared.

urations. The covariance of the dark matter bispectrum significantly increases in the mildly nonlinear regime. To compare with the simulation results, we have also computed the leading disconnected non-Gaussian corrections in the 6-point function. Because the integrals involved are high dimensional yet narrowly peaked, conventional methods would not be able to integrate them. To this goal, we have developed an algorithm to efficiently sample the points that satisfy the Dirac delta function constraints. Including these non-Gaussian corrections we find that the predictions give good agreement with the simulation results in the weakly nonlinear regime. For the equilateral triangle configurations, the diagonal term agrees with the simulation results up to  $k \sim 0.3 \text{ Mpc}^{-1} h$  at  $z = 1$  and  $0.16 \text{ Mpc}^{-1} h$  at  $z = 0$ .

We have also studied the covariance matrix of the halo power spectrum and bispectrum. We have distinguished between the case when the mean Poisson shot noise is subtracted and the Poisson shot noise is estimated and subtracted from each realization. On the theory side we have used the Poisson model to derive the Poisson shot noise contribution to the covariance of the power spectrum and bispectrum. The model corresponds only to the scenario in which the mean Poisson shot noise is subtracted. On the other hand, in simulations or observations, we often have to estimate the shot noise using the volume averaged density. When the mean Poisson

shot noise is subtracted, we find that for the diagonal elements of the power spectrum covariance, the Poisson model yields agreement with the Small/Hires simulation results within 50% up to  $k \sim 1 \text{ Mpc}^{-1} h$ . For the off-diagonal elements, the model only reproduces the correlations at the scales close to the pivot scale, but over-predict the results for larger  $k$ . For the bispectrum, the model underestimates the covariance of the high bias halo groups, while for the low bias groups the model is consistent with the measurements within the scatter. The reason can be because some of the bispectrum contributions, which are negligible for the low bias groups but non-negligible for the high mass groups, are not included in the calculations.

We find that for both the halo power spectrum and bispectrum when the Poisson shot noise is subtracted in each realization, the resultant covariance is substantially reduced and gets closer to the Gaussian limit. It makes sense because part of the fluctuation is absorbed in the shot noise in this case. For example, at  $z = 0$ , the diagonal elements of the bispectrum covariance for the equilateral triangle configurations is roughly within 50% from the Gaussian covariance up to  $k \sim 0.35 \text{ Mpc}^{-1} h$ , and the cross correlation coefficients are within 20% when  $k_i$  and  $k_j$  are less than  $0.2 \text{ Mpc}^{-1} h$ . However we note that there is still significant amount of non-Gaussianity left. We will leave for future work to model the covariance in this case.

We note that the magnitudes of the halo power spectrum and bispectrum covariances are generally higher than those of the dark matter ones. The magnitudes of the halo power spectrum and bispectrum relative to those of the dark matter are quite similar. The shot noise contribution decreases as the number density of halos increases. Even when the number density is  $5 \times 10^{-4} (\text{Mpc } h^{-1})^{-3}$ , the halo covariance is still higher than that of the dark matter by one order of magnitude. Thus matter nonlinearity and halo biasing are not expected to play an important role in this case. Only when the number density is as high as  $10^{-3} (\text{Mpc } h^{-1})^{-3}$ , the covariance of the matter polyspectrum is comparable to that of the halo.

In this work we have largely ignored beat-coupling or super sample variance. We leave for future work to estimate its impact, especially on the bispectrum covariance.

As the power spectrum has been well explored, it is important to ask how much information one can gain by studying the bispectrum in large scale structure. In this work, we have quantified the information content of the power spectrum and the bispectrum with the signal-to-noise ratio S/N using the non-Gaussian covariance matrix measured from simulations.

The S/N of the dark matter power spectrum reaches a plateau in the mildly nonlinear regime. This is because in this regime the covariance increases faster than the signal, causing the S/N to reach a plateau. At  $z = 0$ , it flattens at  $k_{\text{max}} \sim 0.15 \text{ Mpc}^{-1} h$ . This is in line with the findings of previous works. Similarly, we find that the

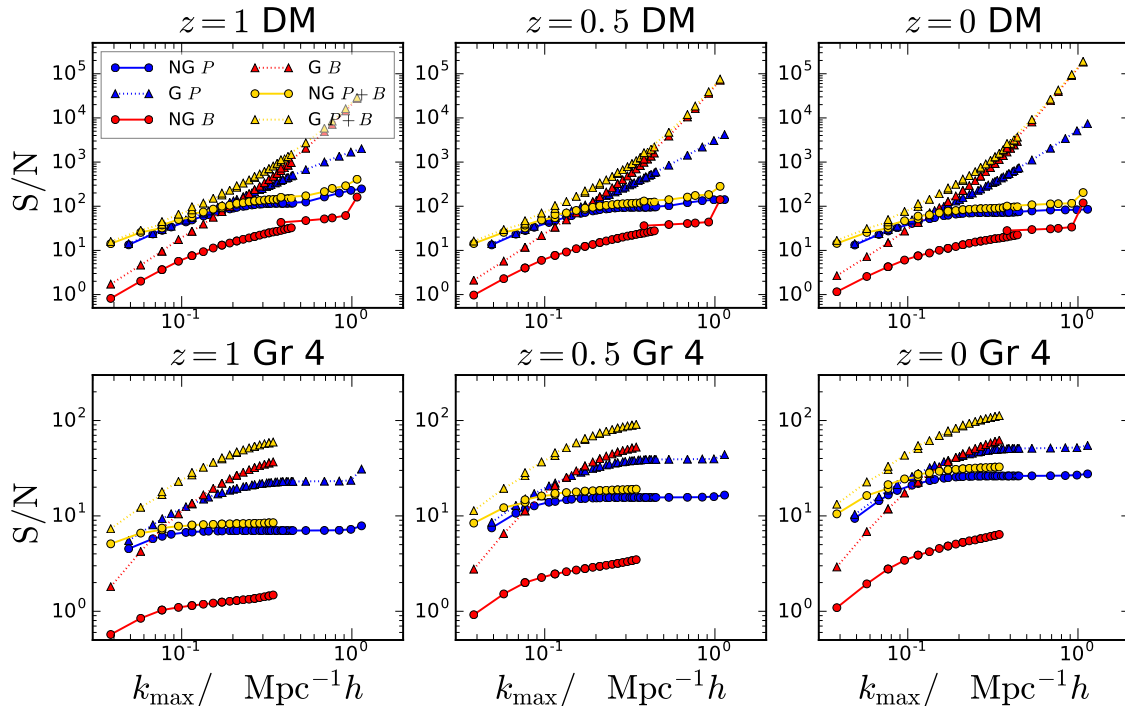


FIG. 19: The signal-to-noise ratio for dark matter and halo power spectrum and bispectrum. Dark matter and halo group 4 data from Small simulation set are used. The signal-to-noise of the power spectrum (blue), bispectrum (red) and, the sum of them (yellow) are shown. The results obtained with the full non-Gaussian covariance results (circles, solid line) are compared with the Gaussian covariance (triangles, dotted line) ones.

S/N of the halo power spectrum flattens in the regime  $k_{\max} \sim 0.1 - 0.2 \text{ Mpc}^{-1} h$  depending on the number density of the samples. The S/N of the halo power spectrum increases as the number density of the sample increases. We find that at  $z = 0$  only the samples with number density  $\gtrsim 5 \times 10^{-4} (\text{Mpc } h^{-1})^{-3}$  yields the S/N comparable to that of dark matter. In contrast, the Gaussian covariance approximation overestimates the S/N by a factor of two to a few at  $k_{\max} \sim 0.4 \text{ Mpc}^{-1} h$ , depending on the number density of the sample.

For the bispectrum, we have computed S/N using all the triangle configurations with lengths less than certain  $k_{\max}$ . For the case of dark matter bispectrum, the S/N increases much slower than the Gaussian approximation suggests. For example, at  $z = 0$ , the S/N at  $k_{\max} = 0.2 \text{ Mpc}^{-1} h$  is an order of magnitude lower than the Gaussian result. Although the Gaussian covariance suggests that the S/N of the matter bispectrum surpasses that of the matter power spectrum at  $k_{\max} \sim 0.14 \text{ Mpc}^{-1} h$  at  $z = 0$ , using the non-Gaussian covariance we find that the S/N is only 30% of the Gaussian one at this scale. In the nonlinear regime, the S/N of the dark matter bispectrum is still mildly increasing but it stalls at  $k_{\max} \sim 0.4 \text{ Mpc}^{-1} h$ . Interestingly, it starts to increase again sharply at  $k_{\max} \sim 1 \text{ Mpc}^{-1} h$ , and overshoots the S/N of the matter power spectrum.

The S/N of the halo bispectrum shares similar trends as that of the matter bispectrum. The Gaussian covariance approximation significantly overestimates the S/N. We find that the overestimation varies from an order of magnitude for the rare sample [ $n \sim 10^{-5} (\text{Mpc } h^{-1})^{-3}$ ] to a factor of a few for the abundant sample [ $n \sim 5 \times 10^{-4} (\text{Mpc } h^{-1})^{-3}$ ] at  $k_{\max} \sim 0.3 \text{ Mpc}^{-1} h$  for redshift range considered.

We conclude that the bispectrum S/N is degraded more seriously by nonlinearities and shot noise relative to the power spectrum S/N. Thus the bispectrum only adds a small amount of increment to the total S/N when the bispectrum is combined with the power spectrum.

Despite more than a decade of efforts to measure the 3-point statistics in Fourier space [58–61] and configuration space [62–68], the information gain that we get is still modest compared to that from the 2-point statistics. It is well-known that the 3-point statistics are more sensitive to nonlinearities and halo biasing, both the local [69] and nonlocal [70, 71] ones. This is both a blessing and curse. On one side, it is easier to estimate the nonlinear coupling and higher order bias parameters using bispectrum. On the other hand, it suffers from stronger nonlinear effects and is harder to model. In this sense our analysis simply reveals the cursing part that strong nonlinearities cause large information loss.

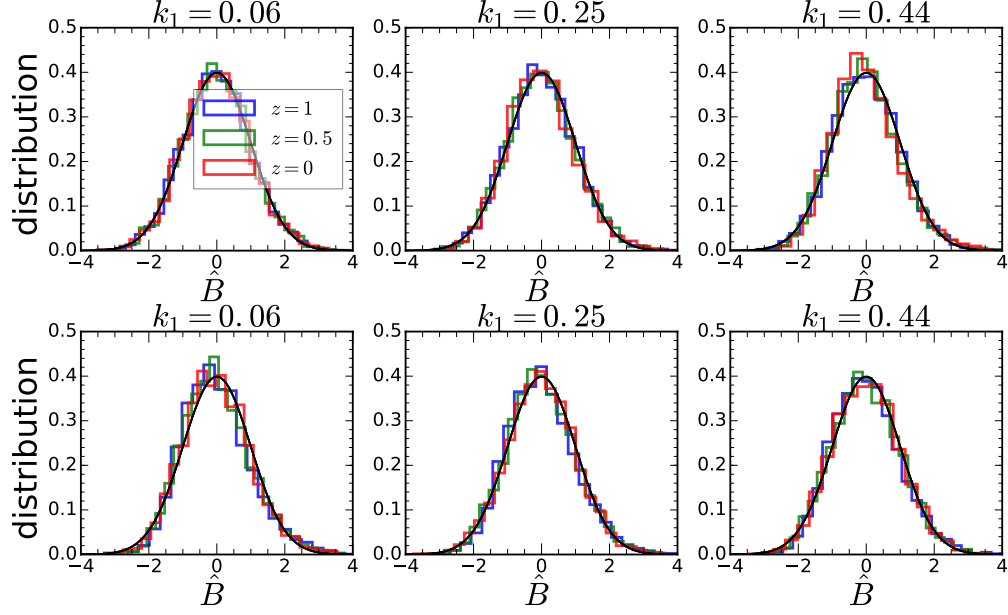


FIG. 20: The distribution of the bispectrum estimator. The results from the Small simulation set at  $z = 1$  (blue), 0.5 (green) and 0 (red) are shown. The upper panels correspond to the results from dark matter while the lower panels are for halo Gr. 4. The equilateral triangle configuration is used. Results at three different wave numbers  $k = 0.06, 0.25$  and  $0.44 \text{ Mpc}^{-1} h$  (from left to right) are displayed. The data have been transformed to the standard variable by Eq. A1. The Gaussian distribution with zero mean and unity variance is overplotted (solid black line).

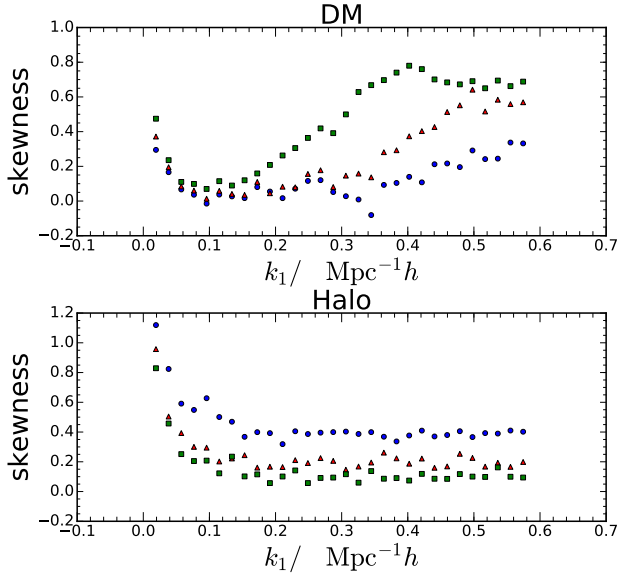


FIG. 21: The skewness of the bispectrum estimator. The results from the dark matter (upper panel) and halo Gr. 4 (lower panel) of the Small simulation set are shown. Only equilateral triangle configuration is used. The results from  $z = 1$  (blue circles), 0.5 (red triangles) and 0 (green squares) are compared.

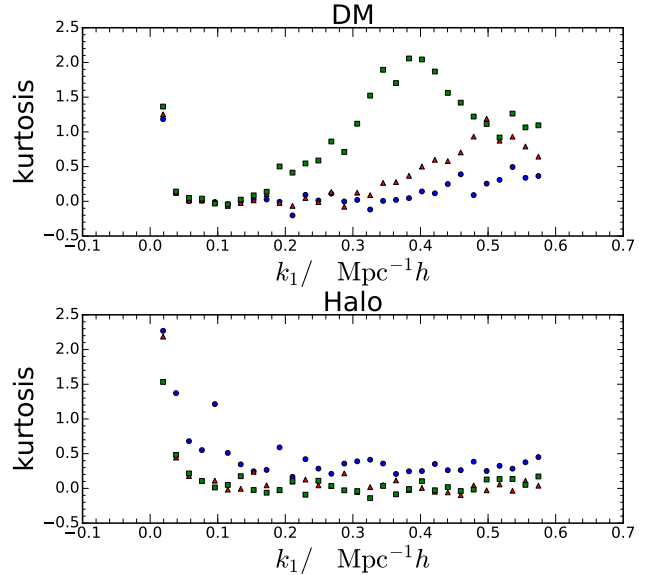


FIG. 22: Similar to Fig. 21 except for kurtosis.

Given the low S/N of the bispectrum, it is not very useful to constrain the cosmological parameters alone. However, there are subtle effects for which the bispectrum analysis is particularly useful. When the power spectrum is combined with the bispectrum, some degeneracies can

be broken. For example the degeneracy between the linear bias  $b_1$  and the growth rate can be broken when the halo power spectrum is combined with the halo bispectrum. The bispectrum is also an important tool to constrain primordial non-Gaussianities. These subtle effects are not reflected in the general signal-to-noise analysis.

However, as we are ultimately interested in how good the polyspectra can constrain the cosmological parameters, the Fisher matrix analysis is more preferable. Previous works found that the power spectrum in the non-linear regime can still constrain some of the cosmological parameters which are not sensitive to the amplitude of the power spectrum [56] and that the weak lensing bispectrum can yield strong constraints on the cosmological parameters even though its S/N is relatively low [57]. We leave the Fisher analysis for future work.

Even though we have only analysed the cases of power spectrum and bispectrum, we speculate that higher points correlators, such as the trispectrum, may suffer of information loss due to nonlinearities and shot noise even more seriously than the Gaussian approximation suggests. If this is true, then it is not a fruitful program to keep on measuring the correlation hierarchy. We can instead consider alternative ways to extract information. Some of the interesting methods include log transformation [53–55], the clustering of voids [72–76] and trying to recover information from the phases of the density field [77–79].

### Acknowledgment

We are grateful to Martín Crocce, Azadeh Moradinezhad, Jorge Noreña, and Ravi Sheth for useful discussions, and Yann Rasera for help with the simulations. We thank the members of the DEUS Consortium<sup>3</sup> for sharing the data with us. This work was granted access to HPC resources of IDRIS through allocations made by GENCI (Grand Équipement National de Calcul Intensif). K.C.C. and L.B. acknowledge the support from the Spanish Ministerio de Economía y Competitividad grant ESP2013-48274-C3-1-P.

### Appendix A: The distribution of the bispectrum estimator

The distribution of the estimator being Gaussian is crucial for many analytic results, e.g. the bias correction formula Eq. 45. Refs. [13, 15] checked that the power spectrum estimator for dark matter follows the Gaussian distribution well. The skewness and kurtosis of the estimator agree with the chi-square distribution, which is a consequence of the underlying density field being

Gaussian. However, [15] also found that the skewness deviates from the chi-square distribution result at low  $z$  for  $k \gtrsim 0.2 \text{ Mpc}^{-1} h$ . In this section, we shall check the distribution of the bispectrum estimator  $\hat{B}$  (Eq. 19).

In Fig. 20, we plot the distribution of the bispectrum estimator. We have transformed the measured data  $\hat{B}_{\text{data}}$  to the standard normal variable as

$$\hat{B}_{\text{snv}} = \frac{\hat{B}_{\text{data}} - \mu}{\sigma}, \quad (\text{A1})$$

where  $\mu$  and  $\sigma$  are the mean and the standard deviation of the data. Here we use the data from the Small simulation set at  $z = 1, 0.5$  and  $0$ , and only the results from the equilateral triangle configuration are shown. Upon comparison with the standard normal distribution, we find that the data across the three redshifts follow the Gaussian distribution reasonably well. Furthermore, the results for both dark matter and halos are similar. The tendency for the distribution to be Gaussian results from the central limit theorem because in the estimator Eq. 19, a large number of modes are averaged over. Nonetheless we note that there are some visible deviations from Gaussianity.

There are two counterinteracting effects at work. First, the central limit theorem works asymptotically for a large number of samples. When  $k$  increases, there are more modes available to be averaged over in the estimator, see Eq. 21. Thus we expect the central limit theorem to perform better for high  $k$ . On the other hand, the underlying density field becomes more non-Gaussian and the modes couple with each other at high  $k$ . This violates the key assumption that the samples are independent in the central limit theorem. Therefore we anticipate that the central limit theorem will fail at both the low  $k$  and high  $k$  regimes.

In order to quantify the deviation from Gaussianity, we compute the sample skewness and kurtosis as

$$S_3 = \frac{\frac{1}{n} \sum_{i=1}^n (\hat{B}_i - \bar{B})^3}{\left[ \frac{1}{n} \sum_{i=1}^n (\hat{B}_i - \bar{B})^2 \right]^{\frac{3}{2}}}, \quad (\text{A2})$$

$$S_4 = \frac{\frac{1}{n} \sum_{i=1}^n (\hat{B}_i - \bar{B})^4}{\left[ \frac{1}{n} \sum_{i=1}^n (\hat{B}_i - \bar{B})^2 \right]^2} - 3. \quad (\text{A3})$$

We plot the skewness in Fig. 21. If the underlying density field is exactly Gaussian, as the skewness is essentially the 9-point function of the underlying density field, it will vanish. Thus the finite value of the skewness is an indication of the deviation of the underlying density field from Gaussianity. For dark matter, at low  $k$  we find that the skewness is large precisely because the number of modes available to estimate the results are small. The skewness first decreases but eventually increases again as  $k$  increases. This is because the mode coupling in the underlying density field increases in the nonlinear regime. The fact that the lower the redshift the larger the value of the skewness supports this interpretation.

<sup>3</sup> <http://www.deus-consortium.org>



For the halo, we again find that the skewness is large at low  $k$ . The skewness increases with redshift and this can be attributed to the fact that there are larger Poisson fluctuations at high redshift due to lower number density. On the other hand, we also find that as  $k$  increases the skewness saturates to some constant value, which is harder to understand.

We present the kurtosis in Fig. 22. For the dark matter bispectrum, similar to the skewness, we find that kurtosis is large at low  $k$ , and then decreases to zero as  $k$  increases,

but eventually increases again at high  $k$ . Curiously, there is a bump in the kurtosis at  $k \sim 0.4 \text{ Mpc}^{-1} h$  at  $z = 0$ . We checked and confirmed that for other shapes, their kurtosis also exhibits a bump. This bump is hard to interpret without a concrete model. Note that there is a small bump in skewness as well, however, this is not present in other shapes. The halo bispectrum kurtosis behaves in a similar way to the skewness.

We have checked that the results are qualitatively similar for other shapes.

---

## Appendix B: A collection of derivations

In this appendix, we present the derivations of some of the formulas used in the main text.

### 1. $V_{123}$

We can analytically integrate  $V_{123}$  (Eq. 21) as

$$\begin{aligned}
 V_{123} &= \int_{k_1} d^3 p \int_{k_2} d^3 q \int_{k_3} d^3 r \delta_{\mathbf{D}}(\mathbf{p} + \mathbf{q} + \mathbf{r}) = \int_{k_1} d^3 p \int_{k_2} d^3 q \int_{k_3} d^3 r \int \frac{d^3 x}{(2\pi)^3} e^{i\mathbf{x} \cdot (\mathbf{p} + \mathbf{q} + \mathbf{r})} \\
 &= \int \frac{d^3 x}{(2\pi)^3} \int_{k_1} dp 4\pi p^2 j_0(px) \int_{k_2} dq 4\pi q^2 j_0(qx) \int_{k_3} dr 4\pi r^2 j_0(rx) \\
 &= \int_{k_1} dp 4\pi p^2 \int_{k_2} dq 4\pi q^2 \int_{k_3} dr 4\pi r^2 \frac{1}{2\pi^2} \int dx x^2 j_0(px) j_0(qx) j_0(rx). \tag{B1}
 \end{aligned}$$

We now make use of an identity for the integral of a product of three spherical Bessel functions [80]

$$\int_0^\infty dr r^2 j_0(k_1 r) j_0(k_2 r) j_0(k_3 r) = \frac{\pi \beta(\Delta)}{4k_1 k_2 k_3}, \tag{B2}$$

where  $\Delta$  and  $\beta(\Delta)$  are defined in Eq. 22 and 23 respectively. Therefore we get

$$V_{123} = 8\pi^2 k_1 k_2 k_3 (\Delta k)^3 \beta(\Delta). \tag{B3}$$

### 2. Gaussian covariance of the bispectrum estimator $\hat{B}$

For Gaussian  $\delta$ , the only non-vanishing contribution to the Gaussian covariance  $C_G^L$  (Eq. 25) reads

$$C_G^L = \int \mathcal{D} \delta_{k_1 k_2 k_3, k'_1 k'_2 k'_3} s_{123} \delta_{\mathbf{D}}(\mathbf{p} + \mathbf{p}') P_L(p) \delta_{\mathbf{D}}(\mathbf{q} + \mathbf{q}') P_L(q) \delta_{\mathbf{D}}(\mathbf{r} + \mathbf{r}') P_L(r). \tag{B4}$$

$\delta_{k_1 k_2 k_3, k'_1 k'_2 k'_3}$  is non-vanishing only if the shape of the triangle  $k_1 k_2 k_3$  is the same as that of  $k'_1 k'_2 k'_3$ . If none of the sides are equal  $s_{123} = 1$ . If the triangles are isosceles,  $s_{123} = 2$ . For equilateral triangle, we have  $s_{123} = 6$ . Both  $\delta_{k_1 k_2 k_3, k'_1 k'_2 k'_3}$  and  $s_{123}$  arise from the fact that the three Dirac delta functions must be satisfied and the number of contractions that they can be satisfied.

Then we can simplify it further as

$$\begin{aligned}
 C_G &= \frac{k_F^6}{V_{123}^2} \delta_{k_1 k_2 k_3, k'_1 k'_2 k'_3} s_{123} \int_{k_1} d^3 p \int_{k_2} d^3 q \int_{k_3} d^3 r \delta_{\mathbf{D}}(\mathbf{p} + \mathbf{q} + \mathbf{r}) \delta_{\mathbf{D}}(\mathbf{0}) P_L(p) P_L(q) P_L(r) \\
 &= \frac{k_F^3}{V_{123}^2} \delta_{k_1 k_2 k_3, k'_1 k'_2 k'_3} s_{123} \int_{k_1} d^3 p \int_{k_2} d^3 q \int_{k_3} d^3 r P_L(p) P_L(q) P_L(r) \int \frac{d^3 x}{(2\pi)^3} e^{i\mathbf{x} \cdot (\mathbf{p} + \mathbf{q} + \mathbf{r})} \\
 &= \frac{k_F^3}{V_{123}^2} \delta_{k_1 k_2 k_3, k'_1 k'_2 k'_3} s_{123} \int \frac{d^3 x}{(2\pi)^3} \int_{k_1} dp 4\pi p^2 j_0(px) \int_{k_2} dq 4\pi q^2 j_0(qx) \int_{k_3} dr 4\pi r^2 j_0(rx) P_L(p) P_L(q) P_L(r) \\
 &\approx \frac{k_F^3}{V_{123}} \delta_{k_1 k_2 k_3, k'_1 k'_2 k'_3} s_{123} P_L(k_1) P_L(k_2) P_L(k_3), \tag{B5}
 \end{aligned}$$

where in the last line we have assumed that the bin is narrow and have taken the power spectra out of the integral. We note that in the first line, there is factor of  $\delta_{\mathbf{D}}(\mathbf{0})$ . This arises from the structure that there are two sets of Dirac delta functions,  $\delta_{\mathbf{D}}(\mathbf{p} + \mathbf{q} + \mathbf{r})$  and  $\delta_{\mathbf{D}}(\mathbf{p}' + \mathbf{q}' + \mathbf{r}')$ , and  $\delta_{\mathbf{D}}(\mathbf{p} + \mathbf{p}')$ ,  $\delta_{\mathbf{D}}(\mathbf{q} + \mathbf{q}')$ , and  $\delta_{\mathbf{D}}(\mathbf{r} + \mathbf{r}')$ . Hence one must be redundant, it results in  $\delta_{\mathbf{D}}(\mathbf{0})$ . We find a similar pattern for the dark matter bispectrum non-Gaussian covariance terms as well.

### 3. Integration domain volume $U$

The volume of the integration domain defined by the Dirac delta functions in the dark matter bispectrum non-Gaussian covariance terms can be computed analytically as

$$\begin{aligned}
U(k_1, k'_1; k_2, k_3, k'_2, k'_3) &\equiv \int_{k_1} d^3p \int_{k_2} d^3q \int_{k_3} d^3r \delta_{\mathbf{D}}(\mathbf{p} + \mathbf{q} + \mathbf{r}) \int_{k'_1} d^3p' \int_{k'_2} d^3q' \int_{k'_3} d^3r' \delta_{\mathbf{D}}(\mathbf{p}' + \mathbf{q}' + \mathbf{r}') \delta_{\mathbf{D}}(\mathbf{p} + \mathbf{p}') \\
&= \delta_{k_1, k'_1} \int_{k_1} d^3p \int_{k_2} d^3q \int_{k_3} d^3r \int_{k'_2} d^3q' \int_{k'_3} d^3r' \delta_{\mathbf{D}}(\mathbf{p} + \mathbf{q} + \mathbf{r}) \delta_{\mathbf{D}}(-\mathbf{p} + \mathbf{q}' + \mathbf{r}') \\
&= \delta_{k_1, k'_1} \int \frac{d^3x}{(2\pi)^3} \int \frac{d^3y}{(2\pi)^3} \int_{k_1} dp 4\pi p^2 j_0(p|\mathbf{x} - \mathbf{y}|) \int_{k_2} dq 4\pi q^2 j_0(qx) \int_{k_3} dr 4\pi r^2 j_0(rx) \\
&\times \int_{k'_2} dq' 4\pi q'^2 j_0(q'y) \int_{k'_3} dr' 4\pi r'^2 j_0(r'y).
\end{aligned} \tag{B6}$$

We can expand  $j_0(|\mathbf{x} - \mathbf{y}|r)$  using the addition theorem for spherical Bessel function (Eq. 10.1.45 in [81])

$$j_0(|\mathbf{x} - \mathbf{y}|r) = 4\pi \sum_{l=0}^{\infty} \sum_{m=-l}^l j_l(xr) j_l(yr) Y_{lm}(\hat{\mathbf{x}}) Y_{lm}^*(\hat{\mathbf{y}}). \tag{B7}$$

After taking the angular integrals of  $\mathbf{x}$  and  $\mathbf{y}$ , we get

$$U = 2^8 \pi \delta_{k_1, k'_1} \int_{k_1} dp p^2 \int_{k_2} dq q^2 \int_{k_3} dr r^2 \int_{k'_2} dq' q'^2 \int_{k'_3} dr' r'^2 \int dx x^2 j_0(px) j_0(qx) j_0(rx) \int dy y^2 j_0(py) j_0(qy) j_0(ry). \tag{B8}$$

Using Eq. B2, we arrive at

$$U(k_1, k'_1; k_2, k_3, k'_2, k'_3) = 2^4 \pi^3 k_2 k_3 k'_2 k'_3 (\Delta k)^5 \beta(\Delta) \beta(\Delta') \delta_{k_1, k'_1}. \tag{B9}$$

The result is the same for  $\delta_{\mathbf{D}}(\mathbf{p} - \mathbf{p}')$ .

### Appendix C: Poisson shot noise contribution to the covariance of the halo power spectrum and bispectrum

In this section, we derive the Poisson shot noise contribution to the covariance of the halo power spectrum and bispectrum. We consider the Poisson model in which the number density of the tracers is given by

$$n(\mathbf{x}) = \sum_i \delta_{\mathbf{D}}(\mathbf{x} - \mathbf{x}_i). \tag{C1}$$

The discrete density contrast  $\delta_{\mathbf{d}}$  is defined as

$$\delta_{\mathbf{d}}(\mathbf{x}) = \frac{n(\mathbf{x})}{\bar{n}} - 1, \tag{C2}$$

where  $\bar{n}$  is the mean number density  $\bar{n} \equiv \langle n(\mathbf{x}) \rangle$ . In this section, all the smooth correlators such as  $\xi$ ,  $P$ ,  $\zeta$ ,  $B$  etc, refer to the nonlinear correlators of the tracers. The tracers can be unbiased such as the dark matter particles in  $N$ -body simulations. Halos are the prototypical example of biased tracers. The Poisson model can be applied to both kinds of tracers.

### 1. Poisson shot noise contribution to the power spectrum covariance

To get the Poisson shot noise contribution to the covariance of the power spectrum we will need the Poisson shot noise contribution to the 2-point and 3-point functions as well. As the computations are similar but less cumbersome, it is instructive to first review the derivations for the 2-point and 3-point functions. One can also include weighting, see Ref.[82], however we will not consider this here. In Ref.[23] the correlators including the Poisson shot noise are derived using an elegant functional method. Ref. [12] applied the Poisson model to compute the shot noise contribution to the covariance of the cross power spectrum between matter and halo. We will compare our results with theirs whenever possible.

The 2-point correlation of the discrete field is given by

$$\xi_d(\mathbf{x}_1, \mathbf{x}_2) = \langle \delta_d(\mathbf{x}_1) \delta_d(\mathbf{x}_2) \rangle = \frac{1}{\bar{n}^2} \langle n(\mathbf{x}_1) n(\mathbf{x}_2) \rangle - 1. \quad (\text{C3})$$

The 2-point correlator of  $n$  can be written as

$$\begin{aligned} \langle n(\mathbf{x}_1) n(\mathbf{x}_2) \rangle &= \left\langle \sum_i \delta_D(\mathbf{x}_1 - \mathbf{x}_i) \delta_D(\mathbf{x}_2 - \mathbf{x}_i) \right\rangle + \left\langle \sum_{i,j} \delta_D(\mathbf{x}_1 - \mathbf{x}_i) \delta_D(\mathbf{x}_2 - \mathbf{x}_j) \right\rangle \\ &= \delta_D(\mathbf{x}_1 - \mathbf{x}_2) \bar{n} + \bar{n}^2 [1 + \xi(|\mathbf{x}_1 - \mathbf{x}_2|)]. \end{aligned} \quad (\text{C4})$$

In this section, all the dummy indices in the summation are unequal. For discrete points, we need to separate the part when two points are the same from the case when the points are different. The latter case can be modelled by the smooth correlation function  $\xi$ . Therefore the discrete correlation function can be written as

$$\xi_d(|\mathbf{x}_1 - \mathbf{x}_2|) = \xi(|\mathbf{x}_1 - \mathbf{x}_2|) + \frac{1}{\bar{n}} \delta_D(\mathbf{x}_1 - \mathbf{x}_2). \quad (\text{C5})$$

Upon Fourier transforming, the discrete power spectrum reads [83]:

$$P_d(k) = P(k) + \frac{1}{\bar{n}}, \quad (\text{C6})$$

where  $P(k)$  is the continuous power spectrum and  $\tilde{n} \equiv (2\pi)^3 \bar{n}$ . This is the well-known shot noise correction for the power spectrum (the presence of  $(2\pi)^3$  is due to the Fourier convention used in this paper).

Similarly the discrete 3-point function reads

$$\zeta_d(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = \langle \delta_d(\mathbf{x}_1) \delta_d(\mathbf{x}_2) \delta_d(\mathbf{x}_3) \rangle = \frac{\langle n(\mathbf{x}_1) n(\mathbf{x}_2) n(\mathbf{x}_3) \rangle}{\bar{n}^3} - \left[ \frac{\langle n(\mathbf{x}_1) n(\mathbf{x}_2) \rangle}{\bar{n}^2} + 2 \text{cyc.} \right] + 2, \quad (\text{C7})$$

where cyc. denotes cyclic permutations. As in Eq. C4, we can express the 3-point correlator of  $n$  as

$$\begin{aligned} &\langle n(\mathbf{x}_1) n(\mathbf{x}_2) n(\mathbf{x}_3) \rangle \\ &= \left\langle \sum_i \delta_D(\mathbf{x}_1 - \mathbf{x}_i) \delta_D(\mathbf{x}_2 - \mathbf{x}_i) \delta_D(\mathbf{x}_3 - \mathbf{x}_i) \right\rangle + \left[ \left\langle \sum_{i,j} \delta_D(\mathbf{x}_1 - \mathbf{x}_i) \delta_D(\mathbf{x}_2 - \mathbf{x}_j) \delta_D(\mathbf{x}_3 - \mathbf{x}_j) \right\rangle + 2 \text{cyc.} \right] \\ &+ \left\langle \sum_{i,j,k} \delta_D(\mathbf{x}_1 - \mathbf{x}_i) \delta_D(\mathbf{x}_2 - \mathbf{x}_j) \delta_D(\mathbf{x}_3 - \mathbf{x}_k) \right\rangle \\ &= \delta_D(\mathbf{x}_1 - \mathbf{x}_2) \delta_D(\mathbf{x}_1 - \mathbf{x}_3) \bar{n} + [\delta_D(\mathbf{x}_2 - \mathbf{x}_3) \bar{n}^2 (1 + \xi_{12}) + 2 \text{cyc.}] + \bar{n}^3 (1 + \xi_{12} + \xi_{23} + \xi_{31} + \zeta), \end{aligned} \quad (\text{C8})$$

where  $\zeta$  is the continuous 3-point function. For convenience, we have used  $\xi_{12}$  to denote  $\xi(|\mathbf{x}_1 - \mathbf{x}_2|)$ , etc.

Using Eq. C4 and C8, we get

$$\zeta_d(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = \frac{1}{\bar{n}^2} \delta_D(\mathbf{x}_1 - \mathbf{x}_2) \delta_D(\mathbf{x}_1 - \mathbf{x}_3) + \left[ \frac{\delta_D(\mathbf{x}_2 - \mathbf{x}_3)}{\bar{n}} \xi_{12} + 2 \text{cyc.} \right] + \zeta_{123}. \quad (\text{C9})$$

In Fourier space we get the discrete bispectrum [23]

$$B_d(k_1, k_2, k_3) = \frac{1}{\bar{n}^2} + \frac{1}{\bar{n}} [P(k_1) + 2 \text{cyc.}] + B(k_1, k_2, k_3), \quad (\text{C10})$$

with  $B$  being the continuous bispectrum.

The discrete 4-point correlation function is given by

$$\begin{aligned} \eta_d(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4) &= \frac{1}{\tilde{n}^4} \langle n(\mathbf{x}_1)n(\mathbf{x}_2)n(\mathbf{x}_3)n(\mathbf{x}_4) \rangle - \left[ \frac{1}{\tilde{n}^3} \langle n(\mathbf{x}_1)n(\mathbf{x}_2)n(\mathbf{x}_3) \rangle + 3 \text{ cyc.} \right] \\ &+ \left[ \frac{1}{\tilde{n}^2} \langle n(\mathbf{x}_1)n(\mathbf{x}_2) \rangle + 5 \text{ cyc.} \right] - 3. \end{aligned} \quad (\text{C11})$$

The 4-point function of  $n$  reads

$$\begin{aligned} &\langle n(\mathbf{x}_1)n(\mathbf{x}_2)n(\mathbf{x}_3)n(\mathbf{x}_4) \rangle \\ &= \tilde{n} \delta_D(\mathbf{x}_1 - \mathbf{x}_2) \delta_D(\mathbf{x}_1 - \mathbf{x}_3) \delta_D(\mathbf{x}_1 - \mathbf{x}_4) + [\delta_D(\mathbf{x}_2 - \mathbf{x}_3) \delta_D(\mathbf{x}_2 - \mathbf{x}_4) \tilde{n}^2 (1 + \xi_{12}) + 3 \text{ cyc.}] \\ &+ [\delta_D(\mathbf{x}_1 - \mathbf{x}_2) \delta_D(\mathbf{x}_3 - \mathbf{x}_4) \tilde{n}^2 (1 + \xi_{13}) + 2 \text{ cyc.}] + [\delta_D(\mathbf{x}_1 - \mathbf{x}_2) \tilde{n}^3 (1 + \xi_{23} + \xi_{24} + \xi_{34} + \zeta_{234}) + 5 \text{ cyc.}] \\ &+ \tilde{n}^4 (1 + \xi_{12} + \xi_{13} + \xi_{14} + \xi_{23} + \xi_{24} + \xi_{34} + \zeta_{123} + \zeta_{124} + \zeta_{134} + \zeta_{234} + \eta_{1234}), \end{aligned} \quad (\text{C12})$$

where  $\eta$  is the continuous 4-point function.

Using Eq. C4, C8, and C12, we get

$$\begin{aligned} \eta_d(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4) &= \frac{1}{\tilde{n}^3} \delta_D(\mathbf{x}_1 - \mathbf{x}_2) \delta_D(\mathbf{x}_1 - \mathbf{x}_3) \delta_D(\mathbf{x}_1 - \mathbf{x}_4) + \left[ \frac{1}{\tilde{n}^2} \delta_D(\mathbf{x}_2 - \mathbf{x}_3) \delta_D(\mathbf{x}_2 - \mathbf{x}_4) \xi_{12} + 3 \text{ cyc.} \right] \\ &+ \left[ \frac{1 + \xi_{13}}{\tilde{n}^2} \delta_D(\mathbf{x}_1 - \mathbf{x}_2) \delta_D(\mathbf{x}_3 - \mathbf{x}_4) + 2 \text{ cyc.} \right] + \left[ \frac{\delta_D(\mathbf{x}_1 - \mathbf{x}_2)}{\tilde{n}} (\xi_{34} + \zeta_{234}) + 5 \text{ cyc.} \right] + \eta_{1234} \text{C13} \end{aligned}$$

In Fourier space, the 4-point correlator

$$\begin{aligned} \mathcal{T}_d(k_1, k_2, k_3, k_4) &= \frac{1}{\tilde{n}^3} + \frac{1}{\tilde{n}^2} (P(k_1) + 3 \text{ cyc.}) + \frac{1}{\tilde{n}^2} [\delta_D(\mathbf{k}_{12}) + P(k_{12}) + \delta_D(\mathbf{k}_{13}) + P(k_{13}) + \delta_D(\mathbf{k}_{14}) + P(k_{14})] \\ &+ \left\{ \frac{1}{\tilde{n}} [\delta_D(\mathbf{k}_{12}) P(k_3) + B(k_{12}, k_3, k_4)] + 5 \text{ cyc.} \right\} + \mathcal{T}(k_1, k_2, k_3, k_4), \end{aligned} \quad (\text{C14})$$

where  $\mathcal{T}$  is the continuous 4-point correlator in Fourier space. Note that  $\mathcal{T}$  is not the trispectrum as it usually refers to the connected part of the 4-point function only, while  $\mathcal{T}$  contains the disconnected part as well. Ref. [23] wrote down the connected trispectrum, which corresponds to the terms without  $\delta_D$  in Eq. C14.

Recall that the covariance of the power spectrum is given by

$$C^P(k, k') = k_F^6 \int_k \frac{d^3 p}{V_s(k)} \int_{k'} \frac{d^3 p'}{V_s(k')} \left[ \langle \delta_d(\mathbf{p}) \delta_d(-\mathbf{p}) \delta_d(\mathbf{p}') \delta_d(-\mathbf{p}') \rangle - \langle \delta_d(\mathbf{p}) \delta_d(-\mathbf{p}) \rangle \langle \delta_d(\mathbf{p}') \delta_d(-\mathbf{p}') \rangle \right]. \quad (\text{C15})$$

The covariance operator in Eq. C15 is now given by

$$\begin{aligned} &\langle \delta_d(\mathbf{p}) \delta_d(-\mathbf{p}) \delta_d(\mathbf{p}') \delta_d(-\mathbf{p}') \rangle - \langle \delta_d(\mathbf{p}) \delta_d(-\mathbf{p}) \rangle \langle \delta_d(\mathbf{p}') \delta_d(-\mathbf{p}') \rangle \\ &= \frac{1}{k_F^3} \left\{ \frac{1}{\tilde{n}^2} [\delta_D(\mathbf{p} + \mathbf{p}') + \delta_D(\mathbf{p} - \mathbf{p}')] + \frac{2}{\tilde{n}} [\delta_D(\mathbf{p} + \mathbf{p}') + \delta_D(\mathbf{p} - \mathbf{p}')] P(p) \right. \\ &+ \frac{1}{\tilde{n}^3} + \frac{1}{\tilde{n}^2} (2P(p) + 2P(p') + P(|\mathbf{p} + \mathbf{p}'|) + P(|\mathbf{p} - \mathbf{p}'|)) \\ &+ \frac{1}{\tilde{n}} [2B(|\mathbf{p} + \mathbf{p}'|, p, p') + 2B(|\mathbf{p} - \mathbf{p}'|, p, p') + B(0, p, p) + B(0, p', p')] \\ &\left. + \mathcal{T}(\mathbf{p}, -\mathbf{p}, \mathbf{p}', -\mathbf{p}') - P(p)P(p') \right\}. \end{aligned} \quad (\text{C16})$$

The first line of the RHS of Eq. C16 are the Gaussian terms. Although we use the terminology ‘‘Gaussian’’ here, these terms are not related to the Gaussian distribution. In fact in the Poisson model, the discrete particles are Poisson distributed. They are called Gaussian because they contribute only to the diagonal covariance as the smooth Gaussian terms. The second and third lines are the non-Gaussian terms and they can couple different bins. The last line is the continuous part of the 4-point function.

We can easily integrate over the Gaussian terms in the first line to get

$$\frac{2k_F^3}{V_s(k)} \delta_{k,k'} \left[ \frac{2P(k)}{\tilde{n}} + \frac{1}{\tilde{n}^2} \right]. \quad (\text{C17})$$

This term can be combined with the Gaussian contribution from the continuous part as

$$C_G^P(k, k') = \frac{2k_F^3}{V_s(k)} \delta_{k, k'} \left[ P(k) + \frac{1}{\tilde{n}} \right]^2. \quad (\text{C18})$$

Eq. C18 agrees with Ref.[31]. In other words, the Gaussian covariance of the power spectrum for the halo is the same as the case for dark matter except with the dark matter power spectrum replaced by the halo one plus the shot noise term.

The non-Gaussian contribution due to the Poisson shot noise is given by

$$\begin{aligned} C_{\text{NG}}^P(k, k') &= k_F^3 \left[ \frac{1}{\tilde{n}^3} + \frac{2}{\tilde{n}^2} (P(k) + P(k')) + \frac{2}{\tilde{n}^2} \int_k \frac{d^3 p}{V_s(k)} \int_{k'} \frac{d^3 p'}{V_s(k')} P(|\mathbf{p} + \mathbf{p}'|) \right] \\ &+ \int_k \frac{d^3 p}{V_s(k)} \int_{k'} \frac{d^3 p'}{V_s(k')} \left\{ \frac{1}{\tilde{n}} [4B(|\mathbf{p} + \mathbf{p}'|, p, p') + B(0, p, p) + B(0, p', p')] \right. \\ &+ \left. \mathcal{T}(\mathbf{p}, -\mathbf{p}, \mathbf{p}', -\mathbf{p}') - P(p)P(p') \right\}. \end{aligned} \quad (\text{C19})$$

Ref. [12] also derived the shot noise contribution to the power spectrum covariance. Comparing Eq. C19 to the result in Ref. [12], besides the minor difference that we have assumed small binning width to simplify the expressions, we note that the terms  $B(0, p, p) + B(0, p', p')$  are missing in [12]. These terms do not vanish in general. Suppose the tree level halo bispectrum is used for  $B(0, p, p)$ , although the local  $b_1$ -term and the nonlocal bias term [70] both vanish because they are generated by large-scale gravitational evolution, the local nonlinear bias term gives finite contribution  $b_1^2 b_2 P^2(p)$ .

## 2. Poisson shot noise contribution to the bispectrum covariance

The complexity of the perturbation series increases rapidly when the number of points in the  $n$ -point function increase. We will first summarize a set of diagrammatic rules to represent the  $n$ -point correlation function for  $n = 2, 3$  and 4. Apart from the Dirac delta function, the rules are similar to the continuous case. In real space, we can represent the Dirac delta function  $\delta_D(\mathbf{x}_i - \mathbf{x}_j)$  by a new link between the points  $\mathbf{x}_i$  and  $\mathbf{x}_j$ . This link is analogous to the continuous correlation function. We can further simplify the diagram by shrinking all the points linked by the Dirac delta functions to a dot. Graphical representations of the 2-point, 3-point, and 4-point correlation functions for Eq. C5, C9, and C13 are shown in Fig. 23. For example the first two diagrams in Fig. 23 denotes the two terms in Eq. C5. The first circle-dot represents the two points connected by a  $\delta_D$ , the wavy line in the second diagram denotes the continuous correlation function  $\xi$ . In the second line, the diagrams represent the three terms in Eq. C9. The first diagram denotes the three points connected by Dirac delta functions. In the second diagram, the circle-dot represents the two points connected by the Dirac delta function and they are connected to the third point by a correlation function. The last one represents the three points connected by the continuous 3-point function. By comparing Eq. C13 with the diagrams for the 4-point in Fig. 23, it is clear that similar rules apply. It is noteworthy that there are disconnected diagrams and the circle-dot can be isolated in the 4-point function.

We now apply the rules to the 6-point function and the results are shown in Fig. 24. They are arranged based on the number of Dirac delta functions, ranging from 5 to 0. From these diagrammatic representations, it is straightforward to write down the discrete 6-point function  $\sigma_d$

$$\begin{aligned} \sigma_d(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4, \mathbf{x}_5, \mathbf{x}_6) &= \frac{1}{\tilde{n}^5} \delta_D(\mathbf{x}_1 - \mathbf{x}_2) \delta_D(\mathbf{x}_1 - \mathbf{x}_3) \delta_D(\mathbf{x}_1 - \mathbf{x}_4) \delta_D(\mathbf{x}_1 - \mathbf{x}_5) \delta_D(\mathbf{x}_1 - \mathbf{x}_6) \\ &+ \frac{1}{\tilde{n}^4} \delta_D(\mathbf{x}_1 - \mathbf{x}_2) \delta_D(\mathbf{x}_1 - \mathbf{x}_3) \delta_D(\mathbf{x}_1 - \mathbf{x}_4) \delta_D(\mathbf{x}_1 - \mathbf{x}_5) \xi_{56} + 5 \text{ cyc.} \\ &+ \frac{1}{\tilde{n}^4} \delta_D(\mathbf{x}_1 - \mathbf{x}_2) \delta_D(\mathbf{x}_1 - \mathbf{x}_3) \delta_D(\mathbf{x}_4 - \mathbf{x}_5) \delta_D(\mathbf{x}_4 - \mathbf{x}_6) \xi_{34} + 9 \text{ cyc.} \\ &+ \frac{1}{\tilde{n}^4} \delta_D(\mathbf{x}_1 - \mathbf{x}_2) \delta_D(\mathbf{x}_1 - \mathbf{x}_3) \delta_D(\mathbf{x}_1 - \mathbf{x}_4) \delta_D(\mathbf{x}_5 - \mathbf{x}_6) \xi_{45} + 14 \text{ cyc.} \\ &+ \frac{1}{\tilde{n}^4} \delta_D(\mathbf{x}_1 - \mathbf{x}_2) \delta_D(\mathbf{x}_3 - \mathbf{x}_4) \delta_D(\mathbf{x}_3 - \mathbf{x}_5) \delta_D(\mathbf{x}_3 - \mathbf{x}_6) + 14 \text{ cyc.} \\ &+ \frac{1}{\tilde{n}^4} \delta_D(\mathbf{x}_1 - \mathbf{x}_2) \delta_D(\mathbf{x}_1 - \mathbf{x}_3) \delta_D(\mathbf{x}_4 - \mathbf{x}_5) \delta_D(\mathbf{x}_4 - \mathbf{x}_6) + 9 \text{ cyc.} \\ &+ \frac{1}{\tilde{n}^3} \delta_D(\mathbf{x}_1 - \mathbf{x}_2) \delta_D(\mathbf{x}_3 - \mathbf{x}_4) \delta_D(\mathbf{x}_5 - \mathbf{x}_6) \zeta_{246} + 14 \text{ cyc.} \end{aligned}$$

$$\begin{aligned}
& + \frac{1}{\tilde{n}^3} \delta_D(\mathbf{x}_1 - \mathbf{x}_2) \delta_D(\mathbf{x}_1 - \mathbf{x}_3) \delta_D(\mathbf{x}_4 - \mathbf{x}_5) \zeta_{346} + 59 \text{ cyc.} \\
& + \frac{1}{\tilde{n}^3} \delta_D(\mathbf{x}_1 - \mathbf{x}_2) \delta_D(\mathbf{x}_1 - \mathbf{x}_3) \delta_D(\mathbf{x}_1 - \mathbf{x}_4) \zeta_{156} + 14 \text{ cyc.} \\
& + \frac{1}{\tilde{n}^3} \delta_D(\mathbf{x}_1 - \mathbf{x}_2) \delta_D(\mathbf{x}_3 - \mathbf{x}_4) \delta_D(\mathbf{x}_5 - \mathbf{x}_6) + 14 \text{ cyc.} \\
& + \frac{1}{\tilde{n}^3} \delta_D(\mathbf{x}_1 - \mathbf{x}_2) \delta_D(\mathbf{x}_3 - \mathbf{x}_4) \delta_D(\mathbf{x}_5 - \mathbf{x}_6) \xi_{24} + 44 \text{ cyc.} \\
& + \frac{1}{\tilde{n}^3} \delta_D(\mathbf{x}_1 - \mathbf{x}_2) \delta_D(\mathbf{x}_1 - \mathbf{x}_3) \delta_D(\mathbf{x}_4 - \mathbf{x}_5) \xi_{16} + 59 \text{ cyc.} \\
& + \frac{1}{\tilde{n}^3} \delta_D(\mathbf{x}_1 - \mathbf{x}_2) \delta_D(\mathbf{x}_1 - \mathbf{x}_3) \delta_D(\mathbf{x}_4 - \mathbf{x}_5) \xi_{56} + 59 \text{ cyc.} \\
& + \frac{1}{\tilde{n}^3} \delta_D(\mathbf{x}_1 - \mathbf{x}_2) \delta_D(\mathbf{x}_1 - \mathbf{x}_3) \delta_D(\mathbf{x}_1 - \mathbf{x}_4) \xi_{56} + 14 \text{ cyc.} \\
& + \frac{1}{\tilde{n}^2} \delta_D(\mathbf{x}_1 - \mathbf{x}_2) \delta_D(\mathbf{x}_5 - \mathbf{x}_6) \eta_{2345} + 44 \text{ cyc.} \\
& + \frac{1}{\tilde{n}^2} \delta_D(\mathbf{x}_1 - \mathbf{x}_2) \delta_D(\mathbf{x}_1 - \mathbf{x}_3) \eta_{1456} + 19 \text{ cyc.} \\
& + \frac{1}{\tilde{n}^2} \delta_D(\mathbf{x}_1 - \mathbf{x}_2) \delta_D(\mathbf{x}_5 - \mathbf{x}_6) \xi_{34} + 44 \text{ cyc.} \\
& + \frac{1}{\tilde{n}^2} \delta_D(\mathbf{x}_1 - \mathbf{x}_2) \delta_D(\mathbf{x}_5 - \mathbf{x}_6) \zeta_{234} + 89 \text{ cyc.} \\
& + \frac{1}{\tilde{n}^2} \delta_D(\mathbf{x}_1 - \mathbf{x}_2) \delta_D(\mathbf{x}_5 - \mathbf{x}_6) \xi_{25} \xi_{34} + 44 \text{ cyc.} \\
& + \frac{1}{\tilde{n}^2} \delta_D(\mathbf{x}_1 - \mathbf{x}_2) \delta_D(\mathbf{x}_5 - \mathbf{x}_6) \xi_{23} \xi_{45} + 89 \text{ cyc.} \\
& + \frac{1}{\tilde{n}^2} \delta_D(\mathbf{x}_1 - \mathbf{x}_2) \delta_D(\mathbf{x}_1 - \mathbf{x}_3) \zeta_{456} + 19 \text{ cyc.} \\
& + \frac{1}{\tilde{n}^2} \delta_D(\mathbf{x}_1 - \mathbf{x}_2) \delta_D(\mathbf{x}_1 - \mathbf{x}_3) \xi_{34} \xi_{56} + 59 \text{ cyc.} \\
& + \frac{1}{\tilde{n}} \delta_D(\mathbf{x}_1 - \mathbf{x}_2) \chi_{23456} + 14 \text{ cyc.} \\
& + \frac{1}{\tilde{n}} \delta_D(\mathbf{x}_1 - \mathbf{x}_2) \xi_{34} \xi_{56} + 44 \text{ cyc.} \\
& + \frac{1}{\tilde{n}} \delta_D(\mathbf{x}_1 - \mathbf{x}_2) \xi_{26} \zeta_{345} + 59 \text{ cyc.} \\
& + \frac{1}{\tilde{n}} \delta_D(\mathbf{x}_1 - \mathbf{x}_2) \eta_{3456} + 14 \text{ cyc.} \\
& + \frac{1}{\tilde{n}} \delta_D(\mathbf{x}_1 - \mathbf{x}_2) \zeta_{234} \xi_{56} + 89 \text{ cyc.} \\
& + \sigma_{123456}, \tag{C20}
\end{aligned}$$

where  $\chi$  and  $\sigma$  are the continuous 5-point and 6-point functions. Note that some of the diagrams correspond to more than one term in Eq. C20.

Then in Fourier space, the 6-point function reads

$$\begin{aligned}
& \mathcal{Y}_d(k_1, k_2, k_3, k_4, k_5, k_6) \\
& = \frac{1}{\tilde{n}^5} + [\frac{1}{\tilde{n}^4} P(k_6) + 5 \text{ cyc.}] + [\frac{1}{\tilde{n}^4} P(k_{123}) + 9 \text{ cyc.}] \\
& + [\frac{1}{\tilde{n}^4} P(k_{56}) + 14 \text{ cyc.}] + [\frac{1}{\tilde{n}^4} \delta_D(\mathbf{k}_{56}) + 14 \text{ cyc.}] + [\frac{1}{\tilde{n}^4} \delta_D(\mathbf{k}_{123}) + 9 \text{ cyc.}] \\
& + [\frac{1}{\tilde{n}^3} B(k_{12}, k_{34}, k_{56}) + 14 \text{ cyc.}] + [\frac{1}{\tilde{n}^3} B(k_{123}, k_{45}, k_6) + 59 \text{ cyc.}] + [\frac{1}{\tilde{n}^3} B(k_{1234}, k_5, k_6) + 14 \text{ cyc.}] + \\
& + [\frac{1}{\tilde{n}^3} \delta_D(\mathbf{k}_{12}) \delta_D(\mathbf{k}_{34}) + 14 \text{ cyc.}] + [\frac{1}{\tilde{n}^3} \delta_D(\mathbf{k}_{56}) P(k_{12}) + 44 \text{ cyc.}] + [\frac{1}{\tilde{n}^3} \delta_D(\mathbf{k}_{45}) P(k_6) + 59 \text{ cyc.}] \\
& + [\frac{1}{\tilde{n}^3} \delta_D(\mathbf{k}_{123}) P(k_6) + 59 \text{ cyc.}] + [\frac{1}{\tilde{n}^3} \delta_D(\mathbf{k}_{56}) P(k_6) + 14 \text{ cyc.}]
\end{aligned}$$

$$\begin{aligned}
& + \left[ \frac{1}{\tilde{n}^2} \mathcal{T}(k_{12}, k_3, k_4, k_{56}) + 44 \text{ cyc.} \right] + \left[ \frac{1}{\tilde{n}^2} \mathcal{T}(k_{123}, k_4, k_5, k_6) + 19 \text{ cyc.} \right] \\
& + \left[ \frac{1}{\tilde{n}^2} \delta_{\text{D}}(\mathbf{k}_{12}) \delta_{\text{D}}(\mathbf{k}_{34}) P(k_3) + 44 \text{ cyc.} \right] + \left[ \frac{1}{\tilde{n}^2} \delta_{\text{D}}(\mathbf{k}_{56}) B(k_{12}, k_3, k_4) + 89 \text{ cyc.} \right] \\
& + \left[ \frac{1}{\tilde{n}^2} \delta_{\text{D}}(\mathbf{k}_{34}) P(k_{12}) P(k_3) + 44 \text{ cyc.} \right] + \left[ \frac{1}{\tilde{n}^2} \delta_{\text{D}}(\mathbf{k}_{123}) P(k_3) P(k_4) + 89 \text{ cyc.} \right] \\
& + \left[ \frac{1}{\tilde{n}^2} \delta_{\text{D}}(\mathbf{k}_{123}) B(k_4, k_5, k_6) + 19 \text{ cyc.} \right] + \left[ \frac{1}{\tilde{n}^2} \delta_{\text{D}}(\mathbf{k}_{56}) P(k_4) P(k_6) + 59 \text{ cyc.} \right] \\
& + \left[ \frac{1}{\tilde{n}} \mathcal{X}(k_{12}, k_3, k_4, k_5, k_6) + 14 \text{ cyc.} \right] + \left[ \frac{1}{\tilde{n}} \delta_{\text{D}}(\mathbf{k}_{12}) \delta_{\text{D}}(\mathbf{k}_{34}) P(k_3) P(k_5) + 44 \text{ cyc.} \right] \\
& + \left[ \frac{1}{\tilde{n}} \delta_{\text{D}}(\mathbf{k}_{126}) P(k_6) B(k_3, k_4, k_5) + 59 \text{ cyc.} \right] + \left[ \frac{1}{\tilde{n}} \delta_{\text{D}}(\mathbf{k}_{12}) \mathcal{T}(k_3, k_4, k_5, k_6) + 14 \text{ cyc.} \right] \\
& + \left[ \frac{1}{\tilde{n}} \delta_{\text{D}}(\mathbf{k}_{56}) B(k_{12}, k_3, k_4) P(k_5) + 89 \text{ cyc.} \right] + \mathcal{Y}_{123456}, \tag{C21}
\end{aligned}$$

where  $\mathcal{X}$  and  $\mathcal{Y}$  are the Fourier transform of the continuous 5-point, and 6-point functions. In Ref.[23], the connected 6-point function including the Poisson shot noise, but limited to the bispectrum order, was written down. They agree with the terms without  $\delta_{\text{D}}$  in the first three lines of Eq. C21.

As we can see,  $\mathcal{Y}_{\text{d}}$  contains 2, 3, 4, 5 and 6-point functions in general, it is formidable to evaluate in the most general case. To proceed, in this paper, we will limit ourselves to the terms up to the power spectrum and compute some of the bispectrum contributions, but we will ignore the ones with higher order correlators. The terms that we will consider are highlighted with colors in Eq. C21.

Using the shorthand notation in Eq. 27, the shot noise contribution to the covariance can be written as

$$C = \int \mathcal{D} \text{cov} \mathcal{Y}_{\text{d}}(\mathbf{p}, \mathbf{q}, \mathbf{r}, \mathbf{p}', \mathbf{q}', \mathbf{r}'), \tag{C22}$$

where  $\text{cov} \mathcal{Y}_{\text{d}}$  is defined as

$$\text{cov} \mathcal{Y}_{\text{d}}(\mathbf{p}, \mathbf{q}, \mathbf{r}, \mathbf{p}', \mathbf{q}', \mathbf{r}') = \mathcal{Y}_{\text{d}}(\mathbf{p}, \mathbf{q}, \mathbf{r}, \mathbf{p}', \mathbf{q}', \mathbf{r}') - \frac{1}{k_{\text{F}}^3} B_{\text{d}}(\mathbf{p}, \mathbf{q}, \mathbf{r}) B_{\text{d}}(\mathbf{p}', \mathbf{q}', \mathbf{r}'). \tag{C23}$$

We now classify the terms in  $C$  based on the number of Dirac delta functions. First the terms with two Dirac delta functions in Eq. C21 are the Gaussian terms. They are non-vanishing only if triangle  $k_1 k_2 k_3$  is the same as  $k'_1 k'_2 k'_3$ . There are altogether three such terms in  $\mathcal{Y}_{\text{d}}(\mathbf{p}, \mathbf{q}, \mathbf{r}, \mathbf{p}', \mathbf{q}', \mathbf{r}')$ , and they are highlighted in red in Eq. C21. However, because of the triangle constraint and the length of the triangle sides are finite, all the Dirac delta functions in these three terms must couple one of the vectors in  $k_1 k_2 k_3$  with another one in  $k'_1 k'_2 k'_3$  to give non-vanishing contribution to the covariance. This leaves us with

$$\left\{ \frac{1}{\tilde{n}^3} + \frac{1}{\tilde{n}^2} [P(p) + P(q) + P(r)] + \frac{1}{\tilde{n}} [P(p)P(q) + P(p)P(r) + P(q)P(r)] \right\} [\delta_{\text{D}}(\mathbf{p} + \mathbf{p}') \delta_{\text{D}}(\mathbf{q} + \mathbf{q}') \delta_{\text{D}}(\mathbf{r} + \mathbf{r}') + 5 \text{ cyc.}] \tag{C24}$$

This term can be combined with the continuous Gaussian terms  $[\delta_{\text{D}}(\mathbf{p} + \mathbf{p}') \delta_{\text{D}}(\mathbf{q} + \mathbf{q}') \delta_{\text{D}}(\mathbf{r} + \mathbf{r}') + 5 \text{ cyc.}] P(p) P(q) P(r)$ . Therefore, in the presence of Poisson shot noise, Eq. 25 is modified to

$$C_{\text{G}} = \frac{k_{\text{F}}^3}{V_{123}} \delta_{k_1 k_2 k_3, k'_1 k'_2 k'_3} s_{123} \left[ P(k_1) + \frac{1}{\tilde{n}} \right] \left[ P(k_2) + \frac{1}{\tilde{n}} \right] \left[ P(k_3) + \frac{1}{\tilde{n}} \right]. \tag{C25}$$

Eq. C25 agrees with the results in [40]. Again, similar to the case of power spectrum, the Gaussian covariance of the halo bispectrum including the shot noise contribution can be obtained by replacing the continuous power spectrum with the halo power spectrum plus the shot noise contribution. Similar to the comments for the Gaussian power spectrum covariance, they are called ‘‘Gaussian’’ here simply because they are on the same footing as the ‘‘true’’ Gaussian terms, not because they arise from the Gaussian distribution.

We now look at the terms with one Dirac delta function in  $\mathcal{Y}_{\text{d}}$ . There are altogether 9 such terms, and they are highlighted in green in Eq. C21. Six of these terms can be computed analytically making use of Eq. 39. Among these terms, there are terms with Dirac delta function connecting three vectors,  $\delta_{\text{D}}(\mathbf{p}_{ijk})$ . When three of the vectors are from the same bispectrum estimator, they are exactly cancelled by the corresponding terms in  $\langle B \rangle \langle B' \rangle$ . The net

results due to the terms with one Dirac delta function in  $\text{cov}\mathcal{Y}_d$  are

$$\begin{aligned}
C_{\text{NG}}^{(1)} &= k_{\text{F}}^3 \frac{U(k_1, k'_1)}{V_{123} V'_{123}} \left\{ \frac{2}{\tilde{n}^4} + \frac{1}{\tilde{n}^3} [P(k_2) + P(k_3) + P(k'_2) + P(k'_3)] \right. \\
&+ \frac{1}{\tilde{n}^3} [P(k_1) + P(k_2) + P(k_3) + P(k'_1) + P(k'_2) + P(k'_3)] + \frac{1}{\tilde{n}^3} P(k_1) \\
&+ \frac{1}{\tilde{n}^2} [P(k_1) + P(k'_2) + P(k'_3)] [P(k'_1) + P(k_2) + P(k_3)] \\
&+ \left. \frac{1}{\tilde{n}^2} P(k_1) [P(k_2) + P(k_3) + P(k'_2) + P(k'_3)] \right\} + 8 \text{ cyc.} \\
&+ \frac{1}{\tilde{n}^2} \int \mathcal{D} \delta_{\text{D}}(\mathbf{p} + \mathbf{p}') (P(p) + \frac{1}{\tilde{n}}) \left[ P(|\mathbf{q} + \mathbf{r}|) + P(|\mathbf{q} + \mathbf{q}'|) + P(|\mathbf{q} + \mathbf{r}'|) \right] + 8 \text{ cyc.} \tag{C26}
\end{aligned}$$

The terms without Dirac delta function and containing up to the power spectrum (highlighted in blue in Eq. C21) give

$$C_{\text{NG}}^{(0)} = \frac{k_{\text{F}}^3}{\tilde{n}^5} + \frac{k_{\text{F}}^3}{\tilde{n}^4} [P(k_1) + 5 \text{ cyc.}] + \int \mathcal{D} \frac{1}{\tilde{n}^4} \left\{ [P(|\mathbf{p} + \mathbf{p}'|) + 8 \text{ cyc.}] + [P(|\mathbf{p} + \mathbf{q}|) + 14 \text{ cyc.}] \right\}. \tag{C27}$$

Note that in the last integral, for the first term the permutations include only the terms connecting the two bispectrum estimator, while in the second term, all the permutations are allowed.

The terms including the bispectrum (cyan in Eq. C21) read

$$\begin{aligned}
C_{\text{NG}}^{(2)} &= \frac{k_{\text{F}}^3 U(k_1, k'_1)}{V V'} \left\{ \frac{1}{\tilde{n}^2} [B(k_1, k'_2, k'_3) + B(k'_1, k_2, k_3)] + 8 \text{ cyc.} \right. \\
&+ \frac{1}{\tilde{n}} [(P(k'_1) + P(k_2) + P(k_3)) B(k_1, k'_2, k'_3) + (P(k_1) + P(k'_2) + P(k'_3)) B(k'_1, k_2, k_3)] + 8 \text{ cyc.} \left. \right\} \\
&+ \int \mathcal{D} \frac{1}{\tilde{n}^2} \delta_{\text{D}}(\mathbf{p} + \mathbf{p}') \left[ B(|\mathbf{q} + \mathbf{r}|, q', r') + B(|\mathbf{q} + \mathbf{q}'|, r, r') + B(|\mathbf{q} + \mathbf{r}'|, r, q') \right. \\
&+ \left. B(|\mathbf{r} + \mathbf{q}'|, q, r') + B(|\mathbf{r} + \mathbf{r}'|, q, q') + B(|\mathbf{q}' + \mathbf{r}'|, q, r) \right] + 8 \text{ cyc.} \\
&+ \int \mathcal{D} \frac{1}{\tilde{n}} \delta_{\text{D}}(\mathbf{p} + \mathbf{p}') P(p) \left[ B(|\mathbf{q} + \mathbf{r}|, q', r') + B(|\mathbf{q} + \mathbf{q}'|, r, r') + B(|\mathbf{q} + \mathbf{r}'|, r, q') \right. \\
&+ \left. B(|\mathbf{r} + \mathbf{q}'|, q, r') + B(|\mathbf{r} + \mathbf{r}'|, q, q') + B(|\mathbf{q}' + \mathbf{r}'|, q, r) \right] + 8 \text{ cyc.} \\
&+ \int \mathcal{D} \frac{1}{\tilde{n}^3} B(|\mathbf{p} + \mathbf{q}|, |\mathbf{r} + \mathbf{p}'|, |\mathbf{q}' + \mathbf{r}'|) + 14 \text{ cyc.} \\
&+ \int \mathcal{D} \frac{1}{\tilde{n}^3} B(|\mathbf{p} + \mathbf{q} + \mathbf{r}|, |\mathbf{p}' + \mathbf{q}'|, r') + 59 \text{ cyc.} \\
&+ \int \mathcal{D} \frac{1}{\tilde{n}^3} B(|\mathbf{p} + \mathbf{q} + \mathbf{r} + \mathbf{p}'|, q', r') + 14 \text{ cyc.} \tag{C28}
\end{aligned}$$

#### Appendix D: The (in)dependence of the signal-to-noise ratio on the binning

In this section, we discuss the possible dependence of the signal-to-noise ratio, S/N, on the binning width  $\Delta k$ . We shall first consider the case of power spectrum and then move to bispectrum. Of course small inaccuracies arise when a coarse binning is used as the field varies across the bin. This is not the case we discuss here, instead we investigate whether the S/N explicitly depends on the binning to the lowest order.

For power spectrum, the Gaussian covariance scales

with the binning as  $(\Delta k)^{-1}$ , while the trispectrum contribution does not depend on  $\Delta k$ . The latter case is true for both the matter power spectrum case Eq. 6 and the Poisson model result Eq. C19. Suppose we change the binning in  $\delta$  from  $\Delta k$  to  $g\Delta k$ . For illustration purposes, let us take  $g = 2$ . When the binning is coarse grained by a factor of  $g = 2$ , the coarse-grained data vector  $S'$  is related to the original data vector  $S$  as

$$S'_i = \frac{S_{2i-1} + S_{2i}}{2}. \tag{D1}$$

i.e. an average over the neighbouring bins. For the sake



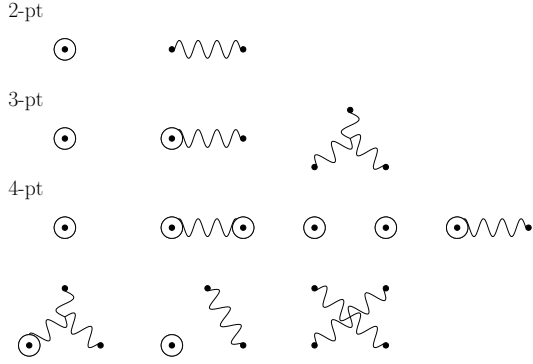


FIG. 23: Diagrammatic representation of the 2-point, 3-point and 4-point correlation including the shot noise contributions.

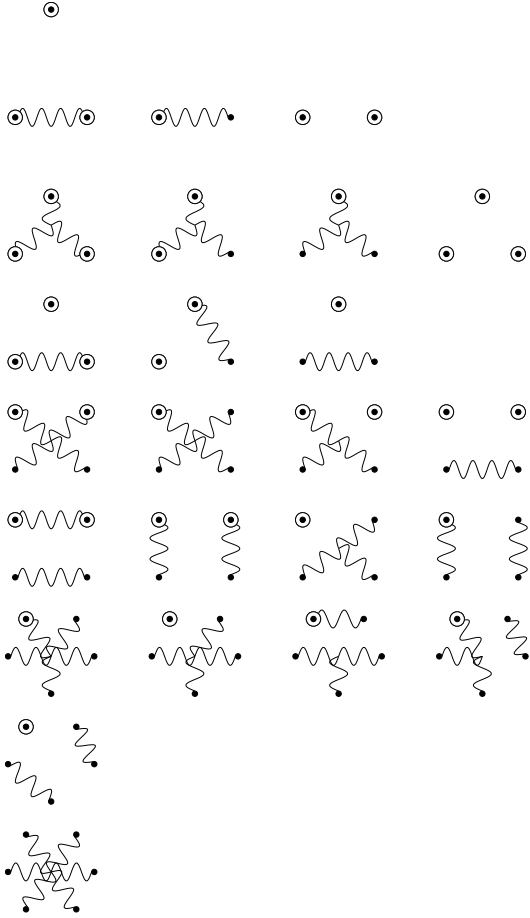


FIG. 24: Diagrammatic representation of the 6-point correlation including the shot noise contributions.

of simplicity here we use a simple average instead of the phase-volume weighted one, but the discussion is still valid when a weighted mean is used. Correspondingly, the covariance matrix of the coarse-grained data vector,  $C'$ , is given by

$$C'_{ij} = \langle S'_i S'_j \rangle$$

$$= \frac{1}{4} (\langle S_{2i-1} S_{2j-1} \rangle + \langle S_{2i-1} S_{2j} \rangle + \langle S_{2i} S_{2j-1} \rangle + \langle S_{2i} S_{2j} \rangle). \quad (\text{D2})$$

Therefore the coarse-grained covariance matrix is obtained by locally averaging the square block in the original matrix. In fact, the scalings of the Gaussian and non-Gaussian parts with respect to  $\Delta k$  respect the averaging prescription Eq. D2. Clearly the non-Gaussian matrix elements are invariant with respect to binning up to the accuracy of the field represented by the binned value. When the Gaussian covariance matrix is coarse grained, the diagonal element of the coarse-grained one is obtained by averaging the  $g$  diagonal elements and  $g(g-1)$  off-diagonal ones, which are zeroes. This is equivalent to simply averaging over the diagonal elements, and we get the  $1/g$  scaling for the diagonal element.

We can expand the inverse of the covariance matrix perturbatively as

$$\begin{aligned} C^{-1} &= (C_G + C_{\text{NG}})^{-1} \\ &= (I + C_G^{-1} C_{\text{NG}})^{-1} C_G^{-1} \\ &= C_G^{-1} - C_G^{-1} C_{\text{NG}} C_G^{-1} + (C_G^{-1} C_{\text{NG}})^2 C_G^{-1} + \dots \end{aligned} \quad (\text{D3})$$

This series expansion is valid when the non-Gaussian part is smaller than the Gaussian one in some appropriate sense.

Plugging Eq. D3 into the signal-to-noise ratio  $S^T C^{-1} S$ , we can check the binning dependence term by term. Clearly the first term  $S^T C_G^{-1} S$  is invariant because although the number of bins is reduced by a factor of  $g$ , the Gaussian precision matrix is enhanced by  $g$ . Because the enhancement by a factor of  $g$  in  $C_G^{-1}$  compensates the reduction in the number of rows in  $C_{\text{NG}}$ , the end result of  $S^T C_G^{-1} C_{\text{NG}} C_G^{-1} S$  is also invariant. By inspecting the perturbation series term by term, we conclude that the signal-to-noise for the power spectrum does not depend on the binning to the lowest order. This result was also stated in Ref. [13].

We now consider the case of bispectrum. When the binning is rescaled by a factor of  $g$ , say  $g = 2$ , the triangles in the bins  $[2i-1, 2i][2j-1, 2j][2k-1, 2k]$  are mapped into the triangle  $[2i]'[2j]'[2k]'$  in the coarse-grained case. Here the triangle sides are in unit of the fundamental mode  $k_F$ . For example, when the binning is changed from  $\Delta k = k_F$  to  $2k_F$ , the scalene triangles,  $[7][5][3]$ ,  $[7][5][4]$ ,  $[7][6][3]$ ,  $[7][6][4]$ ,  $[8][5][3]$ ,  $[8][5][4]$ ,  $[8][6][3]$ , and  $[8][6][4]$  are mapped to  $[8]'[6]'[4]'$ . For triangles with some symmetries, i.e. the isosceles and equilateral triangles, the counting is slightly different. For example, for the equilateral triangle set  $[7,8][7,8][7,8]$ , there are four distinct triangle sets  $[7][7][7]$ ,  $[7][7][8]$ ,  $[7][8][8]$ , and  $[8][8][8]$  and they are mapped to  $[8]'[8]'[8]'$ . We should not include triangles such as  $[8][7][7]$  as it is identical to  $[7][7][8]$  and its inclusion would cause the covariance matrix to be singular. The symmetry factor is important for the consistency of counting. We comment that using a coarse binning

for bispectrum is not very accurate for the bins with the smallest sides as triangles of many different shapes are mapped to a certain coarse-grained one. While for triangles of larger lengths, triangles are mapped to triangle of similar shape by the coarse-graining transformation, and hence we expect that the coarse-grained field reflects the original one accurately.

The Gaussian bispectrum covariance scales with the binning as  $(\Delta k)^{-3}$ . The leading disconnected non-Gaussian contribution to the dark matter bispectrum covariance, and also the non-Gaussian terms  $C_{\text{NG}}^{(1)}$  (Eq. C26), and part of  $C_{\text{NG}}^{(2)}$  (Eq. C28) scale as  $(\Delta k)^{-1}$  and couple only triangles with at least one side equal to each other. On the other hand  $C_{\text{NG}}^{(0)}$  (Eq. C27) does not depend on  $\Delta k$  and it couples all triangles. Again using Eq. D3, we can check if the S/N changes when the binning is rescaled by a factor of  $g$ , such as  $g = 2$ . For scalene triangles, the number of bins is reduced by a factor of 8, while  $C_{\text{G}}^{-1}$  is enhanced by a factor of 8, and hence it is

invariant with respect to binning. For other shapes, such as the equilateral triangles, taking the symmetry factors  $s_{123}$  into account, we can also show that  $S^T C_{\text{G}}^{-1} S$  is invariant. As the non-Gaussian term  $C_{\text{NG}}^{(0)}$  does not scale with  $\Delta k$ , by reasoning similar to the case of power spectrum, we also deduce that this term is invariant under bin width rescaling. For the terms that scale as  $(\Delta k)^{-1}$ , the coupling is non-trivial in the covariance matrix, it is hard to make analytical argument. By considering some explicit examples for scalene, isosceles and equilateral triangles, we check that this particular scaling and coupling also result in the S/N invariant with respect to the binning width.

Thus we have verified that the bispectrum S/N is invariant with respect to  $\Delta k$  to the leading order. However, we can only check the kinds of scaling in the non-Gaussian terms that we have computed, we cannot exclude that there are other kinds of scalings in terms that we did not consider here.

- 
- [1] W. J. Percival et al. (2dFGRS), *Mon. Not. Roy. Astron. Soc.* **327**, 1297 (2001), astro-ph/0105252.
- [2] M. Tegmark et al. (SDSS), *Phys. Rev.* **D69**, 103501 (2004), astro-ph/0310723.
- [3] D. J. Eisenstein et al. (SDSS), *Astrophys. J.* **633**, 560 (2005), astro-ph/0501171.
- [4] S. Cole et al. (2dFGRS), *Mon. Not. Roy. Astron. Soc.* **362**, 505 (2005), astro-ph/0501174.
- [5] M. Tegmark et al. (SDSS), *Phys. Rev.* **D74**, 123507 (2006), astro-ph/0608632.
- [6] L. Anderson et al. (BOSS), *Mon. Not. Roy. Astron. Soc.* **441**, 24 (2014), 1312.4877.
- [7] R. Laureijs, J. Amiaux, S. Arduini, J. Auguères, J. Brinchmann, R. Cole, M. Cropper, C. Dabin, L. Duvet, A. Ealet, et al., ArXiv e-prints (2011), 1110.3193.
- [8] LSST Science Collaboration, P. A. Abell, J. Allison, S. F. Anderson, J. R. Andrew, J. R. P. Angel, L. Armus, D. Arnett, S. J. Asztalos, T. S. Axelrod, et al., ArXiv e-prints (2009), 0912.0201.
- [9] P. Monaco (2016), 1605.07752.
- [10] A. Meiksin and M. J. White, *Mon. Not. Roy. Astron. Soc.* **308**, 1179 (1999), astro-ph/9812129.
- [11] R. Scoccimarro, M. Zaldarriaga, and L. Hui, *Astrophys. J.* **527**, 1 (1999), astro-ph/9901099.
- [12] R. E. Smith, *Mon. Not. Roy. Astron. Soc.* **400**, 851 (2009), 0810.1960.
- [13] R. Takahashi, N. Yoshida, M. Takada, T. Matsubara, N. Sugiyama, I. Kayo, T. Nishimichi, S. Saito, and A. Taruya, *Astrophys. J.* **726**, 7 (2011), 0912.1381.
- [14] J. Harnois-Déraps and U.-L. Pen, *MNRAS* **423**, 2288 (2012), 1109.5746.
- [15] L. Blot, P. S. Corasaniti, J.-M. Alimi, V. Reverdy, and Y. Rasera, *Mon. Not. Roy. Astron. Soc.* **446**, 1756 (2015), 1406.2713.
- [16] R. Scoccimarro and R. K. Sheth, *Mon. Not. Roy. Astron. Soc.* **329**, 629 (2002), astro-ph/0106120.
- [17] E. Sefusatti, M. Crocce, S. Pueblas, and R. Scoccimarro, *Phys. Rev.* **D74**, 023522 (2006), astro-ph/0604505.
- [18] R. Scoccimarro, E. Sefusatti, and M. Zaldarriaga, *Phys. Rev.* **D69**, 103513 (2004), astro-ph/0312286.
- [19] D. Bertolini, K. Schutz, M. P. Solon, J. R. Walsh, and K. M. Zurek, *Phys. Rev.* **D93**, 123505 (2016), 1512.07630.
- [20] I. Mohammed, U. Seljak, and Z. Vlah, Submitted to: *Mon. Not. Roy. Astron. Soc.* (2016), 1607.00043.
- [21] A. Cooray and W. Hu, *Astrophys. J.* **554**, 56 (2001), astro-ph/0012087.
- [22] I. Mohammed and U. Seljak, *Mon. Not. Roy. Astron. Soc.* **445**, 3382 (2014), 1407.0060.
- [23] S. Matarrese, L. Verde, and A. F. Heavens, *Mon. Not. Roy. Astron. Soc.* **290**, 651 (1997), astro-ph/9706059.
- [24] C. D. Rimes and A. J. S. Hamilton, *Mon. Not. Roy. Astron. Soc.* **371**, 1205 (2006), astro-ph/0511418.
- [25] A. J. S. Hamilton, C. D. Rimes, and R. Scoccimarro, *Mon. Not. Roy. Astron. Soc.* **371**, 1188 (2006), astro-ph/0511416.
- [26] R. de Putter, C. Wagner, O. Mena, L. Verde, and W. J. Percival, *JCAP* **4**, 019 (2012), 1111.6596.
- [27] M. Takada and W. Hu, *Phys. Rev.* **D87**, 123504 (2013), 1302.6994.
- [28] Y. Li, W. Hu, and M. Takada, *Phys. Rev.* **D89**, 083519 (2014), 1401.0385.
- [29] E. Sefusatti and R. Scoccimarro, *Phys. Rev.* **D71**, 063001 (2005), astro-ph/0412626.
- [30] R. Angulo, C. M. Baugh, C. S. Frenk, and C. G. Lacey, *Mon. Not. Roy. Astron. Soc.* **383**, 755 (2008), astro-ph/0702543.
- [31] H. A. Feldman, N. Kaiser, and J. A. Peacock, *ApJ* **426**, 23 (1994), astro-ph/9304022.
- [32] J. N. Fry, *ApJ* **279**, 499 (1984).
- [33] M. H. Goroff, B. Grinstein, S.-J. Rey, and M. B. Wise, *ApJ* **311**, 6 (1986).
- [34] F. Bernardeau, S. Colombi, E. Gaztañaga, and R. Scoccimarro, *Phys. Rep.* **367**, 1 (2002), arXiv:astro-ph/0112551.
- [35] Y. Rasera, P.-S. Corasaniti, J.-M. Alimi, V. Bouillot,

- V. Reverdy, and I. Balms, *Mon. Not. Roy. Astron. Soc.* **440**, 1420 (2014), 1311.5662.
- [36] D. N. Spergel, R. Bean, O. Doré, M. R. Nolta, C. L. Bennett, J. Dunkley, G. Hinshaw, N. Jarosik, E. Komatsu, L. Page, et al., *ApJs* **170**, 377 (2007), astro-ph/0603449.
- [37] A. Lewis, A. Challinor, and A. Lasenby, *ApJ* **538**, 473 (2000), arXiv:astro-ph/9911177.
- [38] R. Teyssier, *A&A* **385**, 337 (2002), astro-ph/0111367.
- [39] W. Hu and A. V. Kravtsov, *Astrophys. J.* **584**, 702 (2003), astro-ph/0203169.
- [40] R. Scoccimarro, S. Colombi, J. N. Fry, J. A. Frieman, E. Hivon, and A. Melott, *ApJ* **496**, 586 (1998), arXiv:astro-ph/9704075.
- [41] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes 3rd Edition: The Art of Scientific Computing* (Cambridge University Press, New York, NY, USA, 2007), 3rd ed., ISBN 0521880688, 9780521880688.
- [42] Y. Xu, X. Wang, and X. Chen, *Astrophys. J.* **798**, 40 (2015), 1410.7794.
- [43] Y. Welling, D. van der Woude, and E. Pajer (2016), 1605.06426.
- [44] Y. Xu, J. Hamann, and X. Chen, *ArXiv e-prints* (2016), 1607.00817.
- [45] M. Tegmark, A. Taylor, and A. Heavens, *Astrophys. J.* **480**, 22 (1997), astro-ph/9603021.
- [46] J. Hartlap, P. Simon, and P. Schneider, *Astron. Astrophys.* **464**, 399 (2007), astro-ph/0608064.
- [47] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis, 3rd Edition* (Wiley-Interscience, New York, 2003).
- [48] L. Blot, P. S. Corasaniti, L. Amendola, and T. D. Kitching, *Mon. Not. Roy. Astron. Soc.* **458**, 4462 (2016), 1512.05383.
- [49] T. Jenkyns and E. Muller, *The American Mathematical Monthly* **107**, 634 (2000).
- [50] C. D. Rimes and A. J. S. Hamilton, *Mon. Not. Roy. Astron. Soc.* **360**, L82 (2005), astro-ph/0502081.
- [51] M. C. Neyrinck, I. Szapudi, and C. D. Rimes, *Mon. Not. Roy. Astron. Soc.* **370**, L66 (2006), astro-ph/0604282.
- [52] J. Lee and U.-L. Pen, *Astrophys. J.* **686**, L1 (2008), 0807.1538.
- [53] P. Coles and B. Jones, *MNRAS* **248**, 1 (1991).
- [54] M. C. Neyrinck, I. Szapudi, and A. S. Szalay, *ApJL* **698**, L90 (2009), 0903.4693.
- [55] J. Carron and M. C. Neyrinck, *ApJ* **750**, 28 (2012), 1201.1444.
- [56] A. Repp, I. Szapudi, J. Carron, and M. Wolk, *MNRAS* **454**, 3533 (2015), 1506.00083.
- [57] M. Sato and T. Nishimichi, *Phys. Rev.* **D87**, 123538 (2013), 1301.3588.
- [58] H. A. Feldman, J. A. Frieman, J. N. Fry, and R. Scoccimarro, *Phys. Rev. Lett.* **86**, 1434 (2001), astro-ph/0010205.
- [59] R. Scoccimarro, H. A. Feldman, J. N. Fry, and J. A. Frieman, *Astrophys. J.* **546**, 652 (2001), astro-ph/0004087.
- [60] L. Verde et al., *Mon. Not. Roy. Astron. Soc.* **335**, 432 (2002), astro-ph/0112161.
- [61] H. Gil-Marn, J. Norea, L. Verde, W. J. Percival, C. Wagner, M. Manera, and D. P. Schneider, *Mon. Not. Roy. Astron. Soc.* **451**, 539 (2015), 1407.5668.
- [62] Y. P. Jing and G. Boerner, *Astrophys. J.* **607**, 140 (2004), astro-ph/0311585.
- [63] E. Gaztañaga and R. Scoccimarro, *MNRAS* **361**, 824 (2005), astro-ph/0501637.
- [64] Y. Wang, X.-H. Yang, H. J. Mo, F. C. van den Bosch, and Y.-Q. Chu, *Mon. Not. Roy. Astron. Soc.* **353**, 287 (2004), astro-ph/0404143.
- [65] F. Marín, *ApJ* **737**, 97 (2011), 1011.4530.
- [66] F. A. Marín, C. Blake, G. B. Poole, C. K. McBride, S. Brough, M. Colless, C. Contreras, W. Couch, et al., *MNRAS* **432**, 2654 (2013), 1303.6644.
- [67] Z. Slepian et al. (2015), 1512.02231.
- [68] H. Guo et al. (2016), 1608.03660.
- [69] J. N. Fry and E. Gaztañaga, *ApJ* **413**, 447 (1993), arXiv:astro-ph/9302009.
- [70] K. C. Chan, R. Scoccimarro, and R. K. Sheth, *Phys. Rev. D* **85**, 083509 (2012), arXiv:1201.3614.
- [71] T. Baldauf, U. Seljak, V. Desjacques, and P. McDonald, *Phys. Rev. D* **86**, 083540 (2012), arXiv:1201.4827.
- [72] N. Hamaus, B. D. Wandelt, P. M. Sutter, G. Lavaux, and M. S. Warren, *Physical Review Letters* **112**, 041304 (2014), 1307.2571.
- [73] K. C. Chan, N. Hamaus, and V. Desjacques, *Phys. Rev. D* **90**, 103521 (2014), 1409.3849.
- [74] J. Clampitt, B. Jain, and C. Sánchez, *MNRAS* **456**, 4425 (2016), 1507.08031.
- [75] F.-S. Kitaura, C.-H. Chuang, Y. Liang, C. Zhao, C. Tao, S. Rodríguez-Torres, D. J. Eisenstein, H. Gil-Marín, J.-P. Kneib, C. McBride, et al., *Physical Review Letters* **116**, 171301 (2016), 1511.04405.
- [76] N. Hamaus, A. Pisani, P. M. Sutter, G. Lavaux, S. Escoffier, B. D. Wandelt, and J. Weller, *Physical Review Letters* **117**, 091302 (2016), 1602.01784.
- [77] L.-Y. Chiang, P. Coles, and P. Naselsky, *MNRAS* **337**, 488 (2002), astro-ph/0207584.
- [78] R. Wolstenhulme, C. Bonvin, and D. Obreschkow, *ApJ* **804**, 132 (2015), 1409.3007.
- [79] A. Eggemeier, T. Battefeld, R. E. Smith, and J. Niemeyer, *MNRAS* **453**, 797 (2015), 1504.04036.
- [80] R. Mehrem (2011), arXiv:0909.0494.
- [81] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables* (Dover Publications, New York, 1964).
- [82] K. C. Chan and R. Scoccimarro, *Phys. Rev. D* **86**, 103519 (2012), arXiv:1204.5770.
- [83] P. J. E. Peebles, *The Large-Scale Structure of the Universe* (Princeton University Press, New Jersey, 1980).