Constraining the halo bispectrum in real and redshift space from perturbation theory and non-linear stochastic bias

Francisco-Shu Kitaura[1]*, Héctor Gil-Marín[2], Claudia Scoccola[3,4,5,6], Chia-Hsun Chuang[5,6],

Volker Müller[1], Gustavo Yepes[5] & Francisco Prada[6,7,8]

[1]*Leibniz-Institut für Astrophysik Potsdam (AIP), An der Sternwarte 16, D-14482 Potsdam, Germany*
[2]*Institute of Cosmology & Gravitation, University of Portsmouth, Dennis Sciama Building, Portsmouth PO1 3FX, UK*
[3]*Instituto de Astrofísica de Canarias (IAC), C/Vía Láctea, s/n, E-38200 La Laguna, Tenerife, Spain*
[4]*Departamento de Astrofísica, Universidad de La Laguna (ULL), E-38206 La Laguna, Tenerife, Spain*
[5]*Departamento de Física Teórica, Universidad Autónoma de Madrid, Cantoblanco, 28049, Madrid, Spain*
[6]*Instituto de Física Teórica, (UAM/CSIC), Universidad Autónoma de Madrid, Cantoblanco, E-28049 Madrid, Spain*
[7]*Campus of International Excellence UAM+CSIC, Cantoblanco, E-28049 Madrid, Spain*
[8]*Instituto de Astrofísica de Andalucía (CSIC), Glorieta de la Astronomía, E-18080 Granada, Spain*

7 July 2014

ABSTRACT

We present a method to produce mock galaxy catalogues with efficient perturbation theory schemes, which match the number density, power spectra and bispectra in real and in redshift space from N-body simulations. The essential contribution of this work is the way in which we constrain the bias parameters in the PATCHY-code. In addition of aiming at reproducing the two-point statistics, we seek the set of bias parameters, which constrain the univariate halo probability distribution function (PDF) encoding higher-order correlation functions. We demonstrate that halo catalogues based on the same underlying dark matter field with a fix halo number density, and accurately matching the power spectrum (within 2%), can lead to very different bispectra depending on the adopted halo bias model. A model ignoring the shape of the halo PDF can lead to deviations up to factors of 2. The catalogues obtained additionally constraining the shape of the halo PDF can significantly lower the discrepancy in the three-point statistics, yielding closely unbiased bispectra both in real and in redshift space; which are in general compatible with those corresponding to an N-body simulation within 10% (deviating at most up to 20%). Our calculations show that the constant linear bias of $\sim$2 for Luminous Red Galaxy (LRG) like galaxies seen in the power spectrum, mainly comes from sampling halos in high density peaks, choosing a high density threshold rather than from a factor multiplying the dark matter density field. Our method contributes towards an efficient modelling of the halo/galaxy distribution required to estimate uncertainties in the clustering measurements from galaxy redshift surveys. We have also demonstrated that it represents a powerful tool to test various bias models.

Key words: (cosmology:) large-scale structure of Universe – galaxies: clusters: general – catalogues – galaxies: statistics

1 INTRODUCTION

Mock galaxy catalogues have become an essential tool to assess systematics in the interpretation of galaxy surveys, to test and develop large-scale structure analysis tools, and to understand structure and galaxy formation.

However, the new generation of galaxy surveys with increas-ing volumes and number densities, such as WiggleZ[1] (Drinkwater et al. 2010), VIPERS[2] (Guzzo & The Vipers Team 2013), BOSS[3] (White et al. 2011), DESI[4]/BigBOSS (Schlegel et al. 2011), DES[5]

---

[1] http://wigglez.swin.edu.au/site/
[2] http://vipers.inaf.it/
[3] http://www.sdss3.org/surveys/boss.php
[4] http://desi.lbl.gov/
[5] http://www.darkenergysurvey.org

---

* E-mail: kitaura@aip.de, Karl-Schwarzschild-fellow

(Frieman & Dark Energy Survey Collaboration 2013), LSST [6] (LSST Dark Energy Science Collaboration 2012), J-PAS[7] (Benitez et al. 2014), 4MOST[8] (de Jong et al. 2012) or Euclid[9] (Cimatti et al. 2009; Laureijs 2009), require challenging computational resources to produce the corresponding mock galaxy catalogues necessary to accurately assess the uncertainties in the measurements. Some remarkable attempts producing large N-body simulations can be found in the recent literature (see e.g. Kim et al. 2009; Angulo et al. 2012; Prada et al. 2012; Alimi et al. 2012; Watson et al. 2013). Nevertheless, these simulations do not go beyond the high mass end of haloes modelling luminous red galaxies (LRGs). Moreover, they provide a limited number of realizations, which cannot be used to make a proper study of the cosmic variance because the estimation of covariance matrices requires a large number of mock catalogues (Percival et al. 2014).

Interesting alternatives to produce large numbers of N-body simulations have been recently proposed, such as re-scaling N-body simulations to account for a change in the cosmological parameters (Angulo & White 2010), computing covariance matrices from a small set of simulations (Schneider et al. 2011), or including Lagrangian perturbation theory (LPT) within the Vlasov equations solver to speed up N-body codes (COLA, Tassev, Zaldarriaga & Eisenstein 2013; Koda, Kazin & Blake 2014).

As an alternative to run N-body cosmological simulations, one can calibrate approximate structure formation models to N-body solutions and scan the parameter space using more efficient schemes, as the one presented in this work.

A number of approaches has been proposed in the literature for the generation of mock galaxy catalogues based on LPT, such as PINOCCHIO (Monaco et al. 2002, 2013), PTHALOS (Scoccimarro & Sheth 2002; Manera et al. 2013), and PATCHY (Kitaura, Yepes & Prada 2014). Alternatively, approximate particle mesh based codes: QPM (White, Tinker & McBride 2014) can be applied. Another approach consists of producing low resolution N-body simulations and augment them with a bias model (de la Torre & Peacock 2013; Angulo et al. 2013). We note that the bias model adopted in PATCHY can be applied for such kind of approaches as well. It has been shown that perturbation theory can provide an accurate approach to model the Baryon Acoustic Oscillations (BAOs) (Tassev & Zaldarriaga 2012).

In the era of precision cosmology we need to produce mock catalogues that yield the right matter statistics to high accuracy not only in terms of the number density and correlation function (or power spectrum), but also in terms of the higher-order statistics. The three-point function is essential for an accurate description of the clustering, as it represents a measure of gravitationally induced non-Gaussianities, which characterises the morphology of the cosmic web (Frieman & Gaztanaga 1994) and has a long history applied to galaxy surveys (see e. g. Gaztanaga & Frieman 1994; Scoccimarro et al. 2001a; Verde et al. 2002a; Jing & Börner 2004; Müller & Maulbetsch 2004; Müller, Hoffmann & Nuza 2011). It can be used to test gravity (Shirata et al. 2007; Gil-Marín et al. 2011), to break degeneracies in the galaxy bias (Matarrese, Verde & Heavens 1997; Verde et al. 1998; Scoccimarro et al. 2001b; Verde et al. 2002b; Hoffmann et al. 2014, Gil-Marín et al., in prep.), or to test the existence of primordial non-Gaussianities (Sefusatti &

Komatsu 2007; Jeong & Komatsu 2009). We note that recent efforts have been performed to constrain the dark matter three-point statistics based on an effective field theory description of the large-scale structure (Angulo et al. 2014; Baldauf et al. 2014). An efficient method based on perturbation theory able to produce halo catalogues with the right three-point statistics is still missing.

The distribution of haloes is statistically determined by all its moments. Nevertheless, a method imposing all the corresponding correlation functions (assuming that they are known) to a distribution of haloes is far from trivial and hardly numerically efficient (see Kitaura 2012). Instead, one tries to encode the physics encapturing all the higher-order statistics in the generation of the halo distribution. This is naturally given in a N-body simulation (although the specific halo-finder can lead to some discrepancies, see Knebe et al. 2011). However, when trying to produce thousands of mock halo catalogues on huge volumes, N-body simulations turn out to be computationally very expensive, as they require to solve the interaction between dark matter particles on small scales and high resolutions to resolve the different populations of haloes.

As an alternative one can obtain the large-scale structure cosmic density field (at scales larger than the diameter of the largest haloes, i.e., larger than a few Mpc) with approximate gravity solvers (or from low resolution N-body simulations) and model its relation to the halo distribution with a parametrised statistical bias description to produce the full halo catalogue or augment the missing populations of haloes. The advantage of such an approach is that it is much faster and less memory consuming, as one does not require high resolutions since the individual haloes do not need to be resolved. Moreover, one can get insights into the halo bias and find accurate descriptions, useful for the analysis of the large-scale structure.

This paper is structured as follows: in the next section (§2) we present our method. We then show (§3) our numerical experiments calibrating our mock catalogues with N-body simulations. Finally (§4) we present our conclusions and discussion.

2 METHOD

Let us start defining the generation of a halo distribution as a statistical problem.

2.1 Statistical problem

The distribution of haloes is statistically determined by its mean number density (ξ_1^h: one-point correlation function), (ξ_2^h: two-point) correlation function, skewness (ξ_3^h: three-point correlation function), kurtosis (ξ_4^h: four-point correlation function), and all the rest of higher-order correlation functions:

$$N_h \curvearrowright \mathcal{P}(N_h | \xi_1^h, \xi_2^h, \xi_3^h, \xi_4^h, \dots), \qquad (1)$$

where N_h are the number counts of haloes per cosmic sub-volume (per cells when dividing the cosmic volume in a grid), ξ_i^h the different correlation functions corresponding to the halo distribution and $\mathcal{P}$ the corresponding probability distribution function (PDF). The halo population is defined in a particular mass range ($N_h = N_h[\Delta M_h]$) or alternatively in a maximum circular velocity range ($N_h = N_h[\Delta v_{\max}]$) and consequently all correlation functions also depend on that range ($\xi_i^h = \xi_i^h[\Delta M_h]$ or $\xi_i^h = \xi_i^h[\Delta v_{\max}]$, for $i = 1, \dots, m$ until the highest considered order m).

Let us suppose that the expected continuous halo density field

$\rho_h = \langle N_h \rangle$ and the PDF relating it with the discrete halo number count per cell are known. Then we could get the discrete halo distribution by sampling from this PDF:

$$N_h \curvearrowleft \mathcal{P}(N_h | \rho_h). \qquad (2)$$

This PDF deviates in general from Poissonity, due to the non-vanishing correlation of haloes producing over- or under-dispersed distributions depending on the halo population and density regime (see Somerville et al. 2001; Casas-Miranda et al. 2002; Neyrinck et al. 2014). Under-dispersion is a sub-dominant effect, which appears when the correlation function becomes negative (see also Baldauf et al. 2012, 2013). Therefore we focus on modelling over-dispersion with the negative binomial distribution function, which requires a single stochastic bias parameter β (see Kitaura, Yepes & Prada 2014; Neyrinck et al. 2014). Stochasticity in the bias relation has since long been studied (see e. g. Press & Schechter 1974; Peacock & Heavens 1985; Bardeen et al. 1986; Fry & Gaztanaga 1993; Mo & White 1996; Dekel & Lahav 1999; Sheth & Lemson 1999; Seljak 2000; Berlind & Weinberg 2002; Smith, Scoccimarro & Sheth 2007; Desjacques et al. 2010; Beltrán Jiménez & Durrer 2011; Valageas & Nishimichi 2011; Elia, Ludlow & Porciani 2012; Chan, Scoccimarro & Sheth 2012; Baldauf et al. 2012, 2013) and has a non-trivial relation to the non-local bias (see discussion at the end of next section).

To obtain the expected halo density field one needs to control the bias of the particular halo population, which in general is a non-linear, scale-dependent, and non-local function of the dark matter density field ρ_M: $B(\rho_h | \rho_M)$.

From a statistical perspective, this bias relation is not only a function of the underlying dark matter field, but also of all the moments of the halo distribution: $B(\rho_h | \rho_M) = B(\rho_h | \rho_M, \xi_1^h, \xi_2^h, \xi_3^h, \xi_4^h, \dots)$. Hence, the haloes are sampled from the following PDF:

$$N_h \curvearrowleft \mathcal{P}(N_h | B(\rho_h | \rho_M, \xi_1^h, \xi_2^h, \xi_3^h, \xi_4^h, \dots)). \qquad (3)$$

There is no apparent advantage in the last expression with respect to Eq. 1, as we still need all the moments of the halo distribution and in addition the dark matter field together with the halo bias. However, there is a striking difference, since we have extended the purely statistical problem with a physical model. The physics is encoded in the large-scale structure density field dominated by dark matter and its biased relation to the halo population. While the relation in Eq. 1 does not tell us how to obtain the halo distribution (for some attempts based on expansions of Gaussian/Lognormal fields including higher-order statistics see Kitaura 2012), Eq. 3 gives us a straightforward approach. In principle we just need to define a parametrised bias model and constrain the corresponding parameters with the halo correlation functions extracted from N-body simulations.

2.2 Deterministic bias relations

Let us define the parametrised deterministic bias relation.

(i) One could follow Fry & Gaztanaga (1993) and expand the dark matter overdensity field $\delta_M \equiv \rho_M / \bar{\rho}_M - 1$ (with $\bar{\rho}_M$ being the mean dark matter density) in a Taylor series:

$$\rho_h = f_h^a \sum_i a_i \delta_M^i, \qquad (4)$$

with $f_h^a = \bar{N}_h / \langle \sum_i a_i \delta_M^i \rangle$ and a_i being the corresponding bias coefficients. The ensemble average over the quantity X: $\langle X \rangle$ can

be approximated by the following average: $\sum_i X_i / N_{\text{cells}}$ for sufficiently large volumes (with N_{cells} being the number of cells dividing the entire volume). This bias model has been proposed to augment the resolution of an N-body simulation, populating lower mass haloes than originally resolved in the simulation (see Angulo et al. 2013).

(ii) Alternatively one could follow Cen & Ostriker (1993) and expand the series based on the logarithm of the density field (avoiding in this way negative densities allowed in the previous expansion):

$$\rho_h = f_h^b \exp\left[\sum_i b_i \log\left(1 + \delta_M\right)^i \right], \qquad (5)$$

with $f_h^b = \bar{N}_h / \langle \exp\left[\sum_i b_i \log\left(1 + \delta_M\right)^i \right] \rangle$ and b_i being the corresponding bias coefficients. This model has been used for resolution augmentation of N-body simulations (see de la Torre & Peacock 2013), further augmenting the catalogue with the halo occupation distribution (HOD) (Berlind & Weinberg 2002; Zheng, Coil & Zehavi 2007; de la Torre et al. 2013).

(iii) It has recently been found that the bias is very well fit by a compact relation including an exponential cut-off: $\rho_h \propto \rho_M^\alpha \exp\left[- \left(\frac{\rho_M}{\rho_\epsilon} \right)^\epsilon \right]$ (see Neyrinck et al. 2014). This is a refined version of the thresholding step-function $\theta(\rho_M - \rho_{\text{th}})$ ($= 0$ for $\rho_M < \rho_{\text{th}}$, $= 1$ for $\rho_M \geqslant \rho_{\text{th}}$), which suppresses the generation of haloes below a density threshold ρ_{th} and was previously used in (Kitaura, Yepes & Prada 2014). Such a model is in agreement with the Press & Schechter (1974) and peak-background split picture (Kaiser 1984; Bardeen et al. 1986; Cole & Kaiser 1989; Mo, Jing & White 1996; Sheth, Mo & Tormen 2001). One can combine both descriptions in a single model:

$$\rho_h = f_h \, \theta(\rho_M - \rho_{\text{th}}) \, \rho_M^\alpha \, \exp\left[- \left(\frac{\rho_M}{\rho_\epsilon} \right)^\epsilon \right], \qquad (6)$$

with

$$f_h = \bar{N}_h / \langle \theta(\rho_M - \rho_{\text{th}}) \, \rho_M^\alpha \, \exp\left[- \left(\frac{\rho_M}{\rho_\epsilon} \right)^\epsilon \right] \rangle, \qquad (7)$$

and $\{\rho_{\text{th}}, \alpha, \epsilon, \rho_\epsilon\}$ the parameters of the model. Nevertheless, the exponential cut-off ($\epsilon < 0$) is not very sensible for the massive halo/galaxy population considered in this work. Hereby, the bias is mainly controled by α and ρ_{th}. Different combinations of these parameters can lead to the same power spectrum (two-point statistics). Therefore, we need to introduce additional constraints, as we will discuss below.

Additional bias is introduced by the approximate gravity solver and aliasing effects due to the gridding of the dark matter particles, when defining the dark matter density field ρ_M. We rely in this work on the (iii) bias expression and include additional bias effects in these parameters.

A non-local bias introduces a scatter in the local deterministic bias relations described above (McDonald & Roy 2009). These effects have recently been found to be relevant (Saito et al. 2014), which is in line with the findings of Kitaura, Yepes & Prada (2014). In the latter work, the scatter is described within a stochastic bias relation through the negative binomial distribution function, as we do here. Therefore, stochastic bias may be regarded as an effective description of the non-local bias encoding our ignorance about it. Nevertheless, such effects need to be addressed in more detail in future works, as a stochastic description may lack some physical effects in the halo distribution.

2.3 Halo probability distribution function

The approach described so far to generate the halo distribution based on Eq. 3 is not very efficient. One would need to scan the bias parameter space computing the correlation functions (say up to fourth order) for each set of values until the optimal set is found. A fast method has recently been developed to compute the bisepctrum in the squeezed-limit (Chiang et al. 2014), which however, does not apply for this work. To circumvent these computational problems, we suggest to constrain the one-dimensional halo PDF. In particular we will constrain the shape of the halo PDF, as its integral (the halo number density) will be imposed by f_h (see Eq. 7).

Let us recap some concepts of statistical probability theory (for a review see Bernardeau et al. 2002).

2.3.1 Statistical probability theory

The higher-order moments of a field ρ can be found by calculating the ensemble average of powers of that field ρ over its probability distribution function $\mathcal{P}^1(\rho)$:

$$\mu_n \equiv \int \mathrm{d}\rho \, \mathcal{P}^1(\rho) \, \rho^n = \langle \rho^n \rangle, \tag{8}$$

with n being the order of the moment.

One can then define the moment generating function as the quantity:

$$\mathcal{M}(t) \equiv \sum_{n=0}^{\infty} \mu_n \frac{t^n}{n!} = \int \mathrm{d}\rho \, \mathcal{P}^1(\rho) e^{t\rho} = \langle e^{t\rho} \rangle, \tag{9}$$

which yields the different moments by performing subsequent derivatives of it at the origin $t = 0$.

The cosmological n-point correlation functions are related to the cumulants or connected moments:

$$\kappa_n \equiv \langle \rho^n \rangle_c, \tag{10}$$

defined by the cumulant generating function:

$$\mathcal{C}(t) \equiv \sum_{n=1}^{\infty} \kappa_n \frac{t^n}{n!}, \tag{11}$$

and its relation to the moment generating function:

$$\mathcal{M}(t) = \exp(\mathcal{C}(t)). \tag{12}$$

By inverting Eq. 9 using the inverse Laplace transform, one obtains the PDF as a function of the cumulants:

$$\mathcal{P}^1(\rho) = \int_{-\sqrt{-1}\infty}^{\sqrt{-1}\infty} \frac{\mathrm{d}t}{2\pi\sqrt{-1}} \exp\left(t\rho + \mathcal{C}(t)\right). \tag{13}$$

This equation illustrates how the n-point correlation functions are encoded in the probability distribution function.

2.3.2 Multivariate constraints

We have considered so far the univariate case. However, this alone does not ensure the correct higher-order statistics, as any distribution (for instance a Gaussian field) can be rank-ordered to fulfill a particular PDF (see Weinberg 1992; Sigad, Branchini & Dekel 2000; Szapudi & Pan 2004). The cosmic density field represents a multivariate statistical problem. In practice the statistical dimension is given by the number of cells (voxels) conforming the three dimensional grid we use to model the whole cosmic volume under consideration (for the more general multivariate expressions see Bernardeau et al. 2002; Kitaura 2012).

One needs to constrain the PDF from the multivariate halo distribution. Following Eq. 3 we need to assume that at scales larger than few Mpc (i.e., larger than the diameter of the largest haloes):

• The approximate gravity solver (low N-body resolution or perturbation theory based method) accurately models the higher-order statistics of the dark matter density field.

• The bias model accurately connects the dark matter phase-space distribution with the halo distribution.

In this way, we scan the parameter space spanned by the bias parameters, in our case: $\{\rho_{\mathrm{th}}, \alpha, \epsilon, \rho_\epsilon, \beta\}$, additionaly fixing the number density of the halo population in which we are interested: $\bar{N}_h = \langle \rho_h \rangle \leftarrow \xi_1^h$, to match the power spectrum: $P_h(k) \leftarrow \xi_2^h$ and the halo PDF: $\mathcal{P}_h^1\left(B\left(\rho_h | \rho_{\mathrm{M}}\right)\right) \leftarrow \{\xi_1^h, \xi_2^h, \xi_3^h, \xi_4^h, \dots\}$ obtained from N-body simulations. The computation of the halo PDF is trivial and fast in contrast to the calculation of higher-order correlation functions. This can be summarised with the following sampling process:

$$N_h \curvearrowleft \mathcal{P}(N_h | B(\rho_h | \rho_{\mathrm{M}}), \bar{N}_h, P_h(k), \mathcal{P}_h^1)). \tag{14}$$

The accuracy of our method depends on the level of precision in which each of the two above mentioned conditions are fulfilled. Recently, rank-ordering has been applied to dark matter density fields within second order LPT (2LPT) demonstrating that this approximation encaptures the right matter statistics on scales larger than $8 \, h^{-1}$ Mpc, based on calculations of the bispectrum (see Leclercq et al. 2013). LPT was also shown to accurately model the three-point correlation function on scales relevant to BAOs in configuration space (White, Tinker & McBride 2014). The limitations of LPT can be improved with the spherical collapse model (see Neyrinck 2013; Kitaura & Heß 2013, and the description of the PATCHY-code below). We stress that the advantage of the above probabilistic formulation goes beyond of producing mock catalogues, as it permits us also for statistical inference analysis (see e. g. Kitaura & Enßlin 2008; Kitaura, Jasche & Metcalf 2010; Jasche & Kitaura 2010; Kitaura, Gallerani & Ferrara 2012; Kitaura 2013).

3 NUMERICAL EXPERIMENTS

We present in this section numerical experiments to validate our method.

3.1 Reference halo sample: the BigMutiDark simulation

To calibrate our method, we use a reference halo catalogue at redshift $z = 0.577$ extracted from one of the BigMultiDark (BIGMD) simulations (Heß et al., in prep.; Klypin et al. 2013), which was performed using GADGET-2 (Springel 2005) with 3840^3 particles on a volume of $(2500 \, h^{-1} \, \mathrm{Mpc})^3$ assuming ΛCDM-cosmology with $\{\Omega_{\mathrm{M}} = 0.29, \Omega_{\mathrm{K}} = 0, \Omega_\Lambda = 0.71, \Omega_{\mathrm{B}} = 0.047, \sigma_8 = 0.82, w = -1, n_s = 0.95\}$ and a Hubble constant ($H_0 = 100 \, h$ km s^{-1} Mpc^{-1}) given by $h = 0.7$. Haloes were defined based on density peaks including substructures using the Bound Density Maximum (BDM) halo finder (Klypin & Holtzman 1997) and then selected according to a maximum circular velocity larger than 350 km s^{-1} to match the number density of BOSS CMASS galaxies (Nuza et al. 2013; Klypin et al. 2013). For the impact of these selection criteria in the clustering and scale-dependent bias (Prada et al., in prep.).
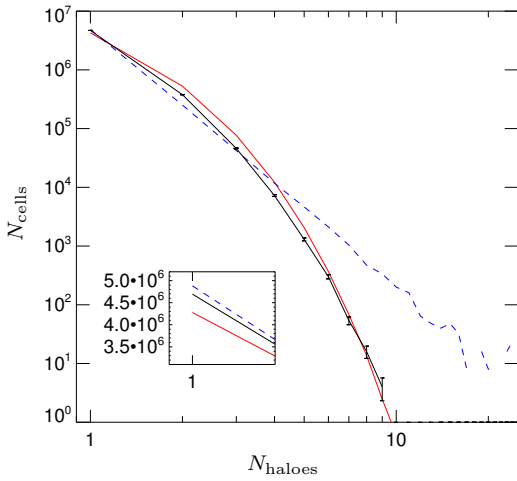
Figure 1. Halo PDF showing the number of cells N_{cells} hosting the same number of haloes N_{haloes} with a cell resolution of $(2.5 \, h^{-1} \, \text{Mpc})^3$. In red PATCHY (with $\rho_{\text{th}}^{\text{High}}$) and in black BIGMD including the error-bars extracted from the PATCHY mocks. In addition one PATCHY mock in dashed-blue (with $\rho_{\text{th}}^{\text{Low}}$) ignoring the shape of the halo PDF. The latter has been multiplied by eight to compensate for the eight times smaller volume, and hence, about eight times smaller number of haloes. The number density (which determines the integral of the PDF) is the same in all catalogues (see Eq. 7). The insert shows the first bin of the halo PDF for cells containing only one halo.

3.2 The PATCHY-code

To maximise the efficiency of the method we rely on augmented Lagrangian perturbation theory (ALPT) splitting the description of the large-scale structure into a long range and a short range component modelled by 2LPT and the spherical collapse model, respectively (see Kitaura & Heß 2013). We note, however, that the method described above can be applied to increase the resolution of N-body simulations, and thus, it can be regarded as an improved method with an extended parametrised bias model with respect to other approaches like the ones presented in de la Torre & Peacock (2013); Angulo et al. (2013). It has been shown with the PATCHY-code that the non-linear bias model also adopted in this work can compensate for the missing power of the perturbative approach and redshift space distortions can be accurately modelled with ALPT and a dispersion term (see Kitaura, Yepes & Prada 2014). We have extended PATCHY to be out-of-core and work with an arbitrary number of chunks to reduce the memory requirements below 28 Gb RAM for 1024^3 particles (or cells), allowing to simulate the distribution of LRG-like galaxies in volumes of $(2.5 \, h^{-1} \, \text{Gpc})^3$. The new version of PATCHY randomly assigns the positions of dark matter particles to haloes, whenever there are more dark matter particles than haloes in a given cell. Otherwise the position within the cell is uniform randomly assigned. This case does, however, occur only for a small fraction of cells considering the particular halo population of this work.

3.3 Bias parameters and statistical constraints

The parameters of our model $\{\rho_{\text{th}}, \alpha, \epsilon, \rho_\epsilon, \beta\}$ are constrained to fit the power spectrum better than 2% in the range $k = [0.07, 0.35] \, h \, \text{Mpc}^{-1}$ (cosmic variance dominates on larger scales: $k \lesssim 0.07 \, h \, \text{Mpc}^{-1}$) and to accurately reproduce the tails of the halo PDF, i.e., matching the low-end better than 10% and being statisti-

cally compatible within 1-sigma at the high-end (around the maximum number of haloes per cell). We find that in this way we obtain also a reasonable fit of the halo PDF in the whole range, with deviations being in general within 50%. Our bispectrum statistics analysis has shown that these deviations are not severe (see below). To achieve better fits we would need to match the low-end of the halo PDF to higher accuracy, as small deviations in the number of cells containing one halo can lead to larger deviations in cells containing higher number of haloes (see insert in Fig. 1). There should also be limitations due to the assumptions adopted in §2.3.2. This should be further investigated in future works. Nevertheless, a massive parameter estimation approach to find the optimal parameters to match the BOSS-CMASS galaxy population will be presented in a forthcoming paper (Scoccola et al., in prep.). We note that the right number density is imposed through Eq. 7.

A multiscale approach has been adopted in this work to efficiently scan the bias parameter space. We start with small volumes of $(312.5 \, h^{-1} \, \text{Mpc})^3$ on grids with 128^3 cells and refine the parameters increasing the volumes by factors of eight with constant resolution, until we reach the final volume of $(2500 \, h^{-1} \, \text{Mpc})^3$ with 1024^3 cells. The majority of the tests were performed on volumes of $(1250 \, h^{-1} \, \text{Mpc})^3$ and grids of 512^3, since the statistical uncertainties are already low at these volumes and they can be computed much faster than the full volume of the BIGMD simulation. We show in this study one of such mocks with the following parameters: $\alpha = 1.68$ and $\delta_{th} = \rho_{\text{th}}/\bar{\rho} - 1 = -0.3$, which ignores the shape of the halo PDF from the BIGMD simulation and solely fits the halo number density (integral of the halo PDF) and the power spectra in real and redshift space (see § 3.4). This setting permits the existence of haloes in low density regions, we therefore will refer to this sample as the PATCHY mock with $\rho_{\text{th}}^{\text{Low}}$. The calculations based on this mock are represented by dashed-blue curves throughout the paper. In addition, we show results from a set of 20 PATCHY mockswith parameters chosen to jointly fit the power spectrum and the halo PDF ($\rho_{\text{th}}^{\text{High}}$), using the same cosmology, number density and volume as the BIGMD simulation described in § 3.1. A $\sim$53 times lower number of dark matter particles (1024^3) was used for the PATCHY mocks. We find that a higher thresholding is crucial to fit the PDF (see Fig. 1). The following parameters are found: $\alpha = 0.365$ and $\delta_{th} = \rho_{\text{th}}/\bar{\rho} - 1 = 1.82$, and therefore we will refer to this sample as the PATCHY mocks with $\rho_{\text{th}}^{\text{High}}$. The calculations based on these mocks are represented by red curves throughout the paper. We find for the parameter modelling the deviation from Poissonity $\beta = 0.35$. In this way we manage to enhance the power towards small scales (high k) fitting the power spectrum of the BIGMD simulation (see Fig. 3). The exponential cut-off parameters are chosen as: $\epsilon = -2$, $\rho_\epsilon = 0.3$. We note, however, that these parameters have a limited impact on the statistics given the particular halo population we are looking at. We expect that they will become more relevant when looking at lower halo masses.

A visual comparison between the halo field from the BIGMD simulation and from the PATCHY mocks is shown in Fig. 2. There are no obvious deviations between the BIGMD simulation and the PATCHY mocks other than a different spatial distribution of haloes due to the different initial random seed perturbations used for each case. A more careful visual inspection reveals that the PATCHY mock with $\rho_{\text{th}}^{\text{Low}}$ exhibits a larger number of cells with low halo number counts filling the voids with respect to the PATCHY mock with $\rho_{\text{th}}^{\text{High}}$ and the BIGMD simulation. This is expected as the density threshold is set to a lower value. We cannot distinguished by eye the larger saturation of the peaks. This is however shown in the PDF, where one can see that the PATCHY mock with $\rho_{\text{th}}^{\text{Low}}$ reaches a
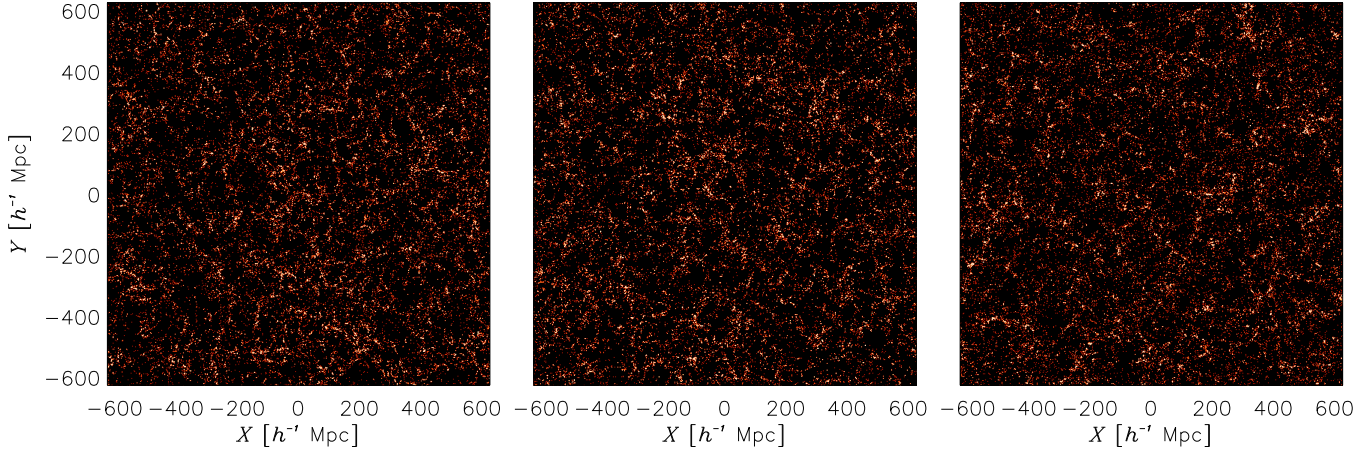
Figure 2. Halo overdensity slices of thickness $20\,h^{-1}$ Mpc taking a subvolume of $(1250\,h^{-1}\,\mathrm{Mpc})^3$ from the BIGMD N-body simulation on the left, from one PATCHY mock including a fit of the halo PDF (with $\rho_{\mathrm{th}}^{\mathrm{High}}$) in the middle, and from one PATCHY mock ignoring the shape of the halo PDF (with $\rho_{\mathrm{th}}^{\mathrm{Low}}$) on the right. Lighter regions indicate a larger halo number count.

factor ~ 3 times larger number of haloes per cell than the PATCHY mock with $\rho_{\mathrm{th}}^{\mathrm{High}}$ or the BIGMD simulation (see Fig. 1). We proceed with a more quantitative statistical analysis.

3.4   Power spectrum and bispectrum

In this section we aim to compare the halo power spectrum and halo bispectrum predictions for PATCHY mocks and BIGMD N-body simulation in real and redshift space.

The halo power spectrum P and the halo bispectrum B are the two- and three-point correlation functions in Fourier space. Given the Fourier transform of the halo overdensity, $\delta_h(\mathbf{x}) \equiv \rho_h(\mathbf{x})/\bar{\rho}_h - 1$,

$$\delta_h(\mathbf{k}) = \int d^3\mathbf{x}\,\delta_h(\mathbf{x})\exp(-i\mathbf{k}\cdot\mathbf{x}), \qquad (15)$$

where $\rho_h(\mathbf{x})$ is the number density of objects and $\bar{\rho}_h$ its mean value, the halo power spectrum and halo bispectrum are defined as,

$$\langle\delta_h(\mathbf{k})\delta_h(\mathbf{k}')\rangle \equiv (2\pi)^3 P(k)\delta^D(\mathbf{k}+\mathbf{k}'), \qquad (16)$$

$$\langle\delta_h(\mathbf{k}_1)\delta_h(\mathbf{k}_2)\delta_h(\mathbf{k}_3)\rangle \equiv (2\pi)^3 B(\mathbf{k}_1,\mathbf{k}_2)\delta^D(\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3), \qquad (17)$$

with δ^D being the Dirac delta function. Note that the bispectrum is only well defined when the set of k-vectors, k_1, k_2 and k_3 close to form a triangle, $\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3 = \mathbf{0}$. It is common to define the reduced bispectrum Q as,

$$Q(\alpha_{12}|\mathbf{k}_1,\mathbf{k}_2) \equiv \frac{B(\mathbf{k}_1,\mathbf{k}_2)}{P(k_1)P(k_2) + P(k_2)P(k_3) + P(k_1)P(k_3)}. \qquad (18)$$

where α_{12} is the angle between $\mathbf{k}_1$ and $\mathbf{k}_2$. Although this quantity does not have any additional information to the power spectrum and bispectrum, it has been historically used as a measurement of the hierarchical amplitude of the bispectrum. This quantity is independent of the overall scale k and redshift at large scales and for a power spectrum that follows a power law. Moreover, it presents a characteristic "U-shape" predicted by the gravitational instability.

Mode coupling and power law deviations in the actual power spectrum induce a slight scale- and time-dependency in this quantity. However, in practice it has been observed that at scales of order $k \sim 0.1\,h\,\mathrm{Mpc}^{-1}$ the reduced bispectrum does not present a high variation in its amplitude.

In order to measure the Fourier statistics from a set of haloes in a box with periodic boundary conditions, we start by discretising the box in grid cells. We use 512^3 grid cells, with a grid-cell resolution of $4.88\,h^{-1}\,\mathrm{Mpc}$, and we assign haloes to the grid cells according to the Cloud-in-Cell (CiC) prescription. We apply the Fourier transform of the density field using fftw[10] and we correct the resulting field by deconvolving the effect of the grid assignment (Jing 2005). Given that the size of the box is $L_b = 2.5\,h^{-1}\,\mathrm{Gpc}$, the k fundamental is $k_f = 2\pi/L_b = 2.51 \times 10^{-3}\,h\,\mathrm{Mpc}^{-1}$.

In order to obtain the power spectrum we radially bin the k-modes, with a bin size of k_f, and we average over all possible k-directions.

We use the real part of $\langle\delta_{\mathbf{k}_1}\delta_{\mathbf{k}_2}\delta_{\mathbf{k}_3}\rangle$ as our bispectrum measurement for those set of $\mathbf{k}_i$-vectors that form a triangle. Since there are a huge number of possible triangular configurations, in this paper we focus on analysing those with $k_1/k_2 = 1$ and $k_1/k_2 = 2$, for a particular values of k_1. We present the bispectrum measurement as a function of k_3, which sweeps from $|\mathbf{k}_1 - \mathbf{k}_2|$ to $|\mathbf{k}_1 + \mathbf{k}_2|$. Our choice for binning the bispectrum is similar to the power spectrum. We bin k_1 and k_3 in k_f bins: $\Delta k_1 = \Delta k_3 = k_f$. In addition, k_2 is binned also in k_f bins for those triangles with $k_1/k_2 = 1$. On the other hand, k_2 is binned in 2 times k_f when $k_2/k_1 = 2$ in order to cover all the available k-space. Generically, we can write $\Delta k_2 = (k_2/k_1)\Delta k_1$.

The measurement of the bispectrum is performed in a similar way of the approach described in Gil-Marín et al. (2012). This method consists of generating k-triangles and randomly-orientate them in k-space. When the number of random triangles is sufficiently large, the mean value of their bispectra tends to the fiducial bispectrum (for details see Gil-Marín et al. 2012).

Discreteness adds a spurious contribution to the measured

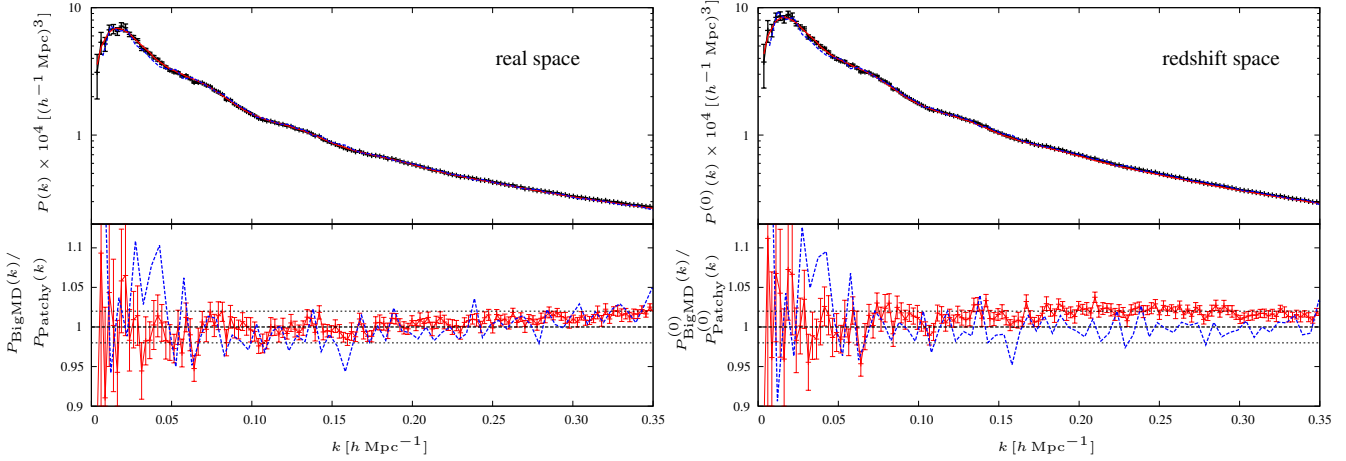
[10]   Fastest Fourier Transform of the West: www.fftw.org

Figure 3. Halo power spectrum for BIGMD N-body simulation (black symbols in the top sub-panels) and for PATCHY mocks (red lines with $\rho_{\rm th}^{\rm High}$ and dashed-blue lines with $\rho_{\rm th}^{\rm Low}$). Left panel shows real space power spectrum and right panel the redshift space monopole power spectrum. The power spectrum for the PATCHY mocks with $\rho_{\rm th}^{\rm High}$ corresponds to the average over 20 independent realizations. Only the errors corresponding to the BIGMD N-body simulations are shown for clarity. In the bottom sub-panels the relative deviation between BIGMD and PATCHY is shown. Black dotted lines indicate 2% deviation.
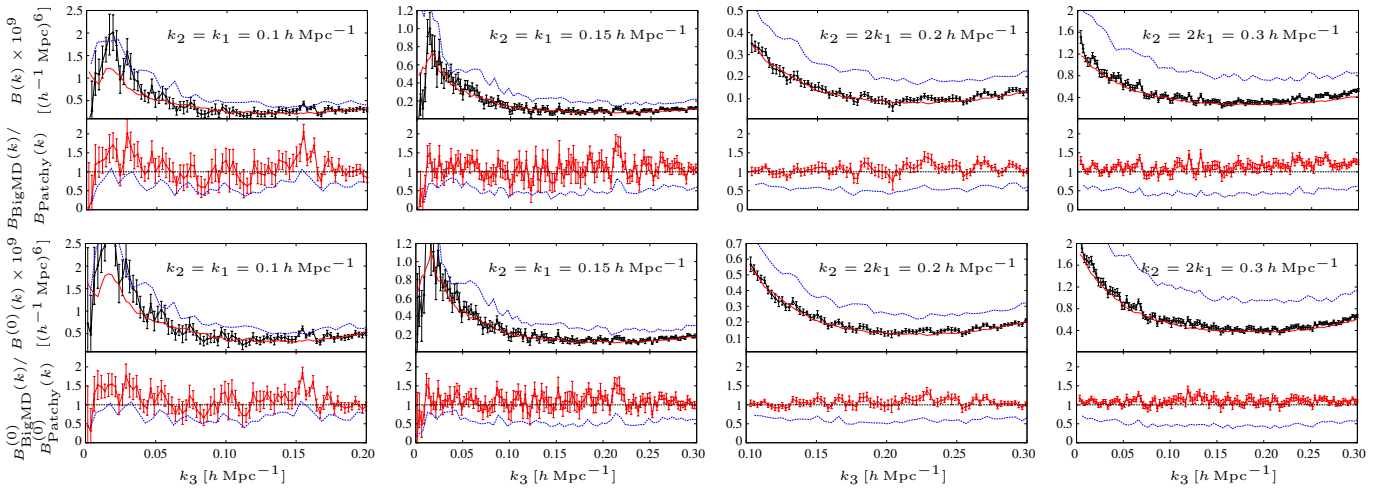
Figure 4. Halo bispectrum for the BIGMD N-body simulations and PATCHY mocks as a function of k_3. Top panels show the bispectrum in real space and bottom panels the bispectrum monopole in redshift space. Different columns present different scales and shapes as labeled. Same colour notation that in Fig. 3 is used.

power spectrum and bispectrum. In this paper we assume that these contributions are of Poisson type and therefore are given by,

$$P_{\rm sn}(k) = \frac{1}{\bar{n}} \qquad (19)$$

$$B_{\rm sn}(\mathbf{k}_1, \mathbf{k}_2) = \frac{1}{\bar{n}}\left[P(k_1) + P(k_2) + P(k_3)\right] + \frac{1}{\bar{n}^2} \qquad (20)$$

where $k_3 = |\mathbf{k}_1 + \mathbf{k}_2|$ and $\bar{n}$ is the number density of haloes. We are aware that deviations from Poissonity are present (see §2.1). However, for the purposes of this paper, we find it sufficient to use the Poisson predictions to correct the power spectrum and bispectrum measurements in a consistent way for both the N-body and the PATCHY simulations.

Both for the power spectrum and bispectrum the errors associated to the measurement come from the dispersion of 20 independent realizations of the PATCHY mocks.

Recall that the BIGMD simulation corresponds to a single box of $2.5\,h^{-1}$Gpc on a side, and hence, has an effective volume of $V_{\rm eff} = 15.625\,[h^{-1}{\rm Gpc}]^3$. On the other hand, for the PATCHY mocks simulations we dispose of 20 boxes of $2.5\,h^{-1}$Gpc on a side with independent initial conditions and with a total effective volume of $V_{\rm eff} = 312.5\,[h^{-1}{\rm Gpc}]^3$. Due to this difference in effective volumes, BIGMD measurements (for both power spectrum and bispectrum) present a more noisy behaviour than PATCHY measurements, with error-bars that are $\sim \sqrt{20}$ times larger for BIGMD compared to PATCHY. This difference is more evident for the bispectrum because the signal-to-noise is weaker, and therefore, the errors larger than for the power spectrum.

For both power spectrum and bispectrum we present the BIGMD error-bars computed from the dispersion among the 20 realizations of PATCHY mocks. We do not display any error-bars
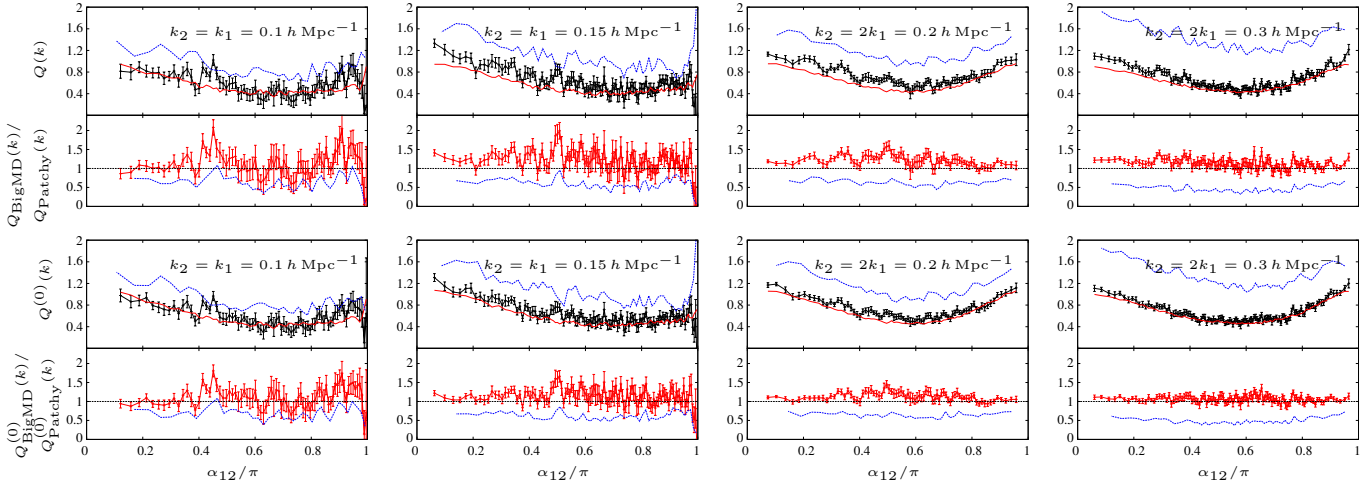
Figure 5. Halo reduced bispectrum for BIGMD N-body simulations and PATCHY mocks as a function of the angle between $\mathbf{k}_1$ and $\mathbf{k}_2$, α_{12}. Same colour and panel notation that in Fig. 4 is used.

for the PATCHY mocks measurements for clarity. These error-bars would be $\sqrt{20}$ times smaller than the ones showed for BIGMD.

Fig. 3 presents the comparison between the power spectrum of BIGMD N-body simulations (black symbols), the PATCHY mocks with ρ^{High} (red lines), and the single PATCHY mock with ρ^{Low} (dashed-blue line). The left panel displays the real space power spectrum and the right panel the redshift space power spectrum monopole. Bottom sub-panels present the relative deviation between them. As a general trend, we see a good agreement between BIGMD and PATCHY power-spectra, which agree within $\leqslant 2\%$ accuracy for $k \leqslant 0.35\,h\,\text{Mpc}^{-1}$, both in real and redshift space. We observe that at small scales, the PATCHY mocks considered in this study tend to slightly under-predict the power spectrum by $\sim 2\%$ respect to BIGMD, being compatible within better than 10% up to $k = 1\,h\,\text{Mpc}^{-1}$. Further tests have, however, shown that a different set of parameters can also over-predict the power spectrum at large scales, indicating that some set will display a better fit on small scales. We do not intend in this study to reach precisions of 2% beyond the scale relevant to BAOs ($k = 0.35\,h\,\text{Mpc}^{-1}$) and therefore have not further improved the parameters. The PATCHY mocks with $\rho^{\text{High}}_{\text{th}}$ slightly over-predict the power spectrum towards high k, however remaining compatible within 2% with the BIGMD simulation. The non-linear contribution of redshift space distortions responsible of damping the power spectrum on small scales can be controlled with the factor of the dispersion term as shown in Kitaura, Yepes & Prada (2014). We plan to investigate this further including calculations of the quadrupole in a future work.

Fig. 4 presents the comparison between BIGMD and PATCHY bispectrum using the same colour notation that in Fig. 3. Top panels correspond to real space bispectrum and bottom panels to redshift space bispectrum monopole. Different columns correspond to different scales and shapes as indicated. Note that in this case, the precision of BIGMD measurements is a limiting factor when we test the accuracy of PATCHY respect to BIGMD bispectra. As a general trend we see that both PATCHY and BIGMD agree within $10 - 20\%$ accuracy. Similarly to the power spectrum case, we see that PATCHY tends to under-predict the bispectrum at small scales by $\sim 10 - 20\%$ with no evidence of any shape dependence.

Fig. 5 presents a similar comparison for the reduced bispectra using the same colour and panel notation that in Fig. 3 and

4. Similarly to the bispectrum case, we observe a general agreement between BIGMD and PATCHY with ρ^{High} reduced bispectra within $10 - 20\%$ accuracy. However, we see that the PATCHY with ρ^{High} prediction for the reduced bispectra tends to under-estimate the BIGMD prediction by $\sim 20\%$. This deviation tends to be more evident in real space that in redshift space. However, we should note the large uncertainties in our single N-body simulation shown in the fluctuations as a function of the angle α_{12}.

We have tested that a different set of bias parameters matching the power spectrum, but disregarding the shape of the halo PDF produces bispectra, which can deviate from the true one by about a factor of 2 as can be seen represented by the dashed-blue curves corresponding to PATCHY with ρ^{Low} in Figs. 3, 4 and 5.

In summary, we consider that the bispectra of the PATCHY-mocks fit well the ones from BIGMD given the uncertainties of our single reference N-body simulation. Further investigation should be done in the future with a larger number of reference N-body simulations.

4 DISCUSSION AND CONCLUSIONS

In this work we have presented a method to produce mock galaxy catalogues with efficient perturbation theory schemes, which match the number density, power spectra and bispectra in real and in redshift space from N-body simulations. The ingredients of our scheme are given by an improved Lagrangian perturbation theory based approach to describe dark matter structure formation and a non-linear stochastic bias model (the PATCHY-code: Kitaura, Yepes & Prada 2014). The essential contribution of this work is the way in which we constrain the parameters of our bias model. In addition of aiming at reproducing the two-point statistics we need to additionally constrain the univariate halo probability distribution function encoding higher-order correlation functions. We have shown that this approach permits us to reasonably reproduce the bispectrum. Nevertheless, we have not included any explicit constraint from the three-point statistics. We therefore expect that our general approach yields also reasonable fits of the four-point correlations, which are relevant, as they quantify the sample variance and covariance of two-point statistics measurements (Cooray & Hu 2001; Takada &

Hu 2013). We leave a thourough analysis of covariance matrices for a forthcoming publication.

We have used a Luminous Red Galaxy (LRG) like reference halo catalogue with the typical BOSS CMASS galaxy number density at $z = 0.577$ based on one of the BIGMD N-body simulations. We have demonstrated that halo catalogues based on the same underlying dark matter field with a fix halo number density (i.e., matching the integral of the halo PDF) and accurately matching the power spectrum (within 2% for $k \leqslant 0.35\,h\,\mathrm{Mpc}^{-1}$ and deviating less than 10% up to $k = 1\,h\,\mathrm{Mpc}^{-1}$), can lead to very different bispectra depending on the halo bias model. A model ignoring the shape of the halo PDF can lead to deviations up to the level of a factor of 2. The catalogues obtained additionally constraining the halo PDF can significantly lower the discrepancy in the three-point statistics yielding closely unbiased bispectra both in real and in redshift space, which are compatible with those corresponding to an N-body simulation in general within 10% (deviating at most up to 20%).

Our calculations have shown that the constant linear bias of ~ 2 for LRG-like galaxies found in the power spectrum (defined as the square root of the halo power spectrum divided by the dark matter power spectrum), mainly comes from sampling halos in the high density peaks choosing a high density threshold rather than from a factor multiplying the dark matter density field.

The method presented in this work can be applied to directly produce galaxy catalogues without requiring the halo distribution, as it just cares about the statistical properties for the type of tracers of interest, and is in this way a model-independent method. The model dependency comes in, when the method is calibrated with a particular mock galaxy catalogue. This dependence could be broken, by using as the reference an observed sample of galaxies. We could then add the large-scale modes and produce different phases with the statistics constrained by the observations with our PATCHY approach. In a separated work we will show how to use a mock galaxy sample based on abundance matching to produce mocks with PATCHY (Scoccola et al., in prep.).

We plan to address in forthcoming works a number of studies which have not been done here, like investigating the impact of (a deterministic) non-local bias, computing the two-point and three-point correlation functions in configuration space and further improving the halo PDF fit.

Our method contributes towards an efficient modelling of the halo/galaxy distribution required to estimate uncertainties in our measurements from galaxy redshift surveys. We have also demonstrated that it represents a powerful tool to test various bias models.

spectrum and bispectrum computations were performed on the Sciama High Performance Compute (HPC) cluster which is supported by the ICG, SEPNet. The MultiDark Database used in this paper and the web application providing online access to it were constructed as part of the activities of the German Astrophysical Virtual Observatory as result of a collaboration between the Leibniz-Institute for Astrophysics Potsdam (AIP) and the Spanish MultiDark Consolider Project CSD2009-00064. The BIGMD simulation suite have been performed in the Supermuc supercomputer at LRZ using time granted by PRACE.

REFERENCES

Alimi J.-M. et al., 2012, ArXiv e-prints
Angulo R. E., Baugh C. M., Frenk C. S., Lacey C. G., 2013, ArXiv e-prints
Angulo R. E., Foreman S., Schmittfull M., Senatore L., 2014, ArXiv e-prints
Angulo R. E., Springel V., White S. D. M., Jenkins A., Baugh C. M., Frenk C. S., 2012, MNRAS, 426, 2046
Angulo R. E., White S. D. M., 2010, MNRAS, 405, 143
Baldauf T., Mercolli L., Mirbabayi M., Pajer E., 2014, ArXiv e-prints
Baldauf T., Seljak U., Desjacques V., McDonald P., 2012, Phys.Rev.D, 86, 083540
Baldauf T., Seljak U., Smith R. E., Hamaus N., Desjacques V., 2013, Phys.Rev.D, 88, 083507
Bardeen J. M., Bond J. R., Kaiser N., Szalay A. S., 1986, ApJ, 304, 15
Beltrán Jiménez J., Durrer R., 2011, Phys.Rev.D, 83, 103509
Benitez N. et al., 2014, ArXiv e-prints
Berlind A. A., Weinberg D. H., 2002, ApJ, 575, 587
Bernardeau F., Colombi S., Gaztañaga E., Scoccimarro R., 2002, Phys.Rep., 367, 1
Casas-Miranda R., Mo H. J., Sheth R. K., Boerner G., 2002, MNRAS, 333, 730
Cen R., Ostriker J. P., 1993, ApJ, 417, 415
Chan K. C., Scoccimarro R., Sheth R. K., 2012, Phys.Rev.D, 85, 083509
Chiang C.-T., Wagner C., Schmidt F., Komatsu E., 2014, J.Cosm.Astr.Phys., 5, 48
Cimatti A., Robberto M., Baugh C., Beckwith S. V. W., Content R., Daddi E., De Lucia G., et al, 2009, Experimental Astronomy, 23, 39
Cole S., Kaiser N., 1989, MNRAS, 237, 1127
Cooray A., Hu W., 2001, ApJ, 554, 56
de Jong R. S. et al., 2012, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 8446, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series
de la Torre S. et al., 2013, Astr.Astrophy., 557, A54
de la Torre S., Peacock J. A., 2013, MNRAS, 435, 743
Dekel A., Lahav O., 1999, ApJ, 520, 24
Desjacques V., Crocce M., Scoccimarro R., Sheth R. K., 2010, Phys.Rev.D, 82, 103529
Drinkwater M. J., Jurek R. J., Blake C., Woods D., Pimbblet K. A., Glazebrook K., Sharp R., et al, 2010, MNRAS, 401, 1429
Elia A., Ludlow A. D., Porciani C., 2012, MNRAS, 421, 3472
Frieman J., Dark Energy Survey Collaboration, 2013, in American Astronomical Society Meeting Abstracts, Vol. 221, American Astronomical Society Meeting Abstracts, p. 335.01
Frieman J. A., Gaztanaga E., 1994, ApJ, 425, 392
Fry J. N., Gaztanaga E., 1993, ApJ, 413, 447
Gaztanaga E., Frieman J. A., 1994, ApJ, 437, L13
Gil-Marín H., Schmidt F., Hu W., Jimenez R., Verde L., 2011, J.Cosm.Astr.Phys., 11, 19
Gil-Marín H., Wagner C., Fragkoudi F., Jimenez R., Verde L., 2012, J.Cosm.Astr.Phys., 2, 47
Guzzo L., The Vipers Team, 2013, The Messenger, 151, 41
Hoffmann K., Bel J., Gaztanaga E., Crocce M., Fosalba P., Castander F. J., 2014, ArXiv e-prints
Jasche J., Kitaura F. S., 2010, MNRAS, 407, 29
Jeong D., Komatsu E., 2009, ApJ, 703, 1230
Jing Y. P., 2005, ApJ, 620, 559

Jing Y. P., Börner G., 2004, ApJ, 617, 782

Kaiser N., 1984, ApJ, 284, L9

Kim J., Park C., Gott, III J. R., Dubinski J., 2009, ApJ, 701, 1547

Kitaura F. S., 2012, MNRAS, 420, 2737

Kitaura F. S., 2013, MNRAS, 429, L84

Kitaura F. S., Enßlin T. A., 2008, MNRAS, 389, 497

Kitaura F. S., Gallerani S., Ferrara A., 2012, MNRAS, 420, 61

Kitaura F. S., Heß S., 2013, MNRAS, 435, L78

Kitaura F. S., Jasche J., Metcalf R. B., 2010, MNRAS, 403, 589

Kitaura F. S., Yepes G., Prada F., 2014, MNRAS, 439, L21

Klypin A., Holtzman J., 1997, ArXiv Astrophysics e-prints

Klypin A., Prada F., Yepes G., Hess S., Gottlober S., 2013, ArXiv e-prints

Knebe A. et al., 2011, MNRAS, 415, 2293

Koda J., Kazin E., Blake C., 2014, in American Astronomical Society Meeting Abstracts, Vol. 223, American Astronomical Society Meeting Abstracts, p. 457.17

Laureijs R., 2009, ArXiv e-prints

Leclercq F., Jasche J., Gil-Marín H., Wandelt B., 2013, J.Cosm.Astr.Phys., 11, 48

LSST Dark Energy Science Collaboration, 2012, ArXiv e-prints

Manera M. et al., 2013, MNRAS, 428, 1036

Matarrese S., Verde L., Heavens A. F., 1997, MNRAS, 290, 651

McDonald P., Roy A., 2009, J.Cosm.Astr.Phys., 8, 20

Mo H. J., Jing Y. P., White S. D. M., 1996, MNRAS, 282, 1096

Mo H. J., White S. D. M., 1996, MNRAS, 282, 347

Monaco P., Sefusatti E., Borgani S., Crocce M., Fosalba P., Sheth R. K., Theuns T., 2013, MNRAS

Monaco P., Theuns T., Taffoni G., Governato F., Quinn T., Stadel J., 2002, ApJ, 564, 8

Müller V., Hoffmann K., Nuza S. E., 2011, Baltic Astronomy, 20, 259

Müller V., Maulbetsch C., 2004, in IAU Colloq. 195: Outskirts of Galaxy Clusters: Intense Life in the Suburbs, Diaferio A., ed., pp. 26–28

Neyrinck M. C., 2013, MNRAS, 428, 141

Neyrinck M. C., Aragón-Calvo M. A., Jeong D., Wang X., 2014, MNRAS, 441, 646

Nuza S. E., Sánchez A. G., Prada F., Klypin A., Schlegel D. J., Gottlöber S., Montero-Dorta A. D., et al, 2013, MNRAS, 432, 743

Peacock J. A., Heavens A. F., 1985, MNRAS, 217, 805

Percival W. J. et al., 2014, MNRAS, 439, 2531

Prada F., Klypin A. A., Cuesta A. J., Betancort-Rijo J. E., Primack J., 2012, MNRAS, 423, 3018

Press W. H., Schechter P., 1974, ApJ, 187, 425

Saito S., Baldauf T., Vlah Z., Seljak U., Okumura T., McDonald P., 2014, ArXiv e-prints

Schlegel D., Abdalla F., Abraham T., Ahn C., Allende Prieto C., Annis J., Aubourg E., et al, 2011, ArXiv e-prints

Schneider M. D., Cole S., Frenk C. S., Szapudi I., 2011, ApJ, 737, 11

Scoccimarro R., Feldman H. A., Fry J. N., Frieman J. A., 2001a, ApJ, 546, 652

Scoccimarro R., Feldman H. A., Fry J. N., Frieman J. A., 2001b, ApJ, 546, 652

Scoccimarro R., Sheth R. K., 2002, MNRAS, 329, 629

Sefusatti E., Komatsu E., 2007, Phys.Rev.D, 76, 083004

Seljak U., 2000, MNRAS, 318, 203

Sheth R. K., Lemson G., 1999, MNRAS, 304, 767

Sheth R. K., Mo H. J., Tormen G., 2001, MNRAS, 323, 1

Shirata A., Suto Y., Hikage C., Shiromizu T., Yoshida N., 2007, Phys.Rev.D, 76, 044026

Sigad Y., Branchini E., Dekel A., 2000, ApJ, 540, 62

Smith R. E., Scoccimarro R., Sheth R. K., 2007, Phys.Rev.D, 75, 063512

Somerville R. S., Lemson G., Sigad Y., Dekel A., Kauffmann G., White S. D. M., 2001, MNRAS, 320, 289

Springel V., 2005, MNRAS, 364, 1105

Szapudi I., Pan J., 2004, ApJ, 602, 26

Takada M., Hu W., 2013, Phys.Rev.D, 87, 123504

Tassev S., Zaldarriaga M., 2012, JCAP, 4, 13

Tassev S., Zaldarriaga M., Eisenstein D. J., 2013, J.Cosm.Astr.Phys., 6, 36

Valageas P., Nishimichi T., 2011, Astr.Astrophy., 527, A87

Verde L., Heavens A. F., Matarrese S., Moscardini L., 1998, MNRAS, 300, 747

Verde L. et al., 2002a, MNRAS, 335, 432

Verde L. et al., 2002b, MNRAS, 335, 432

Watson W. A., Iliev I. T., Diego J. M., Gottlöber S., Knebe A., Martínez-González E., Yepes G., 2013, ArXiv e-prints

Weinberg D. H., 1992, MNRAS, 254, 315

White M., Blanton M., Bolton A., Schlegel D., Tinker J., Berlind A., da Costa L., et al, 2011, ApJ, 728, 126

White M., Tinker J. L., McBride C. K., 2014, MNRAS, 437, 2594

Zheng Z., Coil A. L., Zehavi I., 2007, ApJ, 667, 760