

Comparing approximate methods for mock catalogues and covariance matrices III: Bispectrum

Manuel Colavincenzo^{1,2,3*}, Emiliano Sefusatti^{4,5}, Pierluigi Monaco^{3,4,5},
Linda Blot^{6,7}, Martin Crocce^{6,7}, Martha Lippich^{8,9}, Ariel G. Sánchez⁸,
Marcelo A. Alvarez¹⁰, Aniket Agrawal¹¹, Santiago Avila¹²,
Andrés Balaguera-Antolínez^{13,14}, Richard Bond¹⁵, Sandrine Codis^{15,16},
Claudio Dalla Vecchia^{13,14}, Antonio Dorta^{13,14}, Pablo Fosalba^{6,7},
Albert Izard^{17,18,6,7}, Francisco-Shu Kitaura^{13,14}, Marcos Pellejero-Ibanez^{13,14},
George Stein¹⁵, Mohammadjavad Vakili¹⁹, Gustavo Yepes^{20,21}

¹ *Dipartimento di Fisica, Università di Torino, Via P. Giuria 1, 10125 Torino, Italy*

² *Istituto Nazionale di Fisica Nucleare, Sezione di Torino, Via P. Giuria 1, 10125 Torino, Italy*

³ *Dipartimento di Fisica, Sezione di Astronomia, Università di Trieste, via Tiepolo 11, 34143 Trieste, Italy*

⁴ *Istituto Nazionale di Astrofisica, Osservatorio Astronomico di Trieste, via Tiepolo 11, 34143 Trieste, Italy*

⁵ *Istituto Nazionale di Fisica Nucleare, Sezione di Trieste, Via Valerio, 2, 34127 Trieste, Italy*

⁶ *Institute of Space Sciences (ICE, CSIC), Campus UAB, Carrer de Can Magrans, s/n, 08193 Barcelona, Spain*

⁷ *Institut d'Estudis Espacials de Catalunya (IEEC), 08034 Barcelona, Spain*

⁸ *Max-Planck-Institut für extraterrestrische Physik, Postfach 1312, Giessenbachstr., 85741 Garching, Germany*

⁹ *Universitäts-Sternwarte München, Ludwig-Maximilians-Universität München, Scheinerstrasse 1, 81679 Munich, Germany*

¹⁰ *Berkeley Center for Cosmological Physics, Campbell Hall 341, University of California, Berkeley CA 94720, USA*

¹¹ *Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Str. 1, 85741 Garching, Germany*

¹² *Institute of Cosmology & Gravitation, Dennis Sciama Building, University of Portsmouth, Portsmouth PO1 3FX, UK*

¹³ *Instituto de Astrofísica de Canarias, C/Vía Láctea, s/n, 38200, La Laguna, Tenerife, Spain*

¹⁴ *Departamento Astrofísica, Universidad de La Laguna, 38206 La Laguna, Tenerife, Spain*

¹⁵ *Canadian Institute for Theoretical Astrophysics, University of Toronto, 60 St. George Street, Toronto, ON M5S 3H8, Canada*

¹⁶ *Institut d'Astrophysique de Paris, CNRS & Sorbonne Université, UMR 7095, 98 bis boulevard Arago, 75014 Paris, France*

¹⁷ *Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109, USA*

¹⁸ *Department of Physics and Astronomy, University of California, Riverside, CA 92521, USA*

¹⁹ *Leiden Observatory, Leiden University, P.O. Box 9513, NL-2300 RA, Leiden, The Netherlands*

²⁰ *Departamento de Física Teórica, Módulo 15, Universidad Autónoma de Madrid, 28049 Madrid, Spain*

²¹ *Centro de Investigación Avanzada en Física Fundamental (CIAFF), Universidad Autónoma de Madrid, 28049 Madrid, Spain*

*E-mail: mcolavin@to.infn.it

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

We compare the measurements of the bispectrum and the estimate of its covariance obtained from a set of different methods for the efficient generation of approximate dark matter halo catalogs to the same quantities obtained from full N-body simulations. To this purpose we employ a large set of three-hundred realisations of the same cosmology for each method, run with matching initial conditions in order to reduce the contribution of cosmic variance to the comparison. In addition, we compare how the error on cosmological parameters such as linear and nonlinear bias parameters depends on the approximate method used for the determination of the bispectrum variance. As general result, most methods provide errors within 10% of the errors estimated from N-body simulations. Exceptions are those methods requiring calibration of the clustering amplitude but restrict this to two-point statistics. Finally we test how our results are affected by being limited to a few hundreds measurements from N-body simulation, and therefore to the bispectrum variance, by comparing with a larger set of several thousands realisations performed with one approximate method.

Key words: cosmological simulations – galaxies clustering – error estimation – large-scale structure of Universe.

1 INTRODUCTION

This is the last of a series of three papers exploring the problem of covariance estimation for large-scale structure observables based on dark matter halo catalogs obtained from approximate methods. The importance of a large-set of galaxy catalogs both for purposes of covariance estimation as well as for the testing of the analysis pipeline has become evident over the last decade when such tools have been routinely employed in the exploitation of several major galaxy surveys (see, *e.g.* Manera et al. 2013; de la Torre et al. 2013; Koda et al. 2016; Kitaura et al. 2016; Avila et al. 2017).

In this context, it is crucial to ensure that mock catalogs correctly reproduce the statistical properties of the galaxy distribution. Such properties are characterised not only by the two-point correlation function, but are quantified as well in terms of higher-order correlators like the 3-point and 4-point correlation functions, since the large-scale distributions of both matter and galaxies are highly non-Gaussian random fields.

A correct non-Gaussian component in mock galaxy catalogs has essentially two important implications. In the first place, we expect the trispectrum, *i.e.* the 4-point correlation function in Fourier space, to contribute non-negligibly to the covariance of two-point statistics. This is perhaps more evident in the case of the power spectrum, already in terms of the direct correlation between band power that we measure even in the ideal case of periodic box simulations (see, *e.g.* Meiksin & White 1999; Scoccimarro et al. 1999b; Takahashi et al. 2009; Ngan et al. 2012; Blot et al. 2015). In addition, finite-volume effects such as beat-coupling/super-sample covariance (Hamilton et al. 2006; Rimes & Hamilton 2006; Seifusatti et al. 2006; Takada & Hu 2013) and local average of the density field (de Putter et al. 2012) can be described as consequences of the interplay between the survey window function and both the galaxy bispectrum and trispectrum. In the second place higher-order correlation functions, and particularly the galaxy 3PCF and the bispectrum are emerging as relevant observables in their own right, capable of complementing the more standard analysis of 2PCF and power spectrum (Gaztañaga et al. 2009; Slepian et al. 2017; Gil-Marín et al. 2015a,b, 2017; Pearson & Samushia 2017).

Both these aspects provide strong motivations for ensuring that not only higher-order correlations are properly reproduced in mock catalogs but also their own covariance properties are recovered with sufficient accuracy. In this work we focus, in particular, on the bispectrum of the halo distribution. This is the lowest order non-Gaussian statistic characterising the three-dimensional nature of the large-scale structure. It has also the practical advantage of requiring relatively small numerical resources for its estimation on large sets of catalogs, at least with respect to the 3-point correlation function in real space. On the other hand, a correct prediction of the halo bispectrum does *not* ensure that higher-order correlators such as the halo trispectrum are similarly accurately reproduced. For instance, a matter distribution realised at second order in Lagrangian Perturbation Theory (LPT, the basis for several approximate methods) is characterised by a bispectrum fully reproducing the expected prediction at tree-level in Eulerian PT valid at large scales but that is not the case for the matter trispec-

trum since the scheme only partially reproduces the third order Eulerian nonlinear correction (Scoccimarro 1998).

With this caveat in mind, in this paper we focus on the direct comparison of the halo bispectrum and its covariance, along with a comparison of the errors on the recovered halo bias parameters from a simple likelihood analysis adopting different estimates of the bispectrum variance. Clearly our sets of 300 halo catalogs from N-body simulations and the various approximate methods do not allow a proper comparison at the covariance level, since a reliable estimate of the covariance matrix requires thousands of such realisations. Nevertheless we explore the implications of such limitation taking advantage of a much larger set of 10,000 runs, used for the first time in (Colavincenzo et al. 2017), of one of the approximate methods.

Two companion papers focus on similar comparisons for the 2-point correlation function (Lippich et al. 2018) and for the power spectrum (Blot et al. 2018): we will refer to them, respectively, as Paper I and Paper II throughout this work.

This paper is organised as follows. In section 2 we present the approximate methods considered in this work and how they address the proper prediction of the non-Gaussian properties of the halo distribution. In section 3 we describe the measurements of the halo bispectrum and its covariance for each set of catalogs which are then compared in section 4. In section 5 we extend the comparison to the errors on cosmological parameters while in section 6 we present a few tests to quantify possible systematics due to the limited number of catalogs at our disposal. Finally, we present our conclusions in section 7.

2 THE CATALOGS

For a detailed description of the different approximate methods compared in this, as well as the two companion papers, we refer the reader to section 3 of Paper I, while for a more general examination of the state-of-the-art in the field we refer to the review in Monaco (2016). For a quick reference we reproduce in Table 2 the Table 1 of Paper II, providing a brief summary of the codes considered. Here we briefly discuss the main characteristics of the catalogs and the implications for accurate bispectrum predictions.

For all runs we consider a box of size $L = 1500 h^{-1} \text{Mpc}$ and a cosmology defined by the best-fitting parameters of the analysis in Sánchez et al. (2013). The N-body runs employ a number of particles of 1000^3 leading to a particle mass $m_p = 2.67 \times 10^{11} h^{-1} M_\odot$. In addition to the 100 runs mentioned in Grieb et al. (2016), for this work we consider additional simulations for a total of 300 runs.

We work on the halo catalogs obtained from the N-body identified with a standard Friends-of-friends (FoF) algorithm. FoF halos were then subject to the unbinding procedure provided by the Subfind algorithm (Springel et al. 2001) from snapshots at $z = 1$. We consider two samples characterised by a minimal mass of $M_{min} = 42 m_p = 1.12 \times 10^{13} h^{-1} M_\odot$ (Sample 1) and $M_{min} = 100 m_p = 2.67 \times 10^{13} h^{-1} M_\odot$ (Sample 2). The corresponding number densities are of 2.13×10^{-4} and 5.44×10^{-5} , respectively. For Sample 2 the power spectrum signal is dominated by shot-noise for scales $k \gtrsim 0.15 h \text{Mpc}^{-1}$, while for Sample 1

Method	Algorithm	Computational Requirements	Reference
Minerva	N-body Gadget-2 Halos : SubFind	CPU Time: 4500 hours Memory allocation: 660 Gb	Grieb et al. (2016) https://wwwmpa.mpa-garching.mpg.de/gadget/
ICE-COLA	Predictive 2LPT + PM solver Halos : FoF(0.2)	CPU Time: 33 hours Memory allocation: 340 Gb	Izard et al. (2016) Modified version of: https://github.com/junkoda/cola_halo
PINOCCHIO	Predictive 3LPT + ellipsoidal collapse Halos : ellipsoidal collapse	CPU Time: 6.4 hours Memory allocation: 265 Gb	Monaco et al. (2013); Munari et al. (2017) https://github.com/pigimonaco/Pinocchio
PEAKPATCH	Predictive 2LPT + ellipsoidal collapse Halos : Spherical patches over initial overdensities	CPU Time: 1.72 hours* Memory allocation: 75 Gb*	Bond & Myers (1996a,b,c) Not public
HALOGEN	Calibrated 2LPT + biasing scheme Halos : exponential bias	CPU Time: 0.6 hours Memory allocation: 44 Gb Input: \bar{n} , 2-pt correlation function halo masses and velocity field	Avila et al. (2015). https://github.com/savila/halogen
PATCHY	Calibrated ALPT + biasing scheme Halos : non-linear, stochastic and scale-dependent bias	CPU Time: 0.2 hours Memory allocation: 15 Gb Input: \bar{n} , halo masses and environment	Kitaura et al. (2014) Not Public
Lognormal	Calibrated Lognormal density field Halos : Poisson sampled points	CPU Time: 0.1 hours Memory allocation: 5.6 Gb Input: \bar{n} , 2-pt correlation function	Agrawal et al. (2017) https://bitbucket.org/komatsu5147/lognormal_galaxies
Gaussian	Theoretical Gaussian density field Halos : n/a	CPU Time: n/a Memory allocation: n/a Input: $P(k)$ and \bar{n}	Scoccimarro et al. (1998) for the bispectrum

Table 1. Name of the methods, type of algorithm, halo definition, computing requirements and references for the compared methods. All computing times are given in cpu-hours per run and memory requirements are per run, not including the generation of the initial conditions. The computational resources for halo finding in the N-body and ICE-COLA mocks are included in the requirements. The computing time refers to runs down to redshift 1 except for the N-body where we report the time down to redshift 0 (we estimate an overhead of $\sim 50\%$ between $z = 0$ and $z = 1$). Since every code was run in a different machine the computing times reported here are only indicative. We include the information needed for calibration/prediction of the covariance where relevant. Mocks marked with “*” require an higher resolution run in order to resolve the lower mass halos of our Sample 1 and therefore more computational resources than quoted here.

the shot-noise contribution is always below the signal but still not negligible.

We produced a set of 300 realisations with each of the approximate methods considered, imposing the same initial conditions as the N-body runs in order to reduce any difference due to cosmic variance. The definition of the two samples in the catalogs obtained by the approximate methods depends on the specific algorithm.

We can distinguish between three different classes of algorithms: *predictive methods* (ICE-COLA, PINOCCHIO and PEAKPATCH) that aim at identifying the Lagrangian patches that collapse into halos and do not need to be recalibrated against a simulation (in particular, ICE-COLA is a PM solver, so it is expected to be more accurate at a higher computational cost); *calibrated methods* (HALOGEN and PATCHY) that populate a large-scale density field with halos using a bias model, and need to be recalibrated to match a sample in number density and clustering amplitude; *analytical methods*. These last include the Gaussian prediction for the bispectrum covariance based on the measured

power spectrum, and the Lognormal method, predicting the halo distribution from some assumption on the density field PDF. In particular, the Lognormal realisations do not share the same initial conditions as the N-body runs. Therefore, we employ for this method, the covariance estimated from 1,000 realisation in order to beat down sample variance.

Notice that also for the predictive methods the minimal mass for each sample is set by requiring the same abundance as the N-body samples. A comparison that assumes directly the same mass thresholds as the N-body samples is discussed in appendix A. All other methods assume such density matching by default. For the PEAKPATCH comparison, only the larger mass sample is available.

All methods, with the exception of Lognormal, employ Lagrangian PT at second order (or higher) to determine the large-scale matter density field. We expect therefore, as mentioned before, that at least at large scales where the characteristic LPT suppression of power is still negligible, the measured halo bispectrum presents qualitatively the expected dependence on the shape of the triangular configu-

rations. Any difference with the full N-body results at large scales will likely arise from the specific way each method implements the relation between 2LPT-displaced matter particles and its definition of halos or particle groups. The case of Lognormal is different since it is based on a nonlinear transformation of the Gaussian matter density qualitatively reproducing the nonlinear density probability distribution function (Coles & Jones 1991), but with no guarantee to properly reproduce the proper dependence on configuration of higher-order correlation functions, starting from the matter bispectrum.

These considerations have been already illustrated by the results of the code-comparison project of Chuang et al. (2015b). This work comprises a comparison of both the 3-Point Correlation Function and the bispectrum of halos of minimal mass of $10^{13} h^{-1} M_{\odot}$, very similar to one of the two samples considered in our work, but evaluated at the lower redshift $z \simeq 0.55$. Each measurement was performed for a relatively small set of configurations, covering, in the bispectrum case, the range of scales $0.1 \leq (k/h \text{ Mpc}^{-1}) \leq 0.3$. The codes ICE-COLA, EZMOCK (Chuang et al. 2015a) and PATCHY (the last two requiring calibration of the halo power spectrum) reproduced the N-body results with an accuracy of 10-15%, PINOCCHIO at the 20-25% level, while HALOGEN and PTHALOS (Scoccimarro & Sheth 2002; Manera et al. 2013) at the 40-50%. All these methods correctly recovered the overall shape dependence. On the other hand, the Lognormal method failed to do so, despite the predicted bispectrum showed a comparable, overall magnitude (see also White et al. 2014). It should be noted that in some cases, as *e.g.* PINOCCHIO, the codes employed in this work correspond to an updated version w.r.t. those considered by Chuang et al. (2015b).

We notice that, in the present work, we will go beyond the results of Chuang et al. (2015b), extending the test of the approximate methods to the comparison of the recovered bispectrum *variance*.

3 MEASUREMENTS

For each sample we estimate the Fourier-space density on a grid of 256 of linear size employing the 4th-order interpolation algorithm and the interlacing technique implemented in the POWERI4¹ code described in Sefusatti et al. (2016).

The bispectrum estimator is given by

$$\hat{B}_{tot}(k_1, k_2, k_3) \equiv \frac{k_f^3}{V_B(k_1, k_2, k_3)} \int_{k_1} d^3 q_1 \int_{k_2} d^3 q_2 \int_{k_3} d^3 q_3 \times \delta_D(\mathbf{q}_{123}) \delta_{\mathbf{q}_1} \delta_{\mathbf{q}_2} \delta_{\mathbf{q}_3} \quad (1)$$

where the integrations are taken on shells of size Δk centered on k_i and where

$$V_B(k_1, k_2, k_3) \equiv \int_{k_1} d^3 q_1 \int_{k_2} d^3 q_2 \int_{k_3} d^3 q_3 \delta_D(\mathbf{q}_{123}) \simeq 8\pi^2 k_1 k_2 k_3 \Delta k^3 \quad (2)$$

is a normalisation factor counting the number of fundamental triangles (those defined by the vectors \mathbf{q}_1 , \mathbf{q}_2 and \mathbf{q}_3 on

the discrete Fourier density grid) in a given triangle bin (defined instead by the triplet k_1 , k_2 and k_3 plus the size Δk for all sides). Its implementation is based on the algorithm described in Scoccimarro (2015).

The measured bispectrum will be affected by shot-noise. Under the assumption of Poisson shot-noise, we correct the measurement \hat{B} as follows (Matarrese et al. 1997)

$$B(k_1, k_2, k_3) = \hat{B}_{tot} - \frac{1}{(2\pi)^3 \bar{n}} [P(k_1) + P(k_2) + P(k_3)] - \frac{1}{(2\pi)^6 \bar{n}^2}, \quad (3)$$

where \bar{n} is the halo density of each individual catalog and $P(k)$ is the halo power spectrum, in turn corrected for shot-noise.

We consider all triangular configurations defined by discrete wavenumbers multiples of $\Delta k = 3k_f$ with $k_f \equiv 2\pi/L$ being the fundamental frequency of the box, up to a maximum value of $0.38 h \text{ Mpc}^{-1}$, although we will limit our analysis to scales defined by $k_i \leq 0.2 h \text{ Mpc}^{-1}$, where we conservatively expect analytical predictions in perturbation theory to accurately describe the galaxy bispectrum. These choices lead to a total number of triangle bins of 508.

Given the estimator above, the Gaussian prediction for the variance is given by (Scoccimarro 2000),

$$\Delta B^2(k_1, k_2, k_3) \equiv \langle (\hat{B}^2 - \langle \hat{B} \rangle^2) \rangle \simeq s_B \frac{k_f^3}{V_B} P_{tot}(k_1) P_{tot}(k_2) P_{tot}(k_3), \quad (4)$$

with $s_B = 6, 2, 1$ for equilateral, isosceles and scalene triangles respectively and where $P_{tot}(k) = P(k) + 1/[(2\pi)^3 \bar{n}]$ includes the Poisson shot-noise contribution due to the halo density \bar{n} . We will compare our measurements to this theoretical prediction for the variance. For such comparison we will employ the measured mean value of $P_{h,tot}(k)$ and the exact number of fundamental triangles $V_B(k_1, k_2, k_3)$ as provided by the code, which is slightly different, for certain triangular shapes, from the approximate value on the second line of eq. (2).

Theoretical predictions are computed for “effective” values of the wavenumbers defined, for a given configuration of sides k_1 , k_2 and k_3 by

$$\tilde{k}_{1,23} \equiv \frac{1}{V_B} \int_{k_1} d^3 q_1 q_1 \int_{k_2} d^3 q_2 \int_{k_3} d^3 q_3 \delta_D(\mathbf{q}_{123}), \quad (5)$$

and similarly for the other two values. Differences with respect to evaluations at the center of each k -bin are marginally relevant and only so for the largest scales.

4 BISPECTRUM AND BISPECTRUM VARIANCE COMPARISON

In this section we compare the measurements of the halo bispectrum for the two halo samples both in real and redshift space. Since one of the aims of this work is testing how accurately the non-Gaussian properties of the large-scale halo distribution are recovered, it is relevant to look at the lowest order non-Gaussian statistic also in real space, while the bispectrum as a direct observable motivates all redshift-space tests.

We compare as well the variance estimated from the 300

¹ <https://github.com/sefusatti/PowerI4>

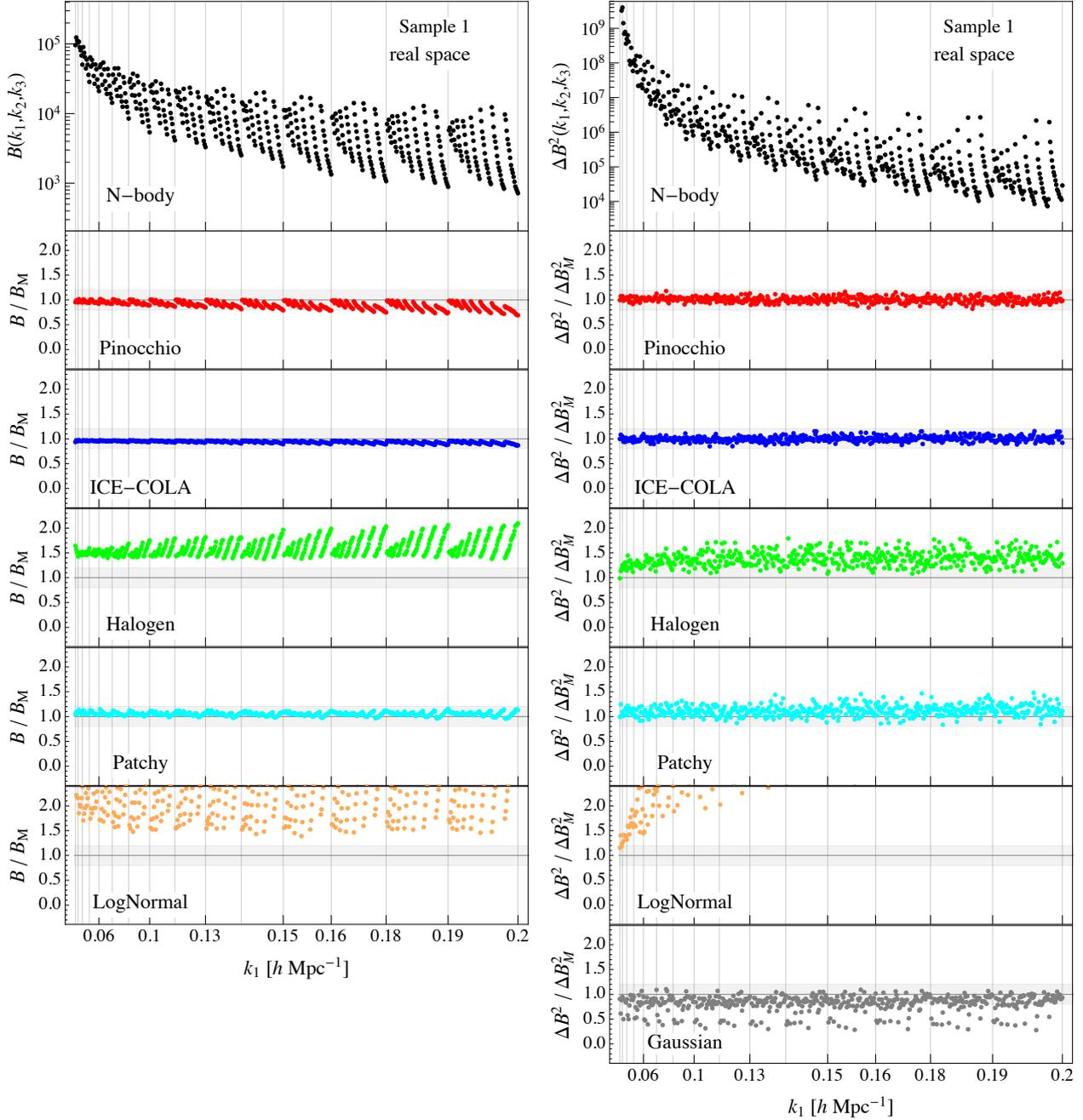


Figure 1. Average bispectrum (left column) and its variance (right column) for all triangle configurations obtained from the 300 realisations for the first mass sample in real space. The top panels show the results for the Minerva (black dots), while all other panels show the ratio between the estimate from an approximate method and the N-body one. In the last panel of the right column the grey dots show the ratio between the Gaussian prediction for the bispectrum variance, eq. (4), and the variance obtained from the N-body. The horizontal shaded area represents a 20% error. The vertical lines mark the triangle configurations where k_1 (the maximum of the triplet) is changing, so that all the points in between such lines correspond to all triangles with the same value for k_1 and all possible values of k_2 and k_3 . Since we assume $k_1 \geq k_2 \geq k_3$, the value of k_1 corresponds also to the maximum side of the triangle. Mocks for PEAKPATCH are not provided in the first sample so its bispectrum is missing in this case.

runs and the covariance among different triangles. Clearly, 300 realisations are not enough to provide a proper estimate of the covariance among 508 triplets. The comparison is then aiming at verifying that the same statistical fluctuations appear across the estimates from different approximate methods, taking advantage of the shared initial conditions.

4.1 Real space

Figures 1 and 2 show, respectively for Sample 1 and Sample 2, in the left column, top panel, the real-space halo bispectrum averaged over the 300 N-body simulations. The panels below show the ratio between the same measurements

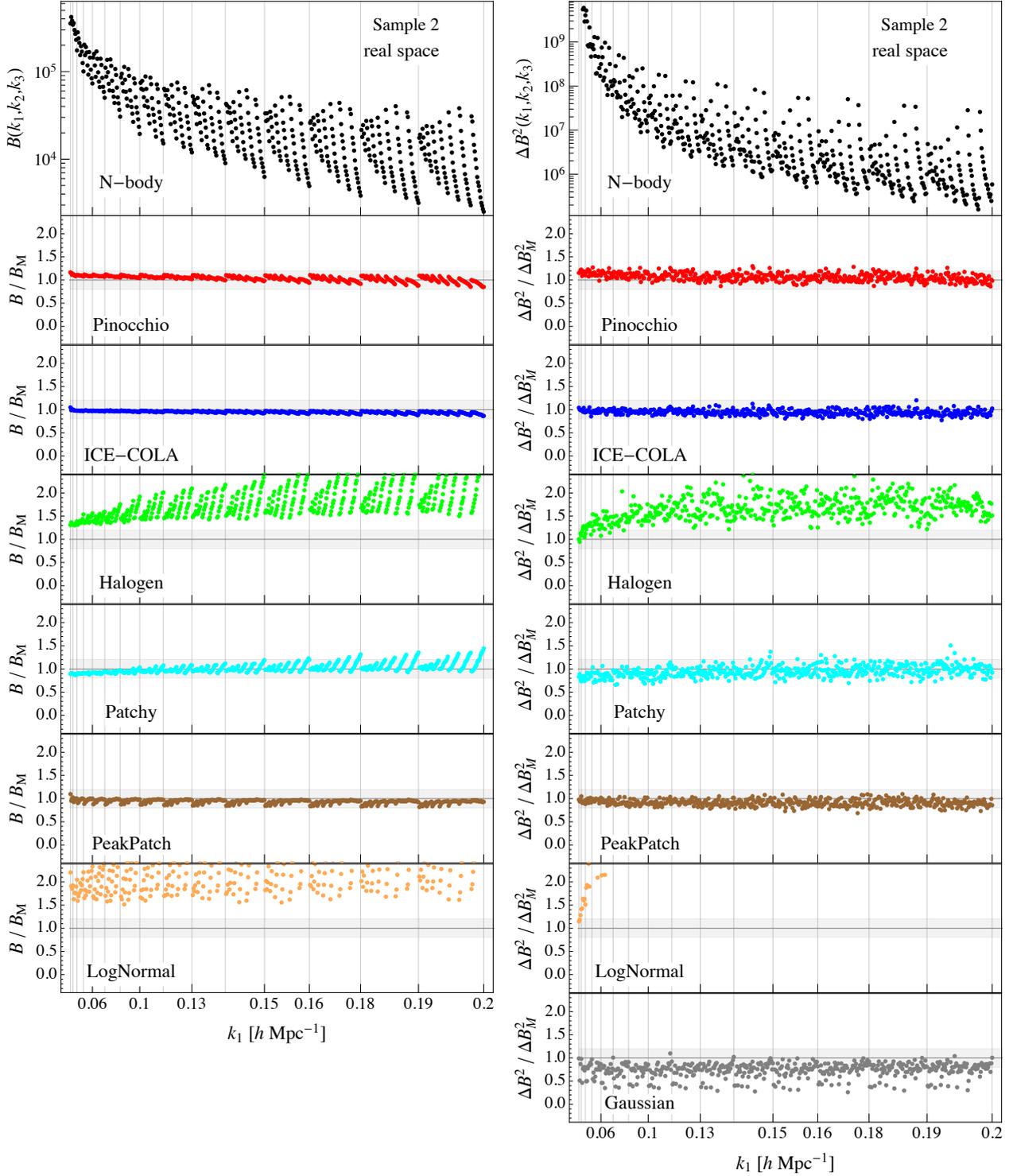


Figure 2. Same as figure 1, but for Sample 2.

obtained from all approximate methods and the N-body results. The right column shows a similar comparison for the halo bispectrum variance. For this quantity we include an additional, bottom panel where we plot the comparison between the Gaussian prediction for the bispectrum variance, eq. 4, and the N-body estimate. We will keep the color-coding for each methods consistently throughout this paper.

Each dot represents the bispectrum for a particular triplet $\{k_1, k_2, k_3\}$. These are plotted in an order where $k_1 \geq k_2 \geq k_3$ with increasing values of each k_i for all allowed configurations. In practice, the first configurations are, in units of the k -bin size Δk

$$\{1, 1, 1\}, \{2, 1, 1\}, \{2, 2, 1\}, \{2, 2, 2\}, \{3, 2, 1\}, \dots$$

The ticks on the abscissa mark the value of k_1 , the largest wavenumber in each triplet, and the vertical grey lines denote the configurations where k_1 changes.

All predictive methods, that is PINOCCHIO, ICE-COLA and PEAKPATCH (this last for Sample 2), reproduce the N-body measurements within 15% for most of the triangle configurations, with some small dependence on the triangle shape. Similar results, among the methods requiring some form of calibration, are obtained for PATCHY, with just some higher discrepancies at the 20-30% level appearing for Sample 2 at small scales, mainly for nearly equilateral triangles. The other calibrated methods fare worse. HALOGEN shows differences above 50%, reaching 100% for nearly equilateral configurations in both samples. The LogNormal approach, as one can expect, shows the largest discrepancy for all the scales and all the configurations in both samples.

Similar considerations can be made for the comparison of the variance. In this case a large component is provided by the shot-noise contribution, so the ratios to the N-body results show a less prominent dependence on the triangle shape. In general, we expect the agreement with N-body to depend to a large extent, particularly for Sample 2, on the correct matching of the object density, and more so for those LPT-based methods that show a lack of power in this regime. The Gaussian prediction underestimates the N-body result by 10-20% for the majority of configurations, and reaching up to 50% for squeezed triangles, *i.e.* those comprising the smallest wavenumber.

4.2 Redshift space

Figures 3 and 4, respectively for the Sample 1 and 2, show the redshift-space bispectrum monopole (left column) and its variance (right column), with the same conventions assumed for the real-space results in figure 1. The overall results are by and large very similar to the real-space ones. Only for the first sample, both HALOGEN and PATCHY show a better agreement with the N-body results than in real space. As before Lognormal is the one that shows the largest disagreement with the N-body results.

Figure 5 shows, for Sample 1, a representative subset of the off-diagonal elements of the bispectrum covariance matrix in redshift space as estimated by the different methods. The quantities shown are the cross-correlation coefficients r_{ij} defined as

$$r_{ij} \equiv \frac{C_{ij}}{\sqrt{C_{ii}C_{jj}}} \quad (6)$$

where

$$C_{ij} \equiv \langle (\hat{B}(t_i) - \langle \hat{B}(t_i) \rangle)(\hat{B}(t_j) - \langle \hat{B}(t_j) \rangle) \rangle, \quad (7)$$

is the covariance between the bispectrum configuration $t_i = \{k_{i,1}, k_{i,2}, k_{i,3}\}$ and the configuration $t_j = \{k_{j,1}, k_{j,2}, k_{j,3}\}$.

The figure shows the correlation of 6 chosen triangles t_i with two subsets of configurations t_j : one at large scale $t_j = \{1, 1, 1\} \Delta k \dots \{6, 4, 3\} \Delta k$ and one at small scales $t_j = \{16, 15, 1\} \Delta k \dots \{16, 16, 16\} \Delta k$, as explicitly denoted on the abscissa in terms of triplets in units of Δk .

With the exception of the diagonal cases $t_i = t_j$, most of the features in the r_{ij} plots reflect random fluctuations rather than actual correlations since 300 realisations are not sufficient to accurately estimate the bispectrum covariance

matrix. A more accurate estimation of the matrix itself, limited to a single method, is presented in section 6, where we show how such fluctuations are of the same order of the expected correlations among triangles sharing, for instance, one or two sides, and it is therefore impossible to tell them apart in this figure. Nevertheless, the random noise itself in the off-diagonal elements of the N-body covariance matrix is well reproduced by all approximate methods matching the initial conditions of Minerva (that is, all except the Lognormal case), with just slightly larger discrepancies from the HALOGEN estimate.

We obtain very similar results for Sample 2, with larger discrepancies (roughly by a factor of two) for the HALOGEN and Lognormal predictions.

5 COMPARISON OF THE ERRORS ON COSMOLOGICAL PARAMETERS

In addition to the direct comparison of bispectrum measurements and their estimated covariance, we explore, as done in Papers I and II, the implications for the determination of cosmological parameters of the choice of an approximate method.

In this case we will consider a simpler likelihood analysis, compared to those assumed for the 2-point correlation function and the power spectrum in the companion papers. In the first place, the model for the halo bispectrum, described in section 5.1, is a tree-level approximation in PT and we will only consider its dependence on the linear and quadratic bias parameters, along with two shot-noise nuisance parameters. We only consider the redshift-space bispectrum monopole as the implementation and testing of loop-corrections to the galaxy bispectrum in redshift space is well beyond the scope of this work. In the second place, we will include in the likelihood only the estimate of the bispectrum variance, since 300 realisations are insufficient for any solid estimation of the covariance of more than 500 triangular configurations. We will explore quantitatively the implications of this last approximation in section 6.

We will not consider any study of the cross-correlation between power spectrum and bispectrum measurements, leaving that subject for future work.

5.1 Halo bispectrum model

We assume a tree-level model both for the matter bispectrum and the halo bispectrum.

The real-space matter bispectrum B_m is therefore given by (see, *e.g.* Bernardeau et al. 2002)

$$B_m(k_1, k_2, k_3) = 2 F_2(\mathbf{k}_1, \mathbf{k}_2) P_m^L(k_1) P_m^L(k_2) + 2 \text{ perm.} \quad (8)$$

where F_2 is the quadratic PT kernel and $P_m^L(k)$ is the linear matter power spectrum.

The halo bias model includes both local and nonlocal corrections (Baldauf et al. 2012; Chan et al. 2012; Sheth et al. 2013) so that, at second order, the halo density contrast takes the form

$$\delta_h = b_1 \delta + \frac{b_2}{2} \delta^2 + \gamma_2 \mathcal{G}_2, \quad (9)$$

where \mathcal{G}_2 is defined as

$$\mathcal{G}_2 \equiv (\nabla_{ij} \Phi_v)^2 - (\nabla^2 \Phi_v)^2, \quad (10)$$

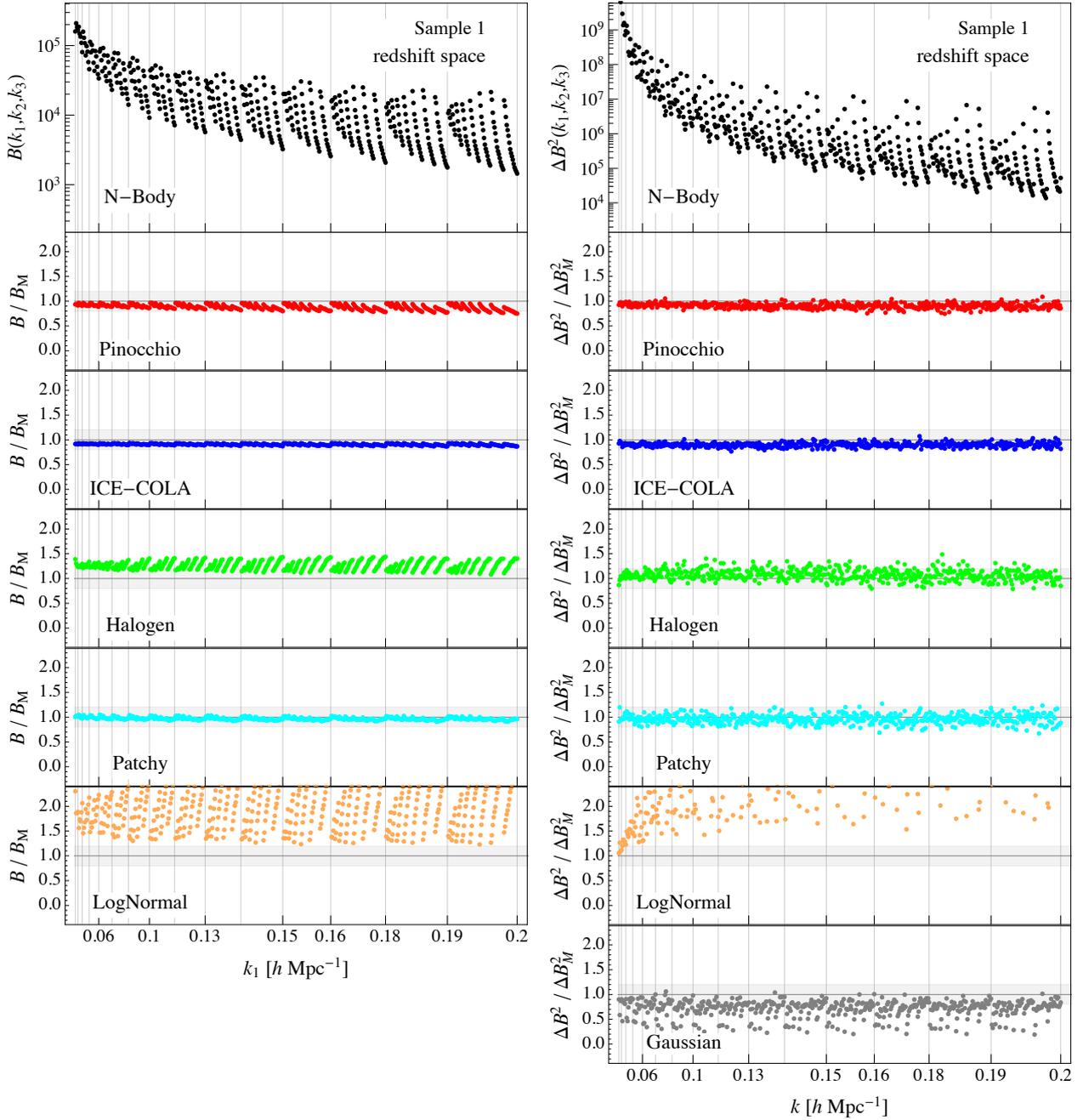


Figure 3. Average bispectrum (left column) and its variance (right column) for all triangle configurations obtained from the 300 realisations for the first mass sample in *redshift* space. The top panels show the results for the Minerva (black dots), while all other panels show the ratio between each a given estimate from an approximate method and the N-body one. In the last panel of the right column the grey dots show the ratio between the Gaussian prediction for the bispectrum variance, eq. (4), and the variance obtained from the N-body. The horizontal shaded area represents a 20% error. The vertical lines mark the triangle configurations where k_1 (the maximum of the triplet) is changing. Mocks for PEAKPATCH are not provided in the first sample so its bispectrum is missing in this case.

with Φ_v being the velocity potential such that $\mathbf{v} = \nabla\Phi_v$.

fore reads

$$\begin{aligned}
 B_h = & b_1^3 B_m(k_1, k_2, k_3) + \\
 & + b_2 b_1^2 \Sigma(k_1, k_2, k_3) + \\
 & + 2\gamma_2 b_1^2 K(k_1, k_2, k_3) + \\
 & + B_{SN}^{(1)} b_1^2 \left[P_m^L(k_1) + P_m^L(k_2) + P_m^L(k_3) \right] \\
 & + B_{SN}^{(2)}, \tag{11}
 \end{aligned}$$

The full model for the real-space halo bispectrum there-

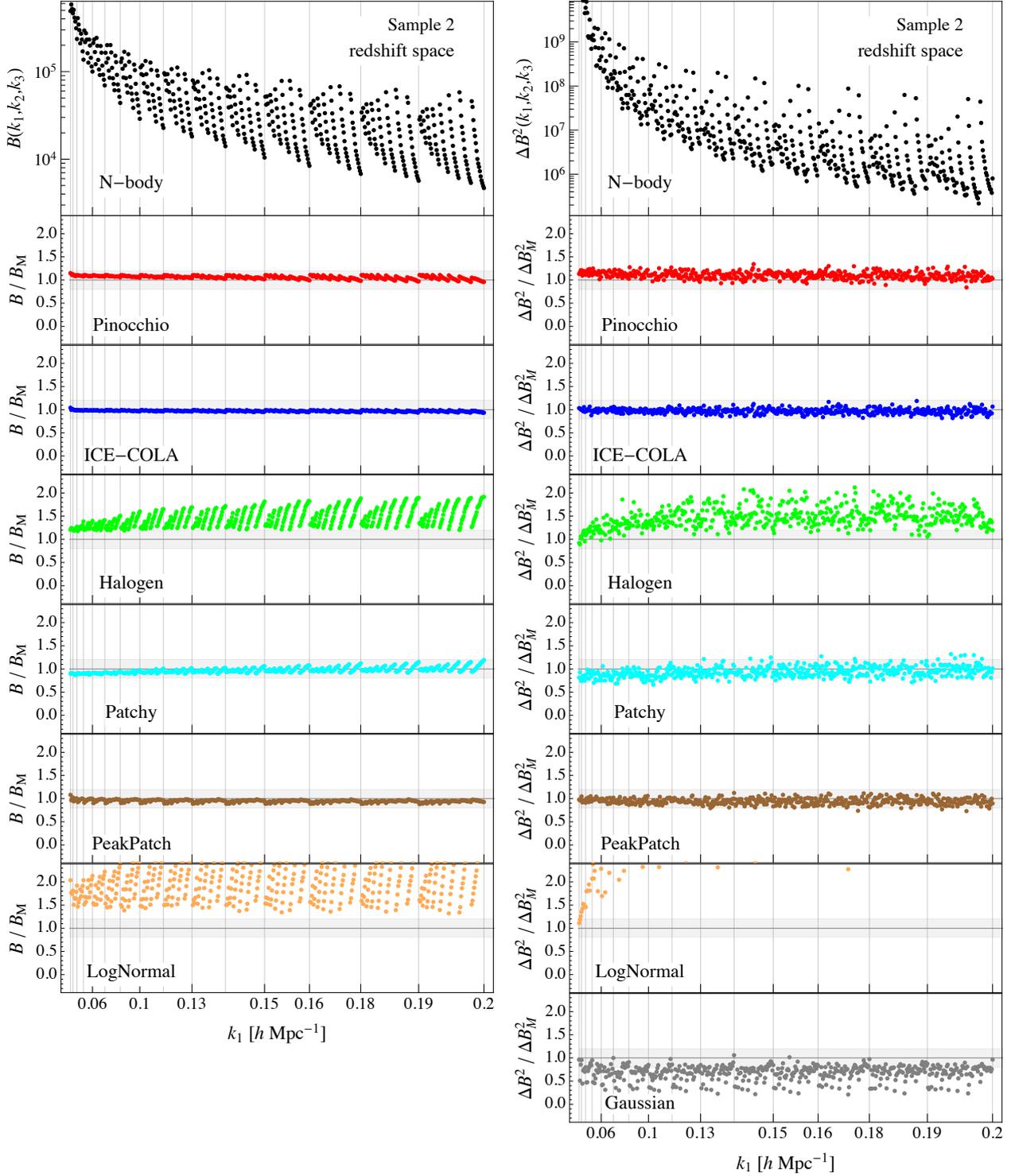


Figure 4. Same as figure 3, but for Sample 2.

where $\Sigma \equiv P_m^L(k_1)P_m^L(k_2) + 2 \text{ cyc}$ and $K \equiv (\mu_{12}^2 - 1)P_m^L(k_1)P_m^L(k_2) + 2 \text{ cyc}$, μ_{12} being the cosine of the angle between \mathbf{k}_1 and \mathbf{k}_2 . The last two contributions account for any departure from the expected shot-noise contribution under the Poisson assumption, see eq. (3). For exactly Poisson

shot-noise $B_{SN}^{(1)} = B_{SN}^{(2)} = 0$ and we will treat them here as free parameters with vanishing fiducial value.

Since we will consider the covariance for the *redshift-space* bispectrum, the corresponding model will be a slight modification accounting for the Kaiser effect on the power

spectrum and bispectrum monopoles. We will have then

$$\begin{aligned}
 B_s &= a_0^B [b_1^3 B_m(k_1, k_2, k_3) + \\
 &+ b_2 b_1^2 \Sigma(k_1, k_2, k_3) + \\
 &+ 2\gamma_2 b_1^2 K(k_1, k_2, k_3)] + \\
 &+ B_{SN}^{(1)} a_0^2 b_1^2 [P_m^L(k_1) + P_m^L(k_2) + P_m^L(k_3)] \\
 &+ B_{SN}^{(2)}, \tag{12}
 \end{aligned}$$

where, following Scoccimarro et al. (1999a); Sefusatti et al. (2006), we model redshift-space effects on the bispectrum monopole simply in terms of the factor $a_0^B = 1 + 2\beta/3 + \beta^2/9$ with $\beta = f/b_1$, f being the growth rate at $z = 1$, while $a_0 = 1 + 2\beta/3 + \beta^2/5$ is the analogous factor for the power spectrum monopole, $P_s(k) = a_0 P_h(k)$. Such correction are not having any substantial effects on our results.

The model above will therefore depend on five parameters: the local bias parameters b_1 , b_2 , the nonlocal bias parameter γ_2 and two shot noise parameters $B_{SN}^{(1)}$ and $B_{SN}^{(2)}$. We will evaluate all matter correlators for our fiducial cosmology, along with the growth rate f , and consider them as known in our analysis.

We expect this model to accurately fit simulations over a quite small range of scales, typically for $k < 0.07 h \text{Mpc}^{-1}$ (see, e.g. Sefusatti et al. 2012; Baldauf et al. 2015; Saito et al. 2014). However, we assume it to represent a full model for the halo bispectrum down to $0.2 h \text{Mpc}^{-1}$ since we are merely interested in assessing the relative effect of different estimate of the variance on parameter determination. The value of $k_{max} = 0.2 h \text{Mpc}^{-1}$ is, nevertheless, a reasonable estimate of the reach of analytical models, once loop corrections are properly included (Baldauf et al. 2015). We leave a more extensive investigation, including a joint power spectrum-bispectrum likelihood to future work.

5.2 Likelihood

We assume a Gaussian likelihood for the bispectrum given by

$$\ln \mathcal{L}_B = -\frac{1}{2} \sum_{ij} \delta B_i [C]_{ij}^{-1} \delta B_j, \tag{13}$$

where $\delta B \equiv B_{data} - B_{model}$ while C_{ij} is the bispectrum covariance matrix with indices i and j denoting individual triangular configurations t_i . The sum runs over all triangular configurations, $i = 1, \dots, N_t$, N_t being their total number corresponding to a chosen value for the smallest scale included in the analysis and determined by the parameter k_{max} . This is given by

$$N_t = \sum_{k_1=\Delta k}^{k_{max}} \sum_{k_2=\Delta k}^{k_1} \sum_{k_3=\max(\Delta k, k_1-k_2)}^{k_2} 1 \tag{14}$$

where the sums ensures that $k_1 \geq k_2 \geq k_3$ and that all triangle bins include closed fundamental triangles. As mentioned above, for $k_{max} = 0.2 h \text{Mpc}^{-1}$ we obtain $N_t = 508$.

Similarly to the analyses in Paper I and II, since we are not interested in evaluating the accuracy of the model we assume, but only to quantify the *relative* effect of replacing the variance estimated from the N-body realisations with those obtained with the approximate methods, we assume as “data” the “model” bispectrum evaluated at some fiducial

values for the parameters, that is $B_{data} = B_{model}(p_\alpha^*)$. While this lead to a vanishing χ^2 for the best fit/fiducial values, it still allows to estimate how the error on the parameters depends on the bispectrum covariance estimation.

Our choice for the parameters allows to obtain an analytical dependence of the likelihood function on them, that does not require a MonteCarlo evaluation. In fact, we can rewrite the model in eq. (11) as

$$B_{model} = \sum_{\alpha=1}^5 p_\alpha \mathcal{B}_\alpha \tag{15}$$

where $\{p_\alpha\} = \{a_0^B b_1^3, a_0^B b_1^2 b_2, a_0^B b_1^2, \gamma_2, a_0^2 b_1^2 B_{SN}^{(1)}, B_{SN}^{(2)}\}$ and $\{\mathcal{B}_\alpha\} = \{B_m, \Sigma, 2K, P_m(k_1) + P_m(k_2) + P_m(k_3), 1\}$. Adding $p_0 = -1$ and $\mathcal{B}_0 = B_{data}$ we can also write

$$-\delta B = B_{model} - B_{measured} = \sum_{\alpha=0}^5 p_\alpha \mathcal{B}_\alpha \tag{16}$$

and therefore it is easy to see that we can rewrite the likelihood as

$$\ln \mathcal{L}_B = -\frac{1}{2} \sum_{\alpha, \beta} p_\alpha p_\beta \mathcal{D}_{\alpha\beta}, \tag{17}$$

where

$$\mathcal{D}_{\alpha\beta} \equiv \sum_{i,j=1}^{N_t} \mathcal{B}_\alpha(t_i) [C^B]_{ij}^{-1} \mathcal{B}_\beta(t_j). \tag{18}$$

In this way the likelihood \mathcal{L}_B is explicitly written as an exact, multivariate Gaussian distribution in the parameters p_α . Clearly, once the quantities $\mathcal{D}_{\alpha\beta}$ are computed, we can evaluate any marginalisation analytically. We could, in principle consider a transformation between these parameters and the set given by $\{b_1, b_2, \gamma_2, B_{SN}^{(1)}, B_{SN}^{(2)}\}$ but this would require an approximation for the likelihood around its maximum and, furthermore, it would not add any information to our goal since any relative variation on the error on the parameter cube b_1^3 , for instance, is of the same order as the relative variation on the error on b_1 .

In practice, with the exception of the tests presented in section 6, we will only compare estimates of the bispectrum *variance* from the 300 runs sets for the various methods, so that we evaluate effectively

$$\mathcal{D}_{\alpha\beta} \simeq \sum_{i=1}^{N_t} \frac{\mathcal{B}_\alpha(t_i) \mathcal{B}_\beta(t_j)}{\Delta B^2(t_i)}, \tag{19}$$

$\Delta B^2(t_i)$ representing the variance for the triangular configuration t_i .

5.3 Parameters constraints comparison

Figure 6 shows the ratio between the marginalised error on each parameter p_α obtained from the variance estimated with a given approximate method and the same marginalised error on the same parameter obtained from the variance estimated from the Minerva N-body set. Such ratio is shown as a function of the maximum wavenumber k_{max} assumed for the likelihood evaluation that defines as well the total number of configurations N_t according to eq. (14). The left column corresponds to Sample 1 while the right column to

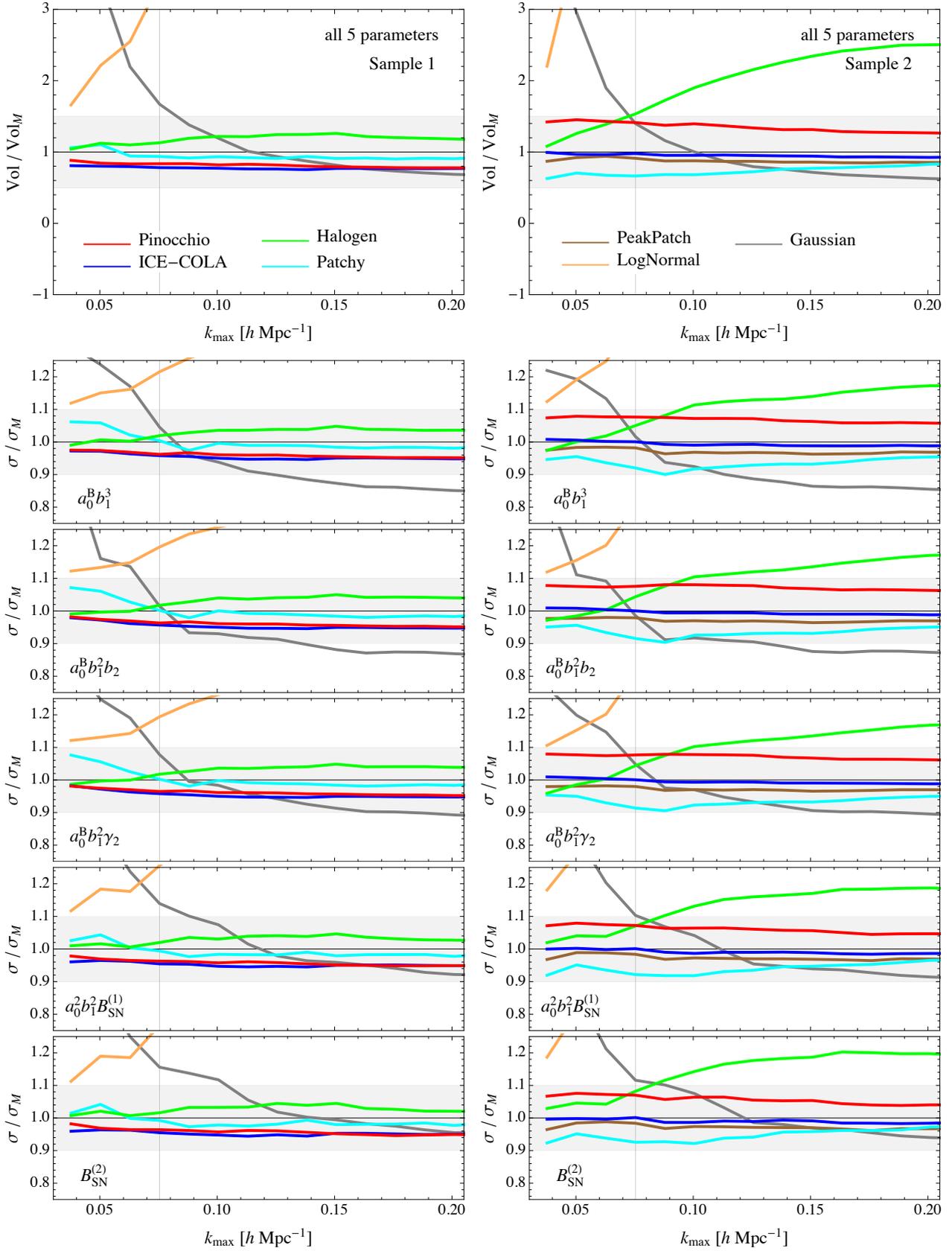


Figure 6. Marginalized errors for the bias parameters in using the bispectrum monopole in redshift space for the two samples (first and second column) compared with the error obtained from N-body estimate of the variance. See text for explanation.

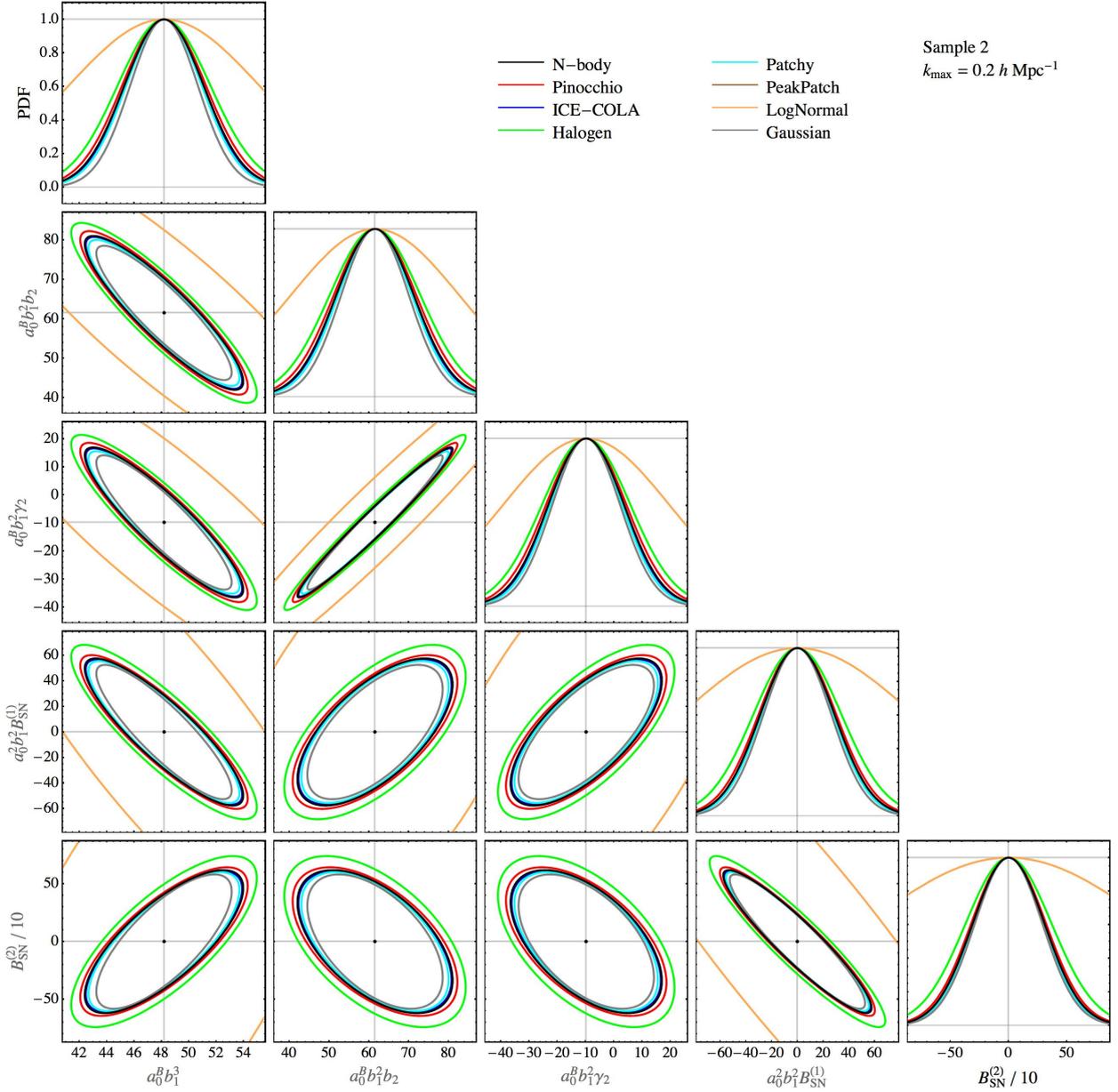


Figure 7. $2\text{-}\sigma$ contour plots for the parameters combinations p_α (see text) from the bispectrum monopole in redshift space for Sample 2. The constraints assume $k_{\text{max}}=0.2 h \text{ Mpc}^{-1}$. Notice that the N-body (black) results are plotted on top so that a few curves, corresponding to methods very closely reproducing the N-body, ones are not easily visible.

Sample 2. The grey shaded area represent a 10% discrepancy between error estimates.

In addition to the errors on individual parameters we consider, as in the companion papers, the volume of the, in our case, 5-dimensional ellipsoid corresponding to the combined errors on all parameters defined as

$$\text{Vol} = \sqrt{\det \mathcal{D}_{\alpha\beta}^{-1}}, \quad (20)$$

where $\mathcal{D}_{\alpha\beta}^{-1}$ represents the parameters covariance matrix. The ratio of this quantity estimated from the approximate methods and from the N-body runs is shown in the two top panels of figure 6 for the two samples. In this case, the shaded area corresponds to a discrepancy of 50%, reflecting the target 10% for individual parameters.

These results reflect those shown in the comparison of the variance. Unsurprisingly the methods that overestimate the variance lead to an overestimate of the error on each parameter, in a similar fashion across all parameters. As already shown in the previous figures, the predictive methods, along with PATCHY, appear to be more accurate, with ICE-COLA, in particular, the one providing more consistent results for both samples. All such methods show discrepancies of less than 10% w.r.t. the N-body case. The behaviour of HALOGEN is also quite good in the low-mass sample but the difference with N-body becomes larger than 10% in the second sample once small scales are included. LogNormal shows the largest difference, with reasonable results only for the very large scales. The Gaussian prediction provides an

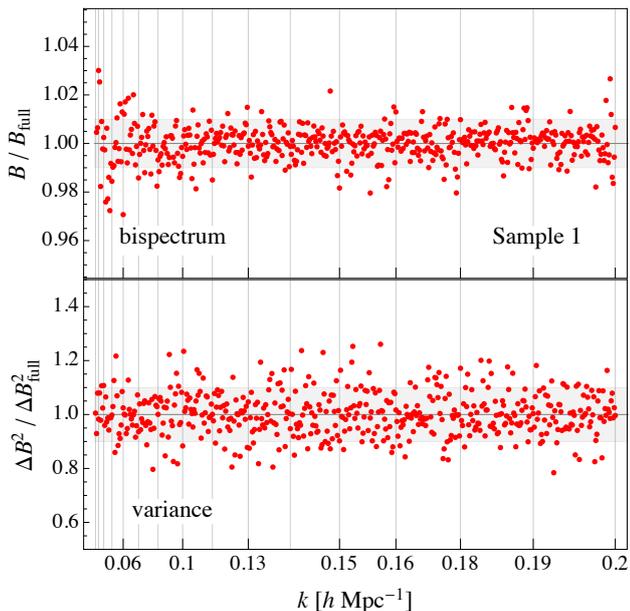


Figure 8. Ratio between the bispectrum (top) and its variance (bottom) as measured in 300 realisations of PINOCCHIO to the same quantities estimated from 10,000 realisations in real space. Both assume Sample 1.

overestimate of the errors at large scales and an underestimate at small scales, particularly in the case of the parameters more directly related to bias, probably due to a missing non-Gaussian component.

Finally, figure 7 as an example, shows the $2\text{-}\sigma$ contour plots for the parameters combinations p_α in redshift space. Similar results are obtained for Sample 1. One can notice, in particular, that no method provides a variance estimate that affects the degeneracies between parameters in any specific way. Such effect might be more relevant when the full covariance is taken into account. We will comment on this in the next section.

6 TESTS WITH A LARGE SET OF REALISATIONS

The number of 300 realisations, despite being quite a large number for many applications, is still rather small when it comes to estimate the covariance of hundreds or thousands of bispectrum configurations. For this reason we limited our likelihood comparison to its dependence on the bispectrum variance alone.

In this section we test the robustness of some of our conclusions taking advantage of a much larger sets of 10,000 PINOCCHIO catalogs characterised by the same configuration and cosmology as the 300 so far considered.

In figure 8 we show the ratios of the real-space bispectrum and its variance obtained from 300 realisations and the same quantities obtained from the 10,000 runs for Sample 1. The scatter on the bispectrum due to the limited number of runs is of the order of a few percent, while for the variance is of the order of 10%, with no particular dependence on

shape. Essentially identical results can be shown for Sample 2.

Figure 9 shows, again for Sample 1, with dashed curves the cross-correlation coefficients defined as

$$r_{ij,full} \equiv \frac{C_{ij}}{\sqrt{C_{ij,full}C_{jj,full}}} \quad (21)$$

where $C_{ij,full}$ represents the covariance between triangles t_i and t_j estimated from the 10,000 runs, while the C_{ij} in the numerator represents the covariance from only 300 realisations. The continuous curves denotes instead the cross-correlation coefficients estimated completely from the 10,000 runs. It is interesting to notice how the noise characterising the first estimates is of the order of the true off-diagonal correlations from the less noisy estimate, present, as expected, between configurations sharing one or two wavenumbers, e.g. $t_i = \{2, 2, 2\}$ and $t_j = \{16, 15, 2\}$.

Finally, figure 10 shows various comparisons between the volume error as defined in eq. (20) for the 5 parameters combinations p_α (in real space) obtained from three different assumptions for the likelihood functions. In the first case, the likelihood is defined in terms of the variance alone determined from 300 realisations, denoted in the figure as “Var(300)”. This is the case adopted for the results in section 5. In the second case, the likelihood still depends only on the variance, now obtained from 10,000 runs, “Var(10,000)”. Finally, in the last case the likelihood is defined in terms of the full bispectrum covariance, as in eq. (13), estimated again from 10,000 runs, “Cov(10,000)”. The comparison is provided in terms of the ratio between the volumes Vol obtained from the different likelihood. We notice first that essentially no difference is found between variance estimates employing 300 or 10,000 runs, at least up to $k_{max} = 0.2 h \text{ Mpc}^{-1}$. More significant differences are evident, instead, between these cases and the likelihood based on the full covariance estimation. For both mass samples the difference is always below the 10%. The fact that we encounter no dramatic difference is somehow reassuring w.r.t. our previous results. It is nevertheless evident that as we move to smaller scales the variance-only likelihood increasingly underestimate the errors, although no more than 10% at $k_{max} = 0.2 h \text{ Mpc}^{-1}$.

7 CONCLUSIONS

In this paper, and in its companion Papers I and II, we have studied the problem of covariance matrix estimation for large-scale structure observables using dark matter halo catalogs produced with approximate methods. This last paper, in particular, focuses on the halo bispectrum and its covariance matrix, with the twofold aim of assessing the correct reproduction of the non-Gaussian properties of the halo distribution as well as considering the halo/galaxy bispectrum as a direct observable in its own right.

The measurements are performed on sets of 300 (1000 for LogNormal) catalogs obtained from several different methods: ICE-COLA, PEAKPATCH, PINOCCHIO, HALOGEN, PATCHY, LogNormal and they are compared with the reference Minerva suite of 300 N-body simulations. All approximate catalogs, apart from LogNormal, assume the same initial conditions of the full N-body simulations, thereby reducing differences due to cosmic variance. Out of each halo

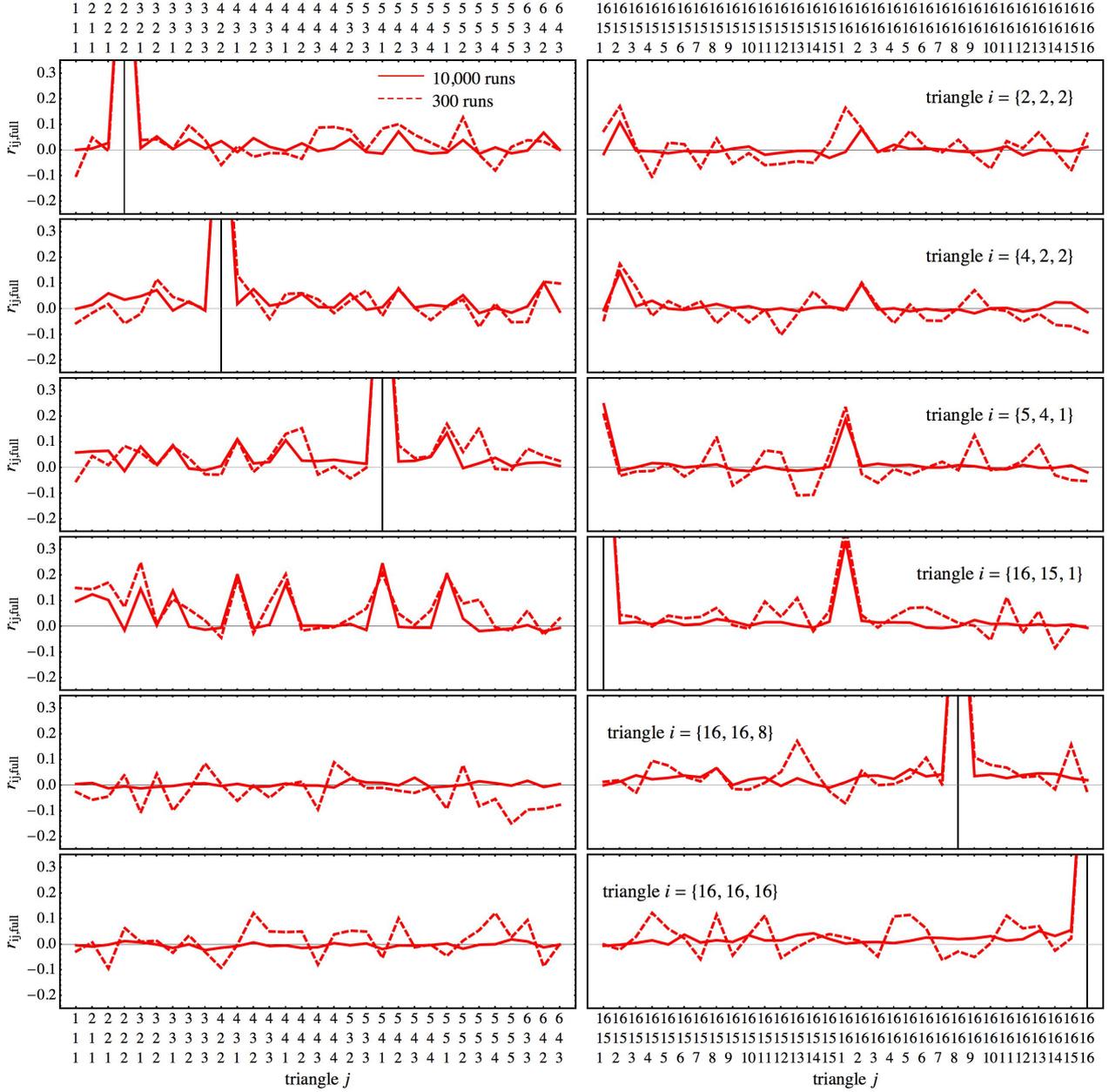


Figure 9. Cut through the cross-correlation coefficient in real space for all the triangle configurations coefficients estimated from 300 realizations (dashed line) and 10,000 realizations (continuous line) for the first sample. On the x-axis there are the triplets for each triangle in fundamental frequency unit. The cross-correlation coefficient is normalized to the Minerva variance.

catalog we select two samples characterised by a different minimal mass in order to gain a better perspective on our results as a function of mass and shot-noise levels.

The approximate methods can be generically subdivided into predictive methods (ICE-COLA, PINOCCHIO, PEAKPATCH), requiring a single redefinition of the halo mass to recover the expected halo number density, and methods (HALOGEN, PATCHY), requiring as well a calibration of the bias function. It should be noted that, in the case of HALOGEN, such bias calibration is limited to the 2-Point Correlation Function and to configuration space, with only one parameter (per mass-bin) controlling the clustering amplitude. In addition, a third type is represented by the Log-

normal method, relying on a non-linear transformation of the matter density field, in turn calibrated on the halo mass function and halo bias. In all our analysis (with the exception of Appendix A) we have changed the limiting mass for each sample in order to ensure the same abundance for all catalogs, including those obtained with more predictive methods.

We have shown that:

(i) the real space bispectrum is reproduced by ICE-COLA, PINOCCHIO, PATCHY and PEAKPATCH within 20% for the most of the triangle configurations while HALOGEN and, particularly, Lognormal present larger disagreements, often beyond 50%;

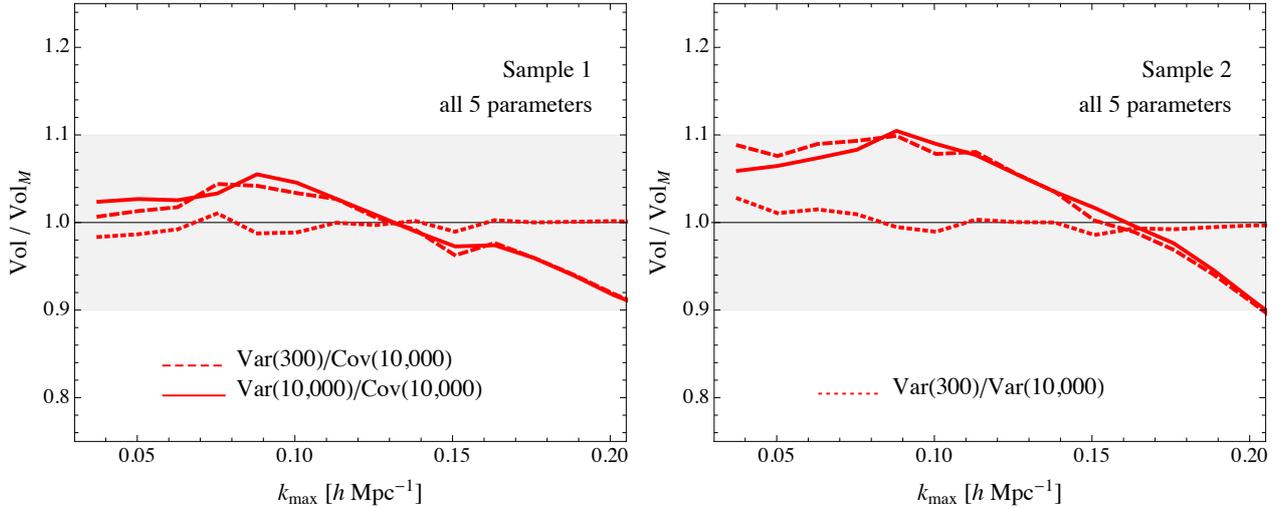


Figure 10. Errors volume, as defined in eq. (20) for the bias parameters using the bispectrum variance from 10,000 realizations (continuous line) and the variance from 300 realizations (dashed and dotted line) compared with the full covariance from 10,000 or 300 realizations. In the first column the results are shown for the first sample, in the second column for the second sample.

(ii) these discrepancies are reflected on the results for the bispectrum variance, where, however, their systematic nature is less evident since there is no clear dependence on the triangle shape, probably due to the fact that for most triangles, the variance is dominated by the shot-noise component; the Gaussian prediction for the variance is generically underestimating the N-body result, particularly for squeezed triangles;

(iii) similar conclusions can be made for the redshift-space bispectrum monopole, where, however, PATCHY and HALOGEN (the latter at least for the small mass sample) show a better agreement with the N-body simulations;

(iv) the inspection of the cross correlation coefficients illustrates how, due to the matching initial conditions, almost all methods (except Lognormal by construction) reproduce the noise present in the N-body estimation, which is dominating the off-diagonal elements of the covariance matrix estimated from only 300 realisations.

Our analysis was not limited to how accurately the bispectrum and its covariance are recovered, but include a comparison of the errors on cosmological parameters, in this case linear and non-linear bias parameters, derived from each approximate estimate of the variance of the halo bispectrum in redshift space. This last step was out of necessity restricted to the variance as the relatively large set of 300 realisations is still not sufficient for a robust estimation of the full covariance of the hundreds of triangular configurations considered.

As in the similar analysis performed in the companion papers, we assumed a model for the bispectrum and produced a data vector from the evaluation of such model at some chosen fiducial value for the parameters. This allowed us to focus our attention exclusively on the errors recovered as a function of the different estimation of the covariance matrix. Differently from the companion papers, the model considered based on tree-level perturbation theory, only depends on bias and shot-noise parameters, allowing a much easier evaluation of the likelihood function. In particular, under these simplified settings we can easily

compute our results as a function of the smallest scale, or maximum wavenumber k_{max} , included in the analysis. More rigorous tests involving additional cosmological parameters, a more accurate modelling of the redshift-space bispectrum in the quasi-linear regime and a solid estimate of the full bispectrum covariance matrix (and cross-correlation with the power spectrum) are clearly well beyond the scope of this comparison project but will be required in the near future for the proper exploitation of the galaxy bispectrum as a relevant observable.

The parameter error comparison has shown that:

(i) the error on the bias and on the shot-noise parameters are reproduced within 10% by all the methods except Lognormal and HALOGEN in the high-mass sample for $k_{max} > 0.07$. This is evident as well in terms of the combined error volume as defined in eq. (20); for the second sample PINOCCHIO, and to a lesser extent PATCHY, show a higher level of disagreement compared with the other predictive methods;

(ii) the Gaussian prediction tends to underestimate the error on some parameters for large values of k_{max} ;

(iii) the results provided by the contour plots, for both mass samples and for different values of k_{max} (not all shown in the figures), do not show any relevant discrepancy in terms of parameter degeneracies, in addition to errors size present in the fully marginalised results.

To sum up, predictive methods, along with PATCHY appear to be the most accurate in reproducing the N-body results, but differences are quite overall relatively small. Of course, our likelihood test has been limited to include the bispectrum variance, due to the relatively small number of N-body runs available. For this reason, we included an additional test employing 10,000 PINOCCHIO realisations to compare, at least for this particular method, the variance estimated from 300 realisation to the variance and the full covariance estimated from the whole set. This has shown that

(i) the variance estimate is not particularly affected by the limited number of 300 runs and essentially no difference is found on the results for the parameters errors;

(ii) the results in terms of the full covariance, instead, do provide differences on the parameters errors but still within 10%, although they highlight a progressive underestimate of the errors based on the variance alone beyond $k_{max} \simeq 0.15 h \text{ Mpc}^{-1}$, where a steady deviation proportional to k_{max} is observed.

Clearly, a more realistic investigation of the relevance of a reliable estimate of the bispectrum covariance matrix requires a proper model for the quasi-linear regime that we will leave for future work. In addition, we should also expect that the relatively small difference between the results obtained from the variance alone and the full covariance will become more relevant once a realistic window function is accounted for as beat-coupling/super-sample covariance effects are expected to provide additional contributions also to off-diagonal elements. Since such effects depend directly on the non-Gaussian properties of the galaxy/halo distribution, we consider the present work only as a first step toward a more complete assessment of the correct recovery of non-Gaussianity by approximate methods for mock catalogs.

From the analysis we have presented it appears that most of the methods we considered are capable to reproduce the halo bispectrum, its variance and the errors on bias parameters based on the variance alone quite accurately. This is particularly true for predictive methods such as ICE-COLA, PINOCCHIO and PEAKPATCH. Similar results are obtained for PATCHY, although the calibration in redshift space might lead to some larger systematic for the real-space bispectrum that in turn could have effects not investigated in this work (*e.g.* finite-volume effects). For what concern HALOGEN, we have already stressed that its calibration is restricted to the two-point statistic so a lower accuracy on the bispectrum might be somehow expected. Nevertheless is worth to point out that the marginalized errors on the parameters in redshift space, in particular for the first sample, are certainly comparable with all the other methods except for Lognormal. This last method, in fact, is the one that fares worst among those considered. This can be expected since, as already mentioned, the nonlinear transformation on the density field that provides a qualitatively reasonable description of the nonlinear power spectrum, while providing a non-Gaussian contribution, does not ensure that such contribution, for instance in the case of the bispectrum, presents the correct functional form and dependence on the triangular configuration shape.

We notice finally how our tests on the bispectrum have highlighted differences among the different methods that are less evident from the similar analysis on two-point statistic performed in the companion papers I and II. This illustrates how the bispectrum can be a useful diagnostic for this type of comparisons, even when we are not directly interested in the bispectrum as an observable. We expect that possible direction of investigation along these lines will include correlators of realistic galaxy distribution and, particularly for Fourier-space statistics, finite-volume effects, in order to better assess the interplay between non-Gaussianity, convolution with a window function and realistic shot-noise contributions.

ACKNOWLEDGMENTS

M. Colavincenzo is supported by the *Departments of Excellence 2018 - 2022* Grant awarded by the Italian Ministero dell'Istruzione, dell'Università e della Ricerca (MIUR) (L. 232/2016), by the research grant *The Anisotropic Dark Universe* Number CSTO161409, funded under the program CSP-UNITO *Research for the Territory 2016* by Compagnia di Sanpaolo and University of Torino; and the research grant TAsP (Theoretical Astroparticle Physics) funded by the Istituto Nazionale di Fisica Nucleare (INFN). P. Monaco and E. Sefusatti acknowledge support from a FRA2015 grant from MIUR PRIN 2015 *Cosmology and Fundamental Physics: illuminating the Dark Universe with Euclid* and from Consorzio per la Fisica di Trieste; they are part of the INFN InDark research group.

L. Blot acknowledges support from the Spanish Ministerio de Economía y Competitividad (MINECO) grant ESP2015-66861. M.Crocce acknowledges support from the Spanish Ramón y Cajal MICINN program. M.Crocce has been funded by AYA2015-71825.

M. Lippich and A.G.Sánchez acknowledge support from the Transregional Collaborative Research Centre TR33 *The Dark Universe* of the German Research Foundation (DFG).

C. Dalla Vecchia acknowledges support from the MINECO through grants AYA2013-46886, AYA2014-58308 and RYC-2015-18078. S. Avila acknowledges support from the UK Space Agency through grant ST/K00283X/1. A. Balaguera-Antolínez acknowledges financial support from MINECO under the Severo Ochoa program SEV-2015-0548. M. Pellejero-Ibanez acknowledges support from MINECO under the grand AYA2012-39702-C02-01. P. Fosalba acknowledges support from MINECO through grant ESP2015-66861-C3-1-R and Generalitat de Catalunya through grant 2017-SGR-885. A. Izard was supported in part by Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. He was also supported in part by NASA ROSES 13-ATP13-0019, NASA ROSES 14-MIRO-PROS-0064, NASA ROSES 12- EUCLID12-0004, and acknowledges support from the JAE program grant from the Spanish National Science Council (CSIC). R. Bond, S. Codis and G. Stein are supported by the Canadian Natural Sciences and Engineering Research Council (NSERC). G. Yepes acknowledges financial support from MINECO/FEDER (Spain) under research grant AYA2015-63810-P.

The Minerva simulations have been performed and analysed on the Hydra and Euclid clusters at the Max Planck Computing and Data Facility (MPCDF) in Garching.

PINOCCHIO mocks were run on the GALILEO cluster at CINECA thanks to an agreement with the University of Trieste.

ICE-COLA simulations were run at the MareNostrum supercomputer - Barcelona Supercomputing Center (BSC-CNS, www.bsc.es), through the grant AECT-2016- 3-0015.

PEAKPATCH simulations were performed on the GPC supercomputer at the SciNet HPC Consortium. SciNet is funded by: the Canada Foundation for Innovation under the auspices of Compute Canada; the Government of Ontario; Ontario Research Fund - Research Excellence; and the University of Toronto.

Numerical computations with HALOGEN were done

on the Sciama High Performance Compute (HPC) cluster which is supported by the ICG, SEPNet and the University of Portsmouth.

PATCHY mocks have been computed in part at the MareNostrum supercomputer of the Barcelona Supercomputing Center thanks to a grant from the Red Española de Supercomputación (RES), and in part at the Teide High-Performance Computing facilities provided by the Instituto Tecnológico y de Energías Renovables (ITER, S.A.).

This work, as the companion papers, has been conceived and developed as part of the joint activity of the “Galaxy Clustering” and the “Cosmological Simulations” Science Working Groups of the Euclid survey consortium.

This paper and companion papers have benefited of discussions and the stimulating environment of the Euclid Consortium, which is warmly acknowledged.

REFERENCES

- Agrawal A., Makiya R., Chiang C.-T., Jeong D., Saito S., Komatsu E., 2017, *J. Cosmology Astropart. Phys.*, **10**, 003
- Avila S., Murray S. G., Knebe A., Power C., Robotham A. S. G., Garcia-Bellido J., 2015, *MNRAS*, **450**, 1856
- Avila S., et al., 2017, preprint, ([arXiv:1712.06232](https://arxiv.org/abs/1712.06232))
- Baldauf T., Seljak U., Desjacques V., McDonald P., 2012, *Phys. Rev. D*, **86**, 083540
- Baldauf T., Mersilli L., Mirbabayi M., Pajer E., 2015, *J. Cosmology Astropart. Phys.*, **5**, 007
- Bernardeau F., Colombi S., Gaztañaga E., Scoccimarro R., 2002, *Phys. Rep.*, **367**, 1
- Blot L., Corasaniti P. S., Alimi J.-M., Reverdy V., Rasera Y., 2015, *MNRAS*, **446**, 1756
- Blot L., et al., 2018, Comparing approximate methods for mock catalogues covariances II: Power spectrum, in preparation
- Bond J. R., Myers S. T., 1996a, *ApJS*, **103**, 1
- Bond J. R., Myers S. T., 1996b, *ApJS*, **103**, 41
- Bond J. R., Myers S. T., 1996c, *ApJS*, **103**, 63
- Chan K. C., Scoccimarro R., Sheth R. K., 2012, *Phys. Rev. D*, **85**, 083509
- Chuang C.-H., Kitaura F.-S., Prada F., Zhao C., Yepes G., 2015a, *MNRAS*, **446**, 2621
- Chuang C.-H., et al., 2015b, *MNRAS*, **452**, 686
- Colavincenzo M., Monaco P., Sefusatti E., Borgani S., 2017, *J. Cosmology Astropart. Phys.*, **3**, 052
- Coles P., Jones B., 1991, *MNRAS*, **248**, 1
- Gaztañaga E., Cabré A., Castander F., Crocce M., Fosalba P., 2009, *MNRAS*, **399**, 801
- Gil-Marín H., Noreña J., Verde L., Percival W. J., Wagner C., Manera M., Schneider D. P., 2015a, *MNRAS*, **451**, 539
- Gil-Marín H., et al., 2015b, *MNRAS*, **452**, 1914
- Gil-Marín H., Percival W. J., Verde L., Brownstein J. R., Chuang C.-H., Kitaura F.-S., Rodríguez-Torres S. A., Olmstead M. D., 2017, *MNRAS*, **465**, 1757
- Grieb J. N., Sánchez A. G., Salazar-Albornoz S., Dalla Vecchia C., 2016, *MNRAS*, **457**, 1577
- Hamilton A. J. S., Rimes C. D., Scoccimarro R., 2006, *MNRAS*, **371**, 1188
- Izard A., Crocce M., Fosalba P., 2016, *MNRAS*, **459**, 2327
- Kitaura F.-S., Yepes G., Prada F., 2014, *MNRAS*, **439**, L21
- Kitaura F.-S., et al., 2016, *MNRAS*, **456**, 4156
- Koda J., Blake C., Beutler F., Kazin E., Marin F., 2016, *MNRAS*, **459**, 2118
- Lippich M., et al., 2018, Comparing approximate methods for mock catalogues covariances I: Correlation Function, in preparation
- Manera M., et al., 2013, *MNRAS*, **428**, 1036
- Matarrese S., Verde L., Heavens A. F., 1997, *MNRAS*, **290**, 651
- Meiksin A., White M., 1999, *MNRAS*, **308**, 1179
- Monaco P., 2016, *Galaxies*, **4**, 53
- Monaco P., Sefusatti E., Borgani S., Crocce M., Fosalba P., Sheth R. K., Theuns T., 2013, *MNRAS*, **433**, 2389
- Munari E., Monaco P., Sefusatti E., Castorina E., Mohammad F. G., Anselmi S., Borgani S., 2017, *MNRAS*, **465**, 4658
- Ngan W., Harnois-Déraps J., Pen U.-L., McDonald P., MacDonald I., 2012, *MNRAS*, **419**, 2949
- Pearson D. W., Samushia L., 2017, preprint, ([arXiv:1712.04970](https://arxiv.org/abs/1712.04970))
- Rimes C. D., Hamilton A. J. S., 2006, *MNRAS*, **371**, 1205
- Saito S., Baldauf T., Vlah Z., Seljak U., Okumura T., McDonald P., 2014, *Phys. Rev. D*, **90**, 123522
- Sánchez A. G., et al., 2013, *MNRAS*, **433**, 1202
- Scoccimarro R., 1998, *MNRAS*, **299**, 1097
- Scoccimarro R., 2000, *ApJ*, **544**, 597
- Scoccimarro R., 2015, *Phys. Rev. D*, **92**, 083532
- Scoccimarro R., Sheth R. K., 2002, *MNRAS*, **329**, 629
- Scoccimarro R., Colombi S., Fry J. N., Frieman J. A., Hivon E., Melott A., 1998, *ApJ*, **496**, 586
- Scoccimarro R., Couchman H. M. P., Frieman J. A., 1999a, *ApJ*, **517**, 531
- Scoccimarro R., Zaldarriaga M., Hui L., 1999b, *ApJ*, **527**, 1
- Sefusatti E., Crocce M., Pueblas S., Scoccimarro R., 2006, *Phys. Rev. D*, **74**, 023522
- Sefusatti E., Crocce M., Desjacques V., 2012, *MNRAS*, **425**, 2903
- Sefusatti E., Crocce M., Scoccimarro R., Couchman H. M. P., 2016, *MNRAS*, **460**, 3624
- Sheth R. K., Chan K. C., Scoccimarro R., 2013, *Phys. Rev. D*, **87**, 083002
- Slepian Z., et al., 2017, *MNRAS*, **469**, 1738
- Springel V., Yoshida N., White S. D. M., 2001, *New Astron.*, **6**, 79
- Takada M., Hu W., 2013, *Phys. Rev. D*, **87**, 123504
- Takahashi R., et al., 2009, *ApJ*, **700**, 479
- White M., Tinker J. L., McBride C. K., 2014, *MNRAS*, **437**, 2594
- de Putter R., Wagner C., Mena O., Verde L., Percival W. J., 2012, *J. Cosmology Astropart. Phys.*, **4**, 19
- de la Torre S., et al., 2013, *A&A*, **557**, A54

APPENDIX A: MASS-CUT VS ABUNDANCE MATCHING

We have seen how predictive methods perform better overall than methods requiring calibration with a set of N-body simulations. However, all our results did assume, including predictive ones, that the halo density matches the one from the N-body catalogs to match the halo density from the N-body catalogs. In this appendix we compare the results presented so far and those obtained from PINOCCHIO, ICE-COLA and PEAKPATCH when their predictions are taken out-of-the-box with no abundance matching. Since each method has a different definition of the mass, a constant mass cut will typically pick up different objects. This is especially true for PEAKPATCH halos which are defined as spherical overdensities in Lagrangian space and are not meant to reproduce FoF masses.

Figure A1 shows the ratio of the bispectrum (left column) and its variance (right column) to the N-body results (similarly to figures 3 and 4) in redshift space comparing the case of density matching (full color) assumed so far to the case where the limiting mass is not changed (faded color). Both mass samples are shown and we remind the reader that the PEAKPATCH catalogs are only available for Sample 2.

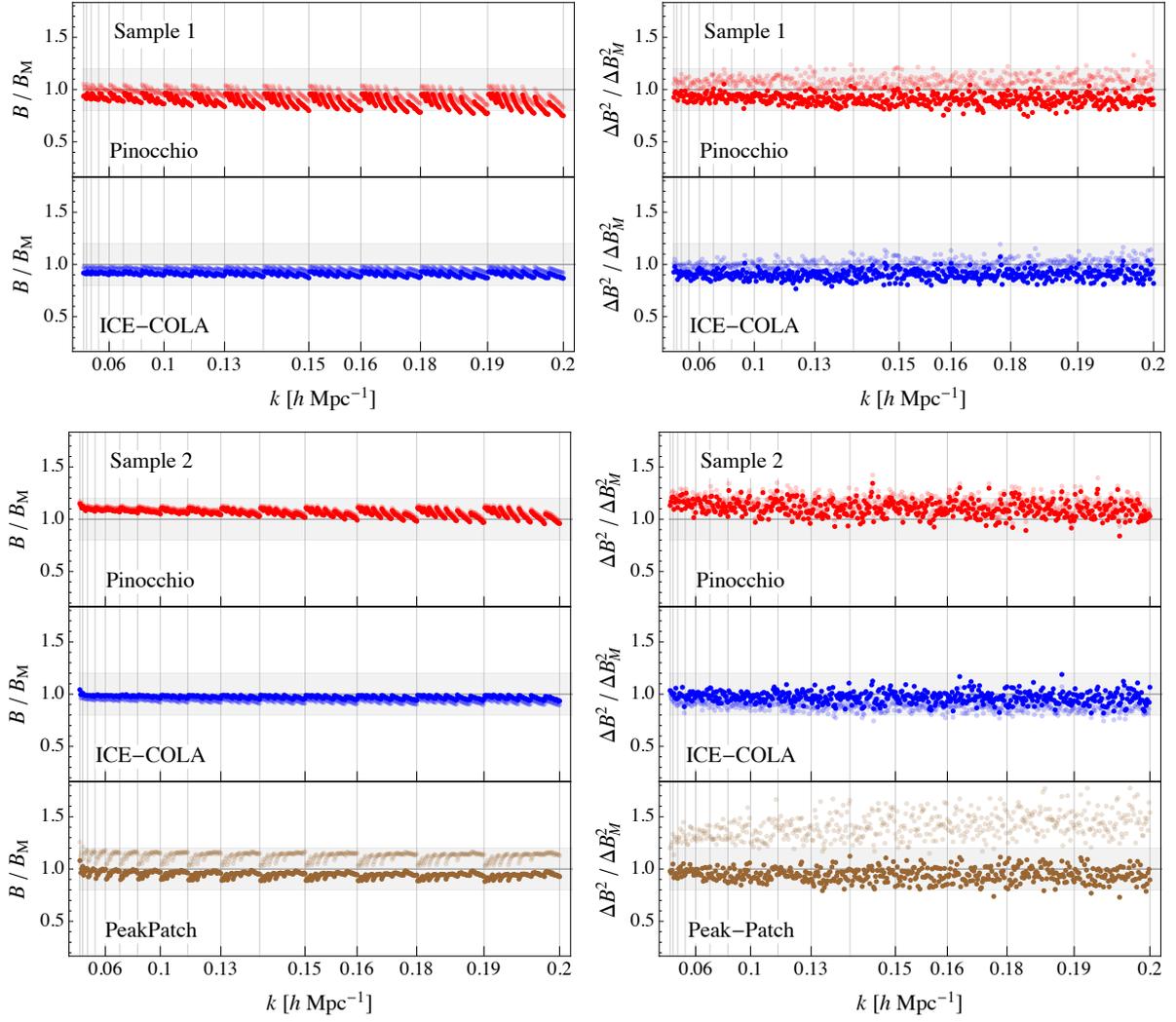


Figure A1. Bispectrum and its variance. Comparison of density matching (full color) to mass-cut (faded color), redshift space.

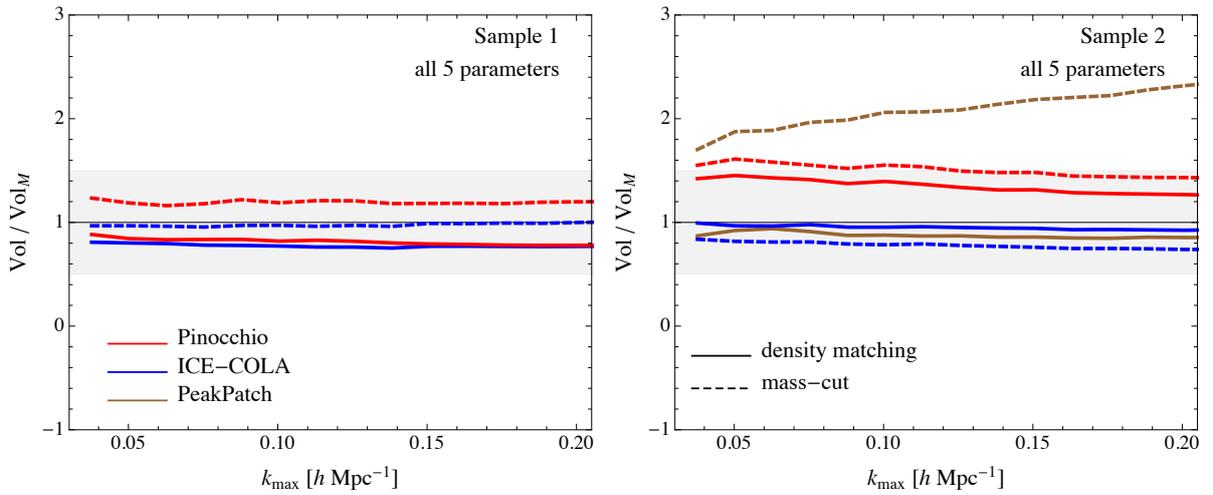


Figure A2. Marginalized errors for the bias parameters using the real bispectrum for the two samples (first and second column) compared with the error obtained using Minerva. Density cuts are displayed with solid lines while dashed lines represent mass cuts. The gray shaded area represent the 10% error on individual parameters, or 50% on the 5-parameters error volume.

For the bispectrum the difference between the density matching and the mass-cut are lower than 10% for PINOCCHIO and ICE-COLA for both the samples, while PEAKPATCH shows a larger difference, but always smaller than 20%, with density matching performing better as we can expect. For the variance the differences appear to be larger. ICE-COLA and PINOCCHIO present, respectively, differences of the order of 15 to 25% for the first sample but smaller in the second sample case. PEAKPATCH, on the other hand shows a difference of about 40% for Sample 2.

Finally, figure A2 shows the combined error volume relative to the N-Body results, as in figure 6, for the two samples, comparing density matching (continuous lines) to the case of direct mass-cut (dashed lines). Using the measurements from the mass-cut case, for both samples, we recover larger errors, as can be expected from the variance comparison, with differences of the order of 10% on the individual parameter error (50% on the 5-parameter volume shown in the figure) for PINOCCHIO. An even larger difference is found for PEAKPATCH, while discrepancies for ICE-COLA are within 5% for both samples.