Non-Gaussianity of the cosmic infrared background anisotropies - I. Diagrammatic formalism and application to the angular bispectrum

F. Lacasa, A. Penin, N. Aghanim

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **439,** 123–142 (2014)
Advance Access publication 2014 January 30

doi:10.1093/mnras/stt2373

Non-Gaussianity of the cosmic infrared background anisotropies – I. Diagrammatic formalism and application to the angular bispectrum

F. Lacasa,[1⋆] A. Pénin[2] and N. Aghanim[1]

[1]*Institut d'Astrophysique Spatiale, CNRS (UMR8617) Université Paris-Sud 11, Bâtiment 121, F-91405 Orsay, France*
[2]*Aix Marseille Université, CNRS, LAM (Laboratoire d'Astrophysique de Marseille) UMR 7326, F-13388 Marseille, France*

Accepted 2013 December 5. Received 2013 December 2; in original form 2013 August 12

ABSTRACT

We present the first halo model based description of the cosmic infrared background (CIB) non-Gaussianity (NG) that is fully parametric. To this end, we introduce, for the first time, a diagrammatic method to compute high order polyspectra of the 3D galaxy density field. It allows an easy derivation and visualization of the different terms of the polyspectrum. We apply this framework to the power spectrum and bispectrum, and we show how to project them on the celestial sphere in the purpose of the application to the CIB angular anisotropies. Furthermore, we show how to take into account the particular case of the shot noise terms in that framework. Eventually, we compute the CIB angular bispectrum at 857 GHz and study its scale and configuration dependences, as well as its variations with the halo occupation distribution parameters. Compared to a previously proposed empirical prescription, such physically motivated model is required to describe fully the CIB anisotropies bispectrum. Finally, we compare the CIB bispectrum with the bispectra of other signals potentially present at microwave frequencies, which hints that detection of CIB NG should be possible above 220 GHz.

Key words: galaxies: statistics – diffuse radiation – large-scale structure of Universe.

1 INTRODUCTION

The structuration of the large-scale structures and galaxies in the Universe is a long-standing field of research in cosmology, theoretically as well as observationally. Of particular interest is the clustering of galaxies as the latter are biased tracers of the underlying dark matter field. Although perturbation theory (see Bernardeau et al. 2002 for a review) may describe the clustering of dark matter up to mildly non-linear scales and epochs, it breaks down in the regime of highly non-linear gravitational infall and does not prescribe the behaviour of galaxies and baryonic physics with respect to dark matter. Neyman & Scott (1952) pioneered the description of galaxies as distributed in clusters, which were later identified as dark matter haloes as the dark matter paradigm became popular. This latter description has become a fruitful tool, assuming that galaxy properties are determined by the physical characteristics of the host halo, as dark matter simulations have become available. Indeed these simulations have permitted to prescribe the distribution of mass inside haloes (a.k.a. the density profile), their abundance and spatial distribution (e.g. Navarro, Frenk & White 1997). Then analytic or semi-analytic models prescribing the distribution and properties of different galaxy populations may be built (e.g. De Lucia & Blaizot 2007). A common analytical tool is the halo model. In this framework, all dark matter is assumed to be bound up in haloes which are populated with galaxies thanks to the halo occupation distribution (HOD). The standard HOD rules the mean number of galaxies in a halo depending on its mass (Berlind et al. 2003; Kravtsov et al. 2004). Such models have been widely used to reproduce the 2-point correlation function of optically selected galaxies, see e.g. Tinker, Wechsler & Zheng (2010), Coupon et al. (2012), and references therein for the most recent analyses. Most applications to date have concentrated on 2-point statistics, i.e. real-space 2-point correlation function or – auto and cross – power spectrum of tracers.

One tracer of galaxies and dark matter that has been studied thanks to the halo model is the cosmic infrared background (CIB). It was first discovered by Puget et al. (1996), and it stems from the cumulative emission of dusty star-forming galaxies (DSFG). The UV emission from young stars heats up the surrounding dust which consequently reemits in the infrared (IR; from 8 μm to 1 mm) with a typical grey-body law. The CIB is consequently a tracer of star formation, with lower frequencies ($\nu < 220$ GHz) tracing star formation at high redshifts (see e.g. Pénin et al. 2012), as their emission is redshifted into the far-IR/submillimetre domain. Resolutions of current instruments permit to resolve directly only a small fraction of the CIB into individual sources, in particular at far-IR frequencies (<857 GHz) where most of the CIB is unresolved so that sources produce brightness fluctuations generating the CIB anisotropies. The CIB fluctuations trace the clustering of the underlying DSFG

⋆E-mail: fabien.lacasa@ias.u-psud.fr

and their angular power spectra have been measured, in the last few years, over a wide range of wavelengths and scales (Lagache et al. 2007; Viero et al. 2009; Amblard et al. 2011; Planck Collaboration 2011b; Pénin et al. 2012; Planck Collaboration XXX 2013; Thacker et al. 2013). These measurements are usually modelled in the context of the halo model associated with a model of evolution of galaxies (Cooray et al. 2010; Pénin et al. 2012). Until recently, only the power spectrum of the CIB anisotropies had been measured, however, statistical information is contained in the higher order moments.

The hierarchy of n-point correlation functions, for n up to infinity, characterizes statistically a field, univoquely under some regularity conditions. In particular, beyond $n = 2$ it probes the non-Gaussianity (NG) of the field. NG studies have emerged as a research field of interest, as they bring information complementary to power spectrum (or 2-point correlation function) analyses. They are of particular interest for the cosmic microwave background (CMB), for instance, the study of primordial NG discriminates inflation models which are degenerate at the power spectrum level. Lately, the Planck NG constraints have ruled out several primordial models, in particular the possibility of ekpyrotic/cyclic Universe (Planck Collaboration XXIV 2013).

Nevertheless, such measurements are delicate as millimetre observations dedicated to the CMB are contaminated by foregrounds which are non-Gaussian. Extragalactic point sources are of particular importance because they are present all over the sky and the fainter ones cannot be detected nor masked. At CMB frequencies, two types of extragalactic point sources are present: radio-loud sources powered by an active galactic nucleus (Toffolatti et al. 1999) and DSFG constituting the CIB (Lagache, Puget & Dole 2005). NG of these point sources has first been looked for at radio frequencies with *Wilkinson Microwave Anisotropy Probe* (*WMAP*), focusing on radio sources which can be considered unclustered (Toffolatti et al. 1998). Bispectrum predictions based on number counts and measurement on *WMAP* data have been found in agreement (Komatsu et al. 2003) and have permitted to quantify the level of unresolved sources in *WMAP* maps. At higher frequencies, NG of DSFG has been pioneered by Argüeso, González-Nuevo & Toffolatti (2003), González-Nuevo, Toffolatti & Argüeso (2005) and lately Lacasa et al. (2012) with a phenomenological prescription based on the clustered power spectrum. Prior to the present study, no physically based model of the CIB NG was proposed.

This paper builds a halo model description of galaxy clustering at high orders that we apply to predict the NG of CIB anisotropies. This allows for a full model for the CIB anisotropies which, given a galaxy emission model and HOD parameters, computes the power spectrum as well as the bispectrum, and possibly higher order moments. The clustering part of the model is fully parametric which will, at longer term, allow us to constrain these parameters using as much statistical information from the data as possible. In a companion paper (Pénin, Lacasa & Aghanim 2014), referred to as Paper II hereafter, we carry out a Fisher analysis forecast of how the degeneracies of these parameters are broken when combining power spectra and bispectra constraints. In addition, we study in detail, amongst others, the variation of the CIB angular bispectrum with respect to the models of evolution of galaxies, and the frequency evolution of the redshift–halo mass contributions to the bispectrum.

This paper is organized as follows. Section 2 details the halo model formalism, accounting for the shot noise due to the discreteness of galaxies, the occupation statistics (HOD) and shows the resulting 3D power spectrum. In Section 3, we derive the galaxy bispectrum and describe a diagrammatic method to carry the derivation to higher orders. Section 4 introduces harmonic transform of correlation functions on the sky and shows how the 3D polyspectra of a signal project on to the sphere. Taking the example of the CIB, we discuss how the shot noise terms must be accounted for and the necessary regularization at low redshift. The resulting CIB angular bispectrum is shown with its different terms in Section 5 as well as their dependences on the HOD parameters. We also investigate the halo mass contributions to the galaxy bispectrum. Eventually, in Section 6, we compare the obtained CIB angular bispectrum to a previously proposed prescription and to the bispectra of other signals present at microwave frequencies, namely radio sources and CMB. We finally conclude in Section 7.

2 GALAXY CLUSTERING WITH THE HALO MODEL

The most common tool to measure the clustering of galaxies is the 2-point correlation function. At first, such measurements were well reproduced by a simple power law (Davis & Peebles 1983; Madgwick et al. 2003; Le Fèvre et al. 2005). Nevertheless, the progress in the field of large-scale surveys enabled more accurate measurements that rule out such simple modelling (e.g. Zehavi et al. 2004) as they display a cut-off at intermediate scales. A more complex modelling was required leading to the wide use of the halo model (e.g. Cooray & Sheth 2002). In that framework, dark matter is assumed to be bound up in haloes which are virialized spherical objects $\Delta_{\rm vir}$ times denser than the background.[1] The introduction of the galaxies within the haloes is done through the HOD (Kravtsov et al. 2004; Tinker & Wetzel 2010). This modelling has proven to be a very convenient analytical tool to reproduce and interpret the non-linear clustering of dark matter haloes as well as that of galaxies (e.g. Coupon et al. 2012). This section summarizes the ingredient of the halo model necessary to predict the galaxy distribution, and we show how to treat self-consistently the discreteness of galaxies.

2.1 Halo framework

In the halo model framework, galaxies reside in dark matter haloes that are assumed spherical. Hence the galaxy density field at a given point $\boldsymbol{x}$ reads (redshift dependences are implicit throughout this paper, we state them explicitly when needed):

$$n_{\rm gal}(\boldsymbol{x}) = \sum_i n_{\rm gal}(\boldsymbol{x}|i), \tag{1}$$

with i being the halo index. In the literature, $n_{\rm gal}(\boldsymbol{x}|i)$ is assumed, implicitly or explicitly, to be a smooth distribution following the halo density profile. However, galaxies are discrete objects; hence we write

$$n_{\rm gal}(\boldsymbol{x}|i) = \sum_{j=1}^{N_{\rm gal}(i)} \theta(\boldsymbol{x} - \boldsymbol{x}_j), \tag{2}$$

with j being the index of the random galaxies, $N_{\rm gal}(i)$ being the – random – number of galaxies in the halo i, x_j is the – random – position of the jth galaxy and $\theta(x)$ is the galaxy profile. We assume here after that galaxies are drawn independently in the halo.

[1] $\Delta_{\rm vir}$ is the density contrast with respect to the critical density at the halo redshift.

Equation (1) can then be rewritten as

$$n_{\rm gal}(\boldsymbol{x}) = \int {\rm d}M\, {\rm d}^3\boldsymbol{x}_{\rm h} \sum_i \delta(M - M_i)\, \delta^{(3)}(\boldsymbol{x}_{\rm h} - \boldsymbol{x}_i) \times \int {\rm d}^3\boldsymbol{x}_{\rm g} \sum_{j=1}^{N_{\rm gal}(i)} \delta^{(3)}(\boldsymbol{x}_{\rm g} - \boldsymbol{x}_j)\, \theta(\boldsymbol{x} - \boldsymbol{x}_j), \quad (3)$$

with M_i and $\boldsymbol{x}_i$ the mass and position of the dark matter halo i. This equation serves as the basis for the computation of the galaxy clustering throughout this paper.

Furthermore, we make the following set of assumptions for the galaxy distribution.

(i) Haloes are spherical. This is a common assumption in halo models; inclusion of halo shapes was shown to have a 5–10 per cent effect on the 3D bispectrum by Smith, Watts & Sheth (2006).

(ii) The number of galaxies $N_{\rm gal}(i)$ ($= N_{\rm gal}(M_i)$) is drawn from the HOD and depends on the mass M_i of the halo (see Section 2.2).

(iii) The galaxy positions follow the dark matter halo profile. They are drawn from a distribution whose pdf is the normalized halo profile $u(\boldsymbol{x}|M)$ centred on the halo centre:

$$p(\boldsymbol{x}_j|i) = u(\boldsymbol{x}_j - \boldsymbol{x}_i|M_i). \quad (4)$$

(iv) The galaxy profile $\theta(x)$ is a Dirac $\delta^{(3)}(x)$. This assumption holds since the scales probed are larger than the galaxy size.

The galaxy density field may then be characterized with a hierarchy of n-point correlation functions of $n_{\rm gal}$. At first order ($n = 1$), the mean number of galaxies per comoving volume is

$$\overline{n}_{\rm gal} = \int {\rm d}M\, \langle N_{\rm gal}(M)\rangle\, \frac{{\rm d}n_{\rm h}}{{\rm d}M}, \quad (5)$$

where $\frac{{\rm d}n_{\rm h}}{{\rm d}M}$ is the number of haloes with mass M per comoving volume, i.e. the halo mass function. It is convenient to define the galaxy density contrast as

$$\delta_{\rm gal}(\boldsymbol{x}) = \frac{n_{\rm gal}(\boldsymbol{x}) - \overline{n}_{\rm gal}}{\overline{n}_{\rm gal}}. \quad (6)$$

In the following, we will derive the correlation functions of $\delta_{\rm gal}$ and their Fourier transform. To this end, we need to specify the behaviour of the number of galaxies occupying a halo.

2.2 Occupation statistics

High-resolution dissipationless simulations as well as semi-analytic and N-body+gas dynamics studies show that the number of galaxies within a single halo depends on halo mass with a shape consisting of a step, a shoulder and a power-law tail at high mass (e.g. Berlind et al. 2003; Kravtsov et al. 2004). This behaviour can be understood when the number of galaxies, described as a random distribution, is split into the contribution from central galaxies and that of satellite ones $N_{\rm gal} = N_{\rm cen} + N_{\rm sat}$. The former is described as a step-like function while the latter is a power law. High-resolution simulations have brought a lot of progress in the modelling of these two contributions (e.g. Tinker & Wetzel 2010).

The HOD provides us with the number of galaxies in a halo; $N_{\rm cen}$ takes either the value of 0 or 1, and the presence of satellite galaxies is conditioned to $\{N_{\rm cen} = 1\}$. If $N_{\rm cen} = 1$, the number of satellites is drawn from a Poisson distribution (see Zheng et al. 2005) with mean $\overline{n}_{\rm sat}$ (in the following, overbar denotes the average conditioned to $\{N_{\rm cen} = 1\}$).[2] Hence we have

$$\langle N_{\rm sat}\rangle = P(N_{\rm cen}=1) \times \overline{n}_{\rm sat} = \langle N_{\rm cen}\rangle \times \overline{n}_{\rm sat}. \quad (7)$$

Using the properties of the Poisson distribution, we can then compute the expectation values that will be used in the following sections:

$$\langle N_{\rm gal}\rangle = P(N_{\rm cen}=1) \times (1 + \overline{n}_{\rm sat}) = \langle N_{\rm cen}\rangle + \langle N_{\rm sat}\rangle, \quad (8)$$

$$\langle N_{\rm gal}(N_{\rm gal} - 1)\rangle = P(N_{\rm cen}=1) \times \overline{(n_{\rm sat} + 1) n_{\rm sat}} \quad s = \langle N_{\rm cen}\rangle(\overline{n}^2_{\rm sat} + 2\,\overline{n}_{\rm sat}) \quad (9)$$

$$\langle N_{\rm gal}\,(N_{\rm gal} - 1)(N_{\rm gal} - 2)\rangle = P(N_{\rm cen}=1) \times \overline{(n_{\rm sat} + 1)\, n_{\rm sat}\, (n_{\rm sat} - 1)} = \langle N_{\rm cen}\rangle \left(\overline{n}^3_{\rm sat} + 3\overline{n}^2_{\rm sat}\right). \quad (10)$$

We take the mean number of central galaxies as in Pénin et al. (2012)

$$\langle N_{\rm cen}\rangle = P(N_{\rm cen}=1) = \frac{1}{2}\left[1 + {\rm erf}\left(\frac{\log M - \log M_{\rm min}}{\sigma_{\log M}}\right)\right] \quad (11)$$

and the mean number of satellites as

$$\langle N_{\rm sat}\rangle = \frac{1}{2}\left[1 + {\rm erf}\left(\frac{\log M - \log 2M_{\rm min}}{\sigma_{\log M}}\right)\right]\left(\frac{M}{M_{\rm sat}}\right)^{\alpha_{\rm sat}}, \quad (12)$$

where as for the rest of this paper, we use base-10 logarithm.

Such expressions are motivated by hydrodynamical cosmological simulations (Berlind et al. 2003) as well as high-resolution collisionless simulations (Kravtsov et al. 2004). In this formulation, the HOD is thus characterized by four parameters: $M_{\rm min}$, the mass threshold above which a halo contains a central galaxy; $\sigma_{\log M}$ describing the width of the transition from 0 to 1 central galaxy; $M_{\rm sat}$ the typical mass above which a halo contains satellite galaxies; and $\alpha_{\rm sat}$ the index of the power law for the number of satellites at large halo masses. Furthermore, $N_{\rm sat}$ has a cut-off of the same form as the central occupation but with a transition mass twice larger than that of the central galaxy. This prevents haloes which have a low probability of hosting a central galaxy to contain satellite galaxies. Throughout this paper, we use $\log M_{\rm min}/{\rm M}_\odot = 12.6$, $\sigma_{\log M} = 0.65$, $\log M_{\rm sat}/{\rm M}_\odot = 13.6$ and $\alpha_{\rm sat} = 1.1$, unless otherwise stated (Paper II).

2.3 Power spectrum

In the last decade, it has been shown that the 2-point correlation function of galaxies departs from a simple power law (Zehavi et al. 2004). In the framework of the halo model, this has been reproduced by splitting the total 2-point correlation function, or the angular power spectrum, into two contributions, the 1- and 2-halo terms. The discreteness of galaxies further adds a shot noise term, which

[2] For example, we have $\overline{n^2_{\rm sat}} = \overline{n}^2_{\rm sat} + \overline{n}_{\rm sat}$, through the properties of the Poisson distribution.

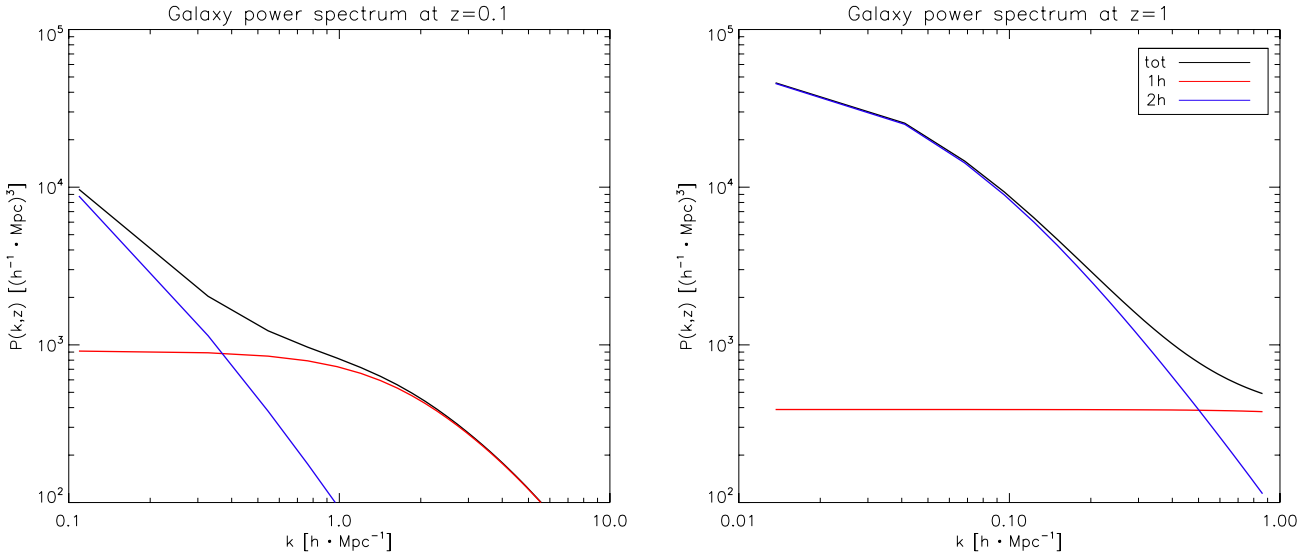

Figure 1. For the Navarro–Frenk–White halo profile and the Sheth & Tormen mass function we plot the 1- and 2-halo terms of the galaxy power spectrum at, respectively, $z = 0.1$ (left-hand panel) and $z = 1$ (right-hand panel). Note that the k range is not identical between the plots.

can be accounted for following a counts-in-cells approach (Peebles 1980). We thus have[3]

$$P_{\rm gal}(k) = P_{\rm gal}^{\rm 1h}(k) + P_{\rm gal}^{\rm 2h}(k) + P_{\rm gal}^{\rm shot}(k). \quad (13)$$

The shot noise contribution is given by

$$P_{\rm gal}^{\rm shot}(k) = \frac{1}{\overline{n}_{\rm gal}}, \quad (14)$$

and will be examined in more detail in Section 4.3. The 1- and 2-halo terms describe, respectively, the contribution of two galaxies within one same halo and that of two galaxies in two different haloes. The 1-halo contribution is

$$P_{\rm gal}^{\rm 1h}(k) = \int {\rm d}M \, \frac{{\rm d}n_{\rm h}}{{\rm d}M} \frac{\langle N_{\rm gal}(M)(N_{\rm gal}(M)-1)\rangle}{\overline{n}_{\rm gal}^2} u(k|M)^2, \quad (15)$$

where $u(k|M)$ is the Fourier transform of the normalized halo profile. Throughout this paper, we use the Navarro–Frenk–White halo profile (Navarro et al. 1997) and the Sheth & Tormen mass function (Sheth & Tormen 1999) that provide us with associated bias functions, in particular the second order bias $b_2(M)$, introduced in Section 3.1, which will be needed for the bispectrum computation later on.

The 2-halo contribution writes

$$P_{\rm gal}^{\rm 2h}(k) = P_{\rm lin}(k) \left(\int {\rm d}M \, \frac{\langle N_{\rm gal}(M)\rangle \, \frac{{\rm d}n_{\rm h}}{{\rm d}M}}{\overline{n}_{\rm gal}} \, b_1(M) \, u(k|M) \right)^2,$$

where $P_{\rm lin}(k)$ is the dark matter power spectrum predicted by linear theory (working at tree level).

On scales larger than the typical halo size, $u(k|M) \to 1$ when $k \to 0$. Hence, the 1-halo term tends towards a constant while the 2-halo term tends towards the linear power spectrum. On small scales, the halo profile smoothes out both terms. For instance, the contribution of massive haloes is smoothed at a smaller cut-off wavenumber compared to that of lower mass haloes. The 1- and 2-halo terms of the power spectrum are exhibited in Fig. 1 at redshifts $z = 0.1$ (left-hand panel) and $z = 1$ (right-hand panel). Note that the k range is different from a panel to another. Indeed, we compute the power spectrum at the wavevectors which will project on observable angular scales on the sky (see Section 5). As expected, we find that the 1-halo term dominates at small scales while the 2-halo term is more important at large scales. We also note that the 1-halo term increases clearly with time, by a factor of ∼2, while the 2-halo does not significantly vary.

[3] See detailed derivation in Appendix A, including the shot noise terms consistently.

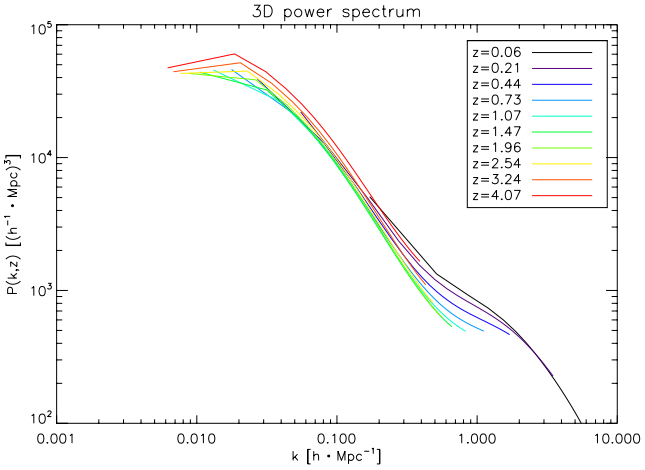

Figure 2. Total 3D galaxy power spectrum as a function of redshift.

In Fig. 2, we show the evolution with redshift of the total galaxy power spectrum (1- and 2-halo terms). The spectrum decreases with increasing redshift up to $z \sim 2$ beyond which it increases with redshift. This behaviour seems counterintuitive as the linear dark matter power spectrum, $P_{\rm lin}(k, z)$, decreases monotonically with increasing redshift. However, galaxies are biased tracers of matter and they are strongly biased at high redshift ($b_{\rm gal} \sim 5.5$ at $z \sim 4$) (Coupon et al. 2012; Jullo et al. 2012), which counterbalances the decrease of $P_{\rm lin}(k, z)$.

3 HIGHER ORDERS WITH THE HALO MODEL

As it is analytical, the halo model can be extended easily to higher orders. Indeed, it has been already used to compute the real-space 3-point correlation function (Wang et al. 2004), comparing favourably with simulations and measurements, as well as the bispectrum in redshift space (Smith, Sheth & Scoccimarro 2008), comparing again favourably with numerical simulations of dark matter.

In the following, we first summarize the 3D bispectrum computation and then we present a new diagrammatic method that can be used to compute the series of high order moments.

3.1 Bispectrum

The 3D galaxy bispectrum at a given redshift can be written as the sum of several terms (see detailed derivation in Appendix B, neglecting primordial NG as argued in the appendix):

$$\begin{aligned} B_{\rm gal}(k_1,k_2,k_3,z) = {} & B_{\rm gal}^{\rm 1h}(k_1,k_2,k_3,z) + B_{\rm gal}^{\rm 2h}(k_1,k_2,k_3,z) \\ & + B_{\rm gal}^{\rm 3-halo}(k_1,k_2,k_3,z) + B_{\rm gal}^{\rm shot2g}(k_1,k_2,k_3,z) \\ & + B_{\rm gal}^{\rm shot1g}(k_1,k_2,k_3,z), \end{aligned} \tag{16}$$

where the 1-halo term writes

$$\begin{aligned} B_{\rm gal}^{\rm 1h}(k_1,k_2,k_3,z) = {} & \int {\rm d}M\, \frac{\langle N_{\rm gal}(N_{\rm gal}-1)(N_{\rm gal}-2)\rangle}{\overline{n}_{\rm gal}^3(z)}\, \frac{{\rm d}n_{\rm h}}{{\rm d}M} \\ & \times u(\boldsymbol{k}_1|M,z)\, u(\boldsymbol{k}_2|M,z)\, u(\boldsymbol{k}_3|M,z). \end{aligned} \tag{17}$$

The 2-halo term writes

$$\begin{aligned} B_{\rm gal}^{\rm 2h}(k_1,k_2,k_3,z) = {} & \mathcal{G}_1(k_1,k_2,z)\; P_{\rm lin}(k_3,z)\, \mathcal{F}_1(k_3,z) \\ & + \mathcal{G}_1(k_1,k_3,z)\; P_{\rm lin}(k_2,z)\, \mathcal{F}_1(k_2,z) \\ & + \mathcal{G}_1(k_2,k_3,z)\; P_{\rm lin}(k_1,z)\, \mathcal{F}_1(k_1,z). \end{aligned} \tag{18}$$

The 3-halo term writes

$$\begin{aligned} B_{\rm gal}^{\rm 3-halo}(k_1,k_2,k_3,z) = {} & \mathcal{F}_1(k_1,z)\, \mathcal{F}_1(k_2,z)\, \mathcal{F}_1(k_3,z) \\ & \times \left[F^s(\boldsymbol{k_1},\boldsymbol{k_2})\, P_{\rm lin}(k_1,z)\, P_{\rm lin}(k_2,z) + {\rm perm.}\right] \\ & + \mathcal{F}_1(k_1,z)\, \mathcal{F}_1(k_2,z)\, \mathcal{F}_2(k_3,z) \\ & \times P_{\rm lin}(k_1,z)\, P_{\rm lin}(k_2,z) + {\rm perm.}, \end{aligned} \tag{19}$$

with

$$\mathcal{F}_1(k,z) = \int {\rm d}M\, \frac{\langle N_{\rm gal}(M)\rangle}{\overline{n}_{\rm gal}(z)}\, \frac{{\rm d}n_{\rm h}}{{\rm d}M}(M,z)\, b_1(M,z)\, u(k|M,z), \tag{20}$$

where $b_1(M,z)$ is the first order bias,

$$\mathcal{F}_2(k,z) = \int {\rm d}M\, \frac{\langle N_{\rm gal}(M)\rangle}{\overline{n}_{\rm gal}(z)}\, \frac{{\rm d}n_{\rm h}}{{\rm d}M}(M,z)\, b_2(M,z)\, u(k|M,z), \tag{21}$$

where $b_2(M,z)$ is the second order bias,

$$\begin{aligned} \mathcal{G}_1(k_1,k_2,z) = {} & \int {\rm d}M\, \frac{\langle N_{\rm gal}(N_{\rm gal}-1)\rangle}{\overline{n}_{\rm gal}(z)^2}\, \frac{{\rm d}n_{\rm h}}{{\rm d}M}(M,z)\, b_1(M,z) \\ & \times u(k_1|M,z)\, u(k_2|M,z) \end{aligned} \tag{22}$$

and

$$F^s(\boldsymbol{k_i},\boldsymbol{k_j}) = \frac{5}{7} + \frac{1}{2}\cos(\theta_{ij})\left(\frac{k_i}{k_j}+\frac{k_j}{k_i}\right) + \frac{2}{7}\cos^2(\theta_{ij}), \tag{23}$$

which stem from non-linear evolution at second order in perturbation theory (e.g., Fry 1984; Gil-Marín et al. 2012).

In the following, we will note 3hcos the part of the 3-halo term containing the F_{ij} kernel (i.e. the first term of equation 19) and we note 3h the part involving the second order bias (i.e. the last lines of equation 19):

$$B_{\rm gal}^{\rm 3-halo}(k_{123},z) = B_{\rm gal}^{\rm 3h}(k_{123},z) + B_{\rm gal}^{\rm 3hcos}(k_{123},z). \tag{24}$$

Eventually, the shot noise terms are

$$B_{\rm gal}^{\rm shot1g}(k_1,k_2,k_3,z) = \frac{1}{\overline{n}_{\rm gal}^2(z)}, \tag{25}$$

$$B_{\rm gal}^{\rm shot2g}(k_1,k_2,k_3,z) = \frac{P_{\rm gal}^{\rm clust}(k_1) + P_{\rm gal}^{\rm clust}(k_2) + P_{\rm gal}^{\rm clust}(k_3)}{\overline{n}_{\rm gal}(z)}, \tag{26}$$

with $P_{\rm gal}^{\rm clust}(k) = P_{\rm gal}^{\rm 1h}(k) + P_{\rm gal}^{\rm 2h}(k)$. These terms will be examined in more detail in Section 4.3.

For illustration, Fig. 3 shows the 1-, 2- and 3-halo terms of the galaxy bispectrum at $z = 0.1$ (left-hand panel) and $z = 1$ (right-hand panel), in the equilateral configuration ($k_1 = k_2 = k_3$). Note that, due to projection effects, here again the k range is not identical between the plots.

We first note that the 2- and 3-halo terms dominate at large scales whereas the 1-halo term dominates at small scales. We see that all terms grow with time, the 1-halo having the fastest growth. We also note that the 3-halo term becomes negative at $z = 0.1$. This comes from the fact that the second order bias can be either positive or negative depending on the halo mass and on the redshift. For haloes with $M \sim 10^{13-14}\,{\rm M}_\odot$ (which typically dominate the 3-halo term) b_2 is positive at high redshifts and it becomes negative at low redshifts (e.g., at $M = 10^{13}\,{\rm M}_\odot$ it changes sign at $z \sim 0.6$, while it changes sign at $z \sim 1.5$ at $M = 10^{12}\,{\rm M}_\odot$ and is always positive at $M = 10^{14}\,{\rm M}_\odot$).

Fig. 4 shows the evolution of the total 3D bispectrum of the galaxies across redshifts. The bispectrum clearly increases with time, as gravitational infall produces non-linear structures so that the density field deviates more and more from its initially Gaussian distribution.

3.2 Diagrammatic and higher orders

The 3D galaxy density field may also be characterized with higher orders than the power spectrum and bispectrum. This is achieved through the use of polyspectra, where $\mathcal{P}_{\rm gal}^{(n)}(\boldsymbol{k}_{1\ldots n}, z_{1\ldots n})$ is the polyspectrum of order n. See Appendix D for a comprehensive definition of polyspectra and their diagonal degrees of freedom.

We introduce here, for the first time, a diagrammatic method permitting the derivation of the equations of the 3D galaxy polyspectrum coming from the halo model. This approach is a generalization of the formalism at second and third order, the power spectrum and bispectrum, respectively. It allows us to have a clear representation and understanding of the different terms involved. It further allows us to avoid cumbersome calculations at high order, by replacing them with diagram drawings.

The first step of the diagrammatic method that we propose here is to draw in the form of diagrams all the possibilities of putting n galaxies in halo(s). Potentially, two or more galaxies can lie at the same point ('contracted') for the shot noise terms. Then for each diagram, the galaxies are labelled, e.g. from 1 to n, as well as the haloes e.g. with α_1 to α_p. An illustration is given in Fig. 5 that displays the three diagrams for the power spectrum; Fig. 6 shows the six diagrams for the bispectrum; and finally Fig. 7 exhibits the 14 diagrams for the trispectrum.

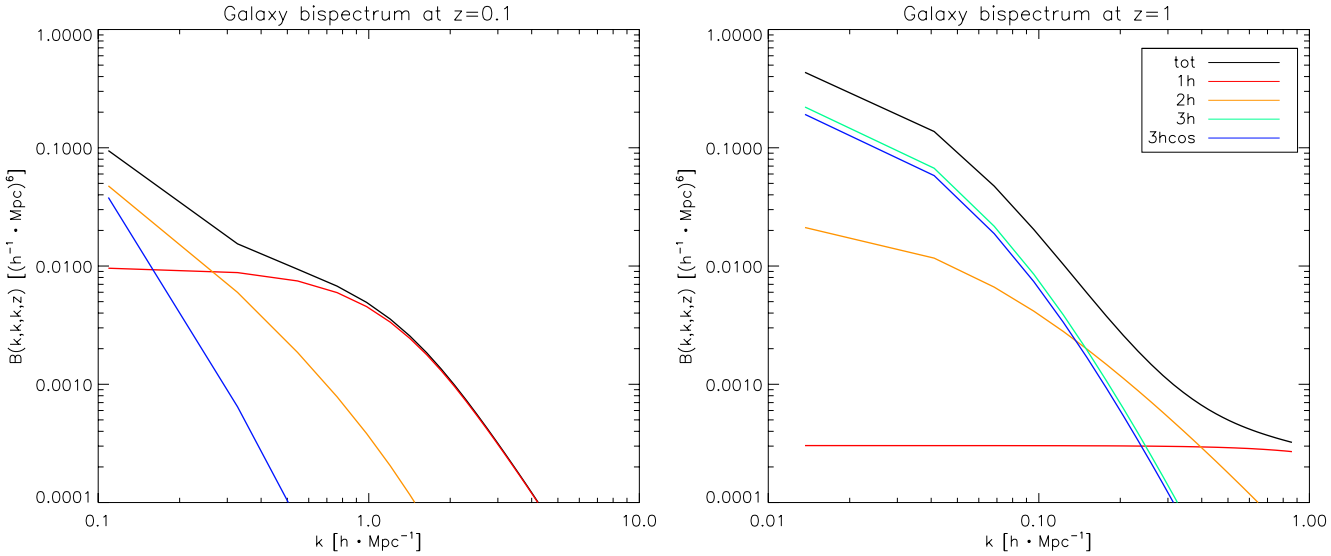

Figure 3. The 1-, 2- and 3-halo terms of the galaxy bispectrum at, respectively, $z = 0.1$ (left-hand panel) and $z = 1$ (right-hand panel). Note that the k range is not identical between the plots.

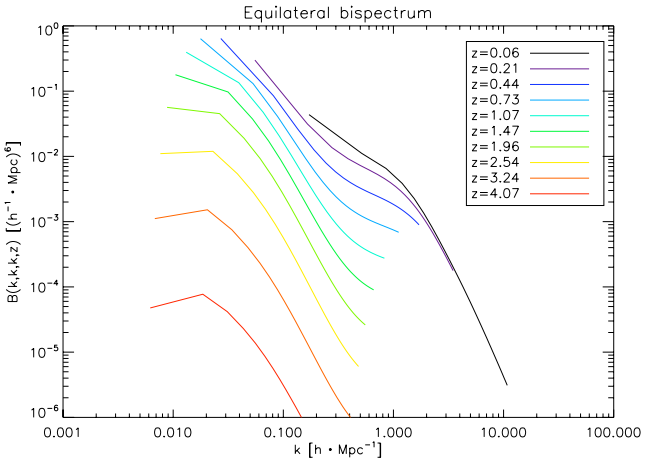

Figure 4. Total galaxy equilateral bispectrum as a function of redshift.

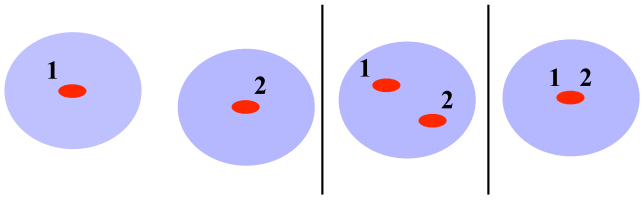

Figure 5. Diagrams for the 3D galaxy power spectrum. From left to right: (2-halo), (1-halo,2-galaxies), (1-halo,1-galaxy). Violet circles represent haloes and red ellipses galaxies.

Each diagram produces a polyspectrum term. This term contains a prefactor $1/\bar{n}_{\rm gal}^{n}$ multiplied by an integral over the halo masses $\int \mathrm{d}M_{\alpha_1 \ldots \alpha_p}$ of several factors. The following 'Feynman'-like rules prescribe these different factors.

(i) For each halo α_j there is a corresponding:

(a) halo mass function $\left.\frac{\mathrm{d}n_{\rm h}}{\mathrm{d}M}\right|_{M_{\alpha_j}}$,

(b) average of the number of galaxy uplets in that halo, e.g. $\langle N_{\rm gal} \rangle$ for a single galaxy in that halo, $\langle N(N-1) \rangle$ for a pair etc.

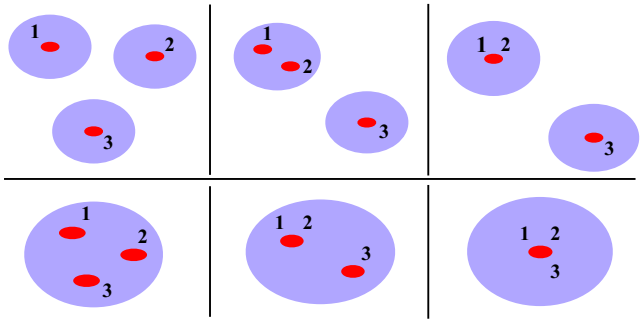

Figure 6. Diagrams for the 3D galaxy bispectrum. From left to right and top to bottom: (3-halo), (2-halo 3-galaxies), (2-halo 2-galaxies), (1-halo 3-galaxies), (1-halo 2-galaxies) and (1-halo 3-galaxies).

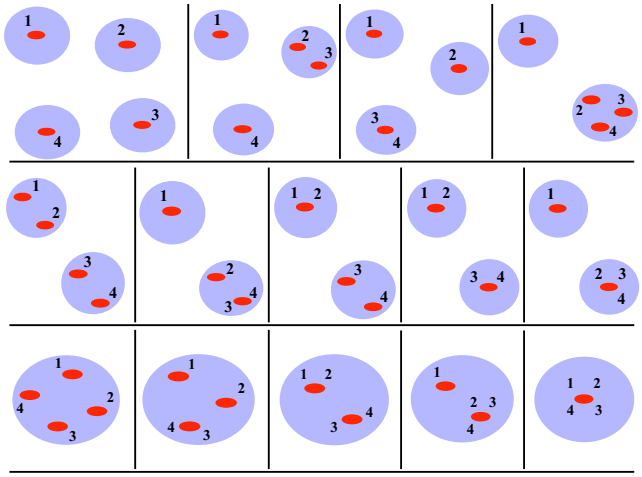

Figure 7. Diagrams for the 3D galaxy trispectrum.

(c) as many halo profiles $u(k|M_{\alpha_j})$ as *different* points, where $k = |\sum_{i \in \rm point} \boldsymbol{k}_i|$.

For example, $k = k_i$ for a non-contracted galaxy i, while $k = |\boldsymbol{k}_{i_1} + \cdots + \boldsymbol{k}_{i_q}|$ for a galaxy contracted q times with labels $i_1 \ldots i_q$.

(ii) The final factor is the halo polyspectrum of order p, conditioned to the masses of the corresponding haloes:

$$\mathcal{P}^{(p)}_{\rm halo}\left(\sum_{i\in\alpha_1}\boldsymbol{k}_i\,,\ldots,\sum_{i\in\alpha_p}\boldsymbol{k}_i\,|\,M_{\alpha_1}\,,\ldots,M_{\alpha_p}\right),$$

where the sum $\sum_{i\in\alpha_j}\boldsymbol{k}_i$ runs over the indexes i of the galaxies inside the halo α_j.

Finally, the possible permutations of the galaxy labels 1 to n in the diagram are taken into account: the contribution is the sum over permutations of $\{1\ldots n\}$ which produce different diagrams. For example, we have seen in Section 3.1 that some contributions to the galaxy bispectrum (namely 1-halo, 3-halo and shot1g) have a single term while others (namely 2-halo and shot 2 g) have three terms.

As an example, the (2-halo, 2-galaxy) term of the bispectrum (upper-right diagram in Fig. 6) yields

$$\begin{aligned}B^{\rm 2h2g}(k_{123}) = \frac{1}{\overline{n}^3_{\rm gal}}\int {\rm d}M_{\alpha 12}\,\langle N_{\rm gal}(M_{\alpha_1})\rangle\,\langle N_{\rm gal}(M_{\alpha_2})\rangle\\ \times\left.\frac{{\rm d}n_{\rm h}}{{\rm d}M}\right|_{M_{\alpha_1}}\left.\frac{{\rm d}n_{\rm h}}{{\rm d}M}\right|_{M_{\alpha_2}}u(|\boldsymbol{k}_1+\boldsymbol{k}_2||M_{\alpha_1})\,u(k_3|M_{\alpha_2})\\ \times\,\mathcal{P}^{(2)}_{\rm halo}(\boldsymbol{k}_1+\boldsymbol{k}_2,\boldsymbol{k}_3|M_{\alpha_1},M_{\alpha_2})+{\rm perm.}\end{aligned}\tag{27}$$

Furthermore, the $\mathcal{P}^{(2)}_{\rm halo}$ term simplifies into $P_{\rm halo}(k_3|M_{\alpha_1},M_{\alpha_2})$ as $\boldsymbol{k}_1+\boldsymbol{k}_2=-\boldsymbol{k}_3$.

We described in the previous sections the mass function, halo profile and the HOD governing the number of galaxies in a halo. The final element needed for the computation of the galaxy polyspectra is a description of the halo polyspectra. To this end, we adopt the local biasing scheme which allows us to compute the halo polyspectrum from the matter polyspectrum as described in Appendix C.

At high order, the halo polyspectrum has thus several possible sources. The first source is the first order biasing of the corresponding dark matter polyspectrum, either primordial (primordial NG) or from perturbation theory. In the bispectrum case, it produces the 3hcos term. The second source is the higher order biasing of lower order dark matter polyspectrum. In the bispectrum case, it produces the 3h term.

Hence, we have proposed a diagrammatic method to compute galaxy polyspectra which gives the power and simplicity of drawings to compute otherwise cumbersome equations at high order. The focus of the next section is to relate these 3D polyspectra to observables on the sky.

4 CIB ANGULAR POLYSPECTRA ON THE SKY

Measurements of the CIB clustering are carried out on the celestial sphere. Hence a statistical characterization of random fields on the sphere is needed, as well as the projection of the statistics of 3D random fields on to the sphere.

In this section, we first describe the formalism of correlation functions on the sphere, then we derive the CIB angular polyspectrum. Eventually, we discuss the shot noise terms and the effect of the flux cut.

4.1 Correlation functions in harmonic space

Given a full-sky map of the temperature $T(\hat{n})$ of some signal on the celestial sphere, it can be decomposed in the harmonic basis as

$$T(\hat{n})=\sum_{\ell m}a_{\ell m}\,Y_{\ell m}(\hat{n}),\tag{28}$$

with

$$a_{\ell m}=\int {\rm d}^2\hat{n}\;Y^*_{\ell m}(\hat{n})\;\Delta T(\hat{n}),\tag{29}$$

with the usual orthonormal spherical harmonics $Y_{\ell m}$.

For a – statistically isotropic – Gaussian field, all the statistical information is contained in the power spectrum C_ℓ, the 2-point correlation function in harmonic space: $\langle a_{\ell m}\,a^*_{\ell' m'}\rangle = C_\ell\,\delta_{\ell\ell'}\,\delta_{mm'}$. For non-Gaussian fields, information is also contained in higher order moments. For instance, the bispectrum $b_{\ell_1\ell_2\ell_3}$ is

$$\langle a_{\ell_1 m_1}a_{\ell_2 m_2}a_{\ell_3 m_3}\rangle = G^{m_1m_2m_3}_{\ell_1\ell_2\ell_3}\times b_{\ell_1\ell_2\ell_3},\tag{30}$$

with the Gaunt coefficient

$$G_{1,2,3}=\int {\rm d}^2\hat{n}\;Y_{123}(\hat{n})\tag{31}$$

$$=\sqrt{\frac{(2\ell+1)_{123}}{4\pi}}\begin{pmatrix}\ell_1&\ell_2&\ell_3\\0&0&0\end{pmatrix}\begin{pmatrix}\ell_1&\ell_2&\ell_3\\m_1&m_2&m_3\end{pmatrix},\tag{32}$$

where $Y_i=Y_{\ell_i m_i}$. In the following, the subscript 123 denotes the product of the corresponding variables, e.g. $X_{123}\equiv X_1X_2X_3$. The Gaunt coefficient is zero unless the triplet $(\ell_1,\,\ell_2,\,\ell_3)$ follows the triangle inequalities and $m_1+m_2+m_3=0$.

Higher order polyspectra $\mathcal{P}^{(n)}(\ell_{1\ldots n}\,,\ell^{\rm d}_{1\ldots(n-3)})$ are defined in Appendix D, along with the case of diagonal independence.

The one-to-one correspondence between the full pdf and the hierarchy of polyspectra for n up to infinity ensures that the polyspectra provide us with a full statistical characterization of a given field on the sphere.

4.2 Anisotropy projection on the sky

The CIB is mostly unresolved in the far-infrared domain which leads to the loss of the redshift information. The emission is thus integrated on a large range of redshift ($1<z<4$), and the CIB temperature in a given direction $\hat{n}$, is a line-of-sight integral of the IR emissivity $\mathtt{j}_\nu$ per comoving volume:

$$T(\hat{n},\nu)=\int {\rm d}z\,\frac{{\rm d}r}{{\rm d}z}\,a(z)\,\mathtt{j}_\nu\,(r(z)\,\hat{n},\,z)\,,\tag{33}$$

with $\mathtt{j}_\nu$ in Jy Mpc^{-1} so that T has units of Jy sr^{-1} and may be converted to a temperature elevation at CMB frequencies through Planck's law. Here r is the comoving distance to redshift z and a is the scalefactor.

Using the Rayleigh/plane wave expansion and the Fourier expansion of the emissivity field, we obtain

$$a_{\ell m}(\nu)=i^\ell\int\frac{{\rm d}^3\boldsymbol{k}}{2\pi^2}\,{\rm d}z\,\frac{{\rm d}r}{{\rm d}z}\,a(z)\,j_\ell(kr)\,Y^*_{\ell m}(\hat{k})\,\mathtt{j}_\nu(\boldsymbol{k},z),\tag{34}$$

where j_ℓ is the spherical Bessel function of order ℓ.

We relate in Appendix E the CIB angular polyspectrum to the 3D emissivity polyspectra, for the terms which are diagonal

independent.[4] This gives

$$\mathcal{P}^{(n)}_{\rm CIB}(\ell_1,\dots,\ell_n) = \left(\frac{2}{\pi}\right)^n \int k^2_{1\dots n}\mathrm{d}k_{1\dots n}\,\mathrm{d}z_{1\dots n}\,x^2\mathrm{d}x \times \left[a(z_i)\left.\frac{\mathrm{d}r}{\mathrm{d}z}\right|_{z_i} j_{\ell_i}(k_i r_i) j_{\ell_i}(k_i x)\right]_{i=1\dots n} \times \mathcal{P}^{(n)}_{\mathrm{j}}(k_{1\dots n}, z_{1\dots n}) \quad (35)$$

with parity invariance imposing $\ell_1 + \dots + \ell_n$ to be even.

In the Limber's approximation, equation (35) simplifies to

$$\mathcal{P}^{(n)}_{\rm CIB}(\ell_1,\dots,\ell_n) = \int \frac{r^2\mathrm{d}r}{r^{2n}} a^n(z)\,\mathcal{P}^{(n)}_{\mathrm{j}}(k^*_{1\dots n}, z), \quad (36)$$

with $k^\star = (\ell + 1/2)/r(z)$. Note that, as a consequence of Limber's approximation, equation (36) now involves the emissivity polyspectrum at a single redshift (i.e. $z_1 = \dots = z_n = z$).

Galaxy power spectrum and galaxy emissivity are often related assuming implicitly or explicitly that all galaxies at a given redshift z have the same luminosity (e.g. Knox et al. 2001; Pénin et al. 2012; Xia et al. 2012). We will refer to this approach as the constant-luminosity assumption. It yields

$$\delta\mathrm{j}_\nu(\boldsymbol{k}, z) = \bar{\mathrm{j}}_\nu(z)\cdot\delta_{\rm gal}(\boldsymbol{k}, z), \quad (37)$$

where the average emissivity $\bar{\mathrm{j}}_\nu$ is computed from galaxy number counts.

In the present study, we propose a more general assumption referred to as the flux-abundance independence assumption. That is, we assume that the flux is stochastic with a distribution given by the number counts, and furthermore that the stochasticity is independent of the galaxy position/abundance. The commonly used constant-luminosity assumption is a special case of the flux-abundance independence assumption when the luminosity function $\frac{\mathrm{d}^2N}{\mathrm{d}L\,\mathrm{d}V}$ is a dirac. Note that, although the flux-abundance independence is a more general assumption than the constant-luminosity one, it does not capture the possibility that the luminosity may depend on underlying variables having an influence on galaxy abundance (e.g. the host halo mass).

For non-shot noise terms, the flux-abundance independence assumption yields the same relation as the constant-luminosity assumption:

$$\mathcal{P}^{(n)}_{\mathrm{j}}(k^*_{1\dots n}, z) = \bar{\mathrm{j}}_\nu(z)^n\,\mathcal{P}^{(n)}_{\rm gal}(k^*_{1\dots n}, z). \quad (38)$$

The difference between the two assumptions impacts the shot noise terms as discussed in detail in Section 4.3.

For non-shot noise terms, the angular polyspectrum then takes the form

$$\mathcal{P}^{(n)}_{\rm CIB}(\ell_1,\dots,\ell_n) = \int \frac{\mathrm{d}z}{r^{2n-2}}\frac{\mathrm{d}r}{\mathrm{d}z}\,a^n(z)\,\bar{\mathrm{j}}_\nu(z)^n\,\mathcal{P}^{(n)}_{\rm gal}(k^*_{1\dots n}, z). \quad (39)$$

In particular at second order the power spectrum is

$$C_\ell(\nu) = \int \frac{\mathrm{d}z}{r^2}\frac{\mathrm{d}r}{\mathrm{d}z}\,a^2(z)\,\bar{\mathrm{j}}^2_\nu(z)\,P_{\rm gal}(k^*, z) \quad (40)$$

as given e.g. by Knox et al. (2001), Pénin et al. (2012) and Xia et al. (2012).

And at third order, the bispectrum is

$$b_{\ell_1\ell_2\ell_3} = \int \frac{\mathrm{d}z}{r^4}\frac{\mathrm{d}r}{\mathrm{d}z}\,a^3(z)\,\bar{\mathrm{j}}^3_\nu(z)\,B_{\rm gal}(k^*_{123}, z). \quad (41)$$

[4] Note that not all terms may be diagonal independent, in which case equation (E2) must be used in all generality.

4.3 Shot noise

The shot noise terms, for the power spectrum or for the bispectrum, correspond to terms in the correlation function involving multiple times the same galaxy. For the power spectrum, the halo model gives the galaxy shot noise power spectrum

$$P^{\rm shot}_{\rm gal}(k, z) = \frac{1}{\bar{n}_{\rm gal}(z)}. \quad (42)$$

With the constant-emissivity assumption, we get the angular power spectrum

$$C^{\rm shot}_\ell = \int_{z=0}^{\infty}\frac{\mathrm{d}z}{r^2}\frac{\mathrm{d}r}{\mathrm{d}z}\,a^2(z)\,\bar{\mathrm{j}}^2_\nu(z)\,P_{\rm shot}(k, z). \quad (43)$$

The shot noise level can be predicted from number counts (see e.g. Lacasa et al. 2012) as

$$C^{\rm shot}_\ell = \int_{z=0}^{\infty}\int_0^{S_{\rm cut}} S^2\,\frac{\mathrm{d}^2n}{\mathrm{d}S\,\mathrm{d}z}\,\mathrm{d}S\,\mathrm{d}z. \quad (44)$$

These two formulae agree if indeed all sources have the same luminosity, potentially depending on redshift. However, in reality, sources do not have the same luminosity; and specifically equation (44) will give more weight to bright galaxies – averaging S^2 – as compared to equation (43) – averaging S.

With the flux-abundance independence assumption, the distribution of luminosities can be incorporated in the model by introducing the nth order emissivities:

$$\mathrm{j}^{(n)}_\nu(z) = \frac{(1+z)^n\,r(z)^{2n-2}}{\frac{\mathrm{d}r}{\mathrm{d}z}}\int_0^{S_{\rm cut}} S^n\,\frac{\mathrm{d}^2N}{\mathrm{d}S\,\mathrm{d}z}\mathrm{d}S, \quad (45)$$

which effectively average S^n instead of S. Removing $\bar{n}_{\rm gal}(z)$ factors from the shot noise 3D power spectrum, the shot noise angular power spectrum becomes

$$C^{\rm shot}_\ell = \int \frac{\mathrm{d}z}{r^2}\frac{\mathrm{d}r}{\mathrm{d}z}\,a^2(z)\,\mathrm{j}^{(2)}_\nu(z)\times 1, \quad (46)$$

which can be shown to be equivalent to equation (44).

At any order, with the diagrammatic approach described in Section 3.2, the shot noise terms can be computed and integrated over redshifts with equation (39) provided the following modification.

For each contraction, $\mathrm{j}^p_\nu(z)/\bar{n}^{p-1}_{\rm gal}$ must be replaced by the pth order emissivity $\mathrm{j}^{(p)}_\nu(z)$, where p is the order of the contraction under consideration.

For example, for the bispectrum, the 3D shot noise contains three terms (see Appendix B):

$$B^{\rm shot}_{\rm gal}(k_1, k_2, k_3, z) = B^{\rm 1h-1g}_{\rm gal} + B^{\rm 1h-2g}_{\rm gal} + B^{\rm 2h-2g}_{\rm gal}, \quad (47)$$

with

$$B^{\rm 1h-1g}_{\rm gal}(k_1, k_2, k_3, z) = \frac{1}{\bar{n}^2_{\rm gal}(z)}, \quad (48)$$

$$B^{\rm 1h-2g}_{\rm gal}(k_1, k_2, k_3, z) = \frac{P^{\rm 1h-2g}_{\rm gal}(k_1) + P^{\rm 1h-2g}_{\rm gal}(k_2) + P^{\rm 1h-2g}_{\rm gal}(k_3)}{\bar{n}_{\rm gal}(z)}, \quad (49)$$

$$B^{\rm 2h-2g}_{\rm gal}(k_1, k_2, k_3, z) = \frac{P^{\rm 2h-2g}_{\rm gal}(k_1) + P^{\rm 2h-2g}_{\rm gal}(k_2) + P^{\rm 2h-2g}_{\rm gal}(k_3)}{\bar{n}_{\rm gal}(z)}. \quad (50)$$

Now taking into account luminosity distribution, the 1-galaxy angular shot noise takes the form

$$b^{\rm shot1g}_{\ell_{123}} = \int \frac{{\rm d}z}{r^4}\frac{{\rm d}r}{{\rm d}z}\, a^3(z)\, {\rm j}^{(3)}_\nu(z) \times 1, \tag{51}$$

and the 2-galaxy shot noise

$$b^{\rm shot2g}_{\ell_{123}} = \int \frac{{\rm d}z}{r^4}\frac{{\rm d}r}{{\rm d}z}\, a^3(z)\, {\rm j}^{(1)}_\nu(z)\, {\rm j}^{(2)}_\nu(z) \times \left[P^{\rm clust}_{\rm gal}(k^*_1) + P^{\rm clust}_{\rm gal}(k^*_2) + P^{\rm clust}_{\rm gal}(k^*_3)\right], \tag{52}$$

with $P^{\rm clust}_{\rm gal}(k) = P^{\rm 1h}_{\rm gal}(k,z) + P^{\rm 2h}_{\rm gal}(k,z)$ and as usual $k^* = \frac{\ell+1/2}{r(z)}$.

In this formulation, the flux cut, or more generally the selection function, is implemented in the p-order emissivities. The shot noise terms are quite sensitive to the value of this flux limit, which in turn depends on the instrumental setup used for the observation. Hence, the amplitude of the shot noise terms may change depending on the instrument considered.

4.4 Flux cut and low-redshift contribution

As can be seen in equations (46) and (51), the shot noise equations diverge if the emissivities tend to a constant as $z \to 0$. This is indeed the case in the Euclidean case ($\frac{{\rm d}n}{{\rm d}S} \propto S^{-5/2}$) where equation (44) diverges if no flux cut is applied.

In practice, the flux cut reduces the contribution from low-redshift sources which dominate the counts. This needs to be reflected in the halo model, where the number of objects is dictated by the HOD. For simplicity, we implement the flux cut in terms of a cut-off in the redshift integrals at $z_{\rm cut}$. Below this redshift, a typical galaxy with luminosity $L_\star$ (the knee of the luminosity function) has a flux $S \geq S_{\rm cut}$:

$$S_{\rm cut} = \frac{L_*}{4\pi\, d_{\rm L}^2(z_{\rm cut})}, \tag{53}$$

with $d_{\rm L}$ being the luminosity distance.

The effect of the flux cut on the galaxy clustering has not been considered previously in the CIB power spectrum literature, although it is theoretically necessary. We found that it has a small effect on the power spectrum, mostly on the 1-halo term and for some values of the HOD parameters ($\alpha > 1$). The redshift cut has more effect on the bispectrum and potentially at higher order, as they may be more sensitive to low redshift because of the $r(z)^{2n-2}$ denominator in equation (39) which goes to zero as $z \to 0$.

5 RESULTS

5.1 The CIB angular bispectrum

In the following, we apply our formalism to the computation of the bispectrum of the CIB. To this end, the HOD best-fitting parameters were obtained so as to reproduce CIB power spectrum constraints (Paper II). We also use the galaxy emission model by Béthermin et al. (2011) at 350 μm = 857 GHz with a flux cut at 0.82 Jy (Planck Collaboration 2011a), which implies a redshift cut at $z_{\rm cut} = 0.03$.

The obtained CIB bispectrum and its different terms are displayed in Fig. 8 in some particular configurations: equilateral, orthogonal isoceles, flat isoceles and squeezed. We consider a multipole range of $\ell = 30$–2000. In this range of scales, shot noise terms are found to be negligible, as happens for the power spectrum (Pénin et al. 2012). However, they are expected to dominate on smaller scales ($\ell \sim 5000$ for the particular galaxy emission model used in the present study). The 2- and 3-halo terms dominate on large scales $\ell \leq 300$, while the 1-halo dominates on small scales, except in the squeezed configuration where the 2-halo term dominates for $\ell > 300$.

An interesting point is the configuration dependence of the bispectrum. Fig. 9 shows the CIB bispectrum at 350 μm plotted in the geometrical parametrization proposed in Lacasa et al. (2012). The values of the bispectrum are colour coded from violet to red. In this parametrization, each subplot shows a slice of the bispectrum at a given perimeter, indicated in the bottom-left corner. All triangles within the perimeter bin are shown in the subplot, with squeezed triangles in the upper-left corner, equilateral triangle in the upper-right corner and flat isosceles triangle in the bottom corner. The geometrical parametrization permits us to visualize, at the same time, both the scale and configuration dependence of the bispectrum. We do not display the 1-galaxy shot noise term since it is constant.

As seen already in Fig. 8, the CIB bispectrum decreases generally with scale, as it is the case for the power spectrum. A distinctive feature of the CIB bispectrum is that it peaks in the squeezed configuration, as already noted by Lacasa et al. (2012). Figs 10–14 show the different terms of the CIB bispectrum plotted in the geometrical parametrization with the colour range adapted to each term to highlight its variations.

It is worth noting that most terms peak strongly in the squeezed configurations, except the 1-halo term. The latter has little configuration dependence but strong scale dependence and is slightly weaker in squeezed. The 3hcos term from perturbation theory is, unlike other terms, more important in the flat configuration than in the equilateral. This is due to the F^s kernel (equation 23) which is more important in flat than in equilateral configurations.

5.2 Dependences on the HOD parameters

We now investigate how the bispectrum and its different components vary with respect to the HOD parameters. We use a fiducial set of HOD parameters inspired from previous studies of CIB anisotropies, namely Viero et al. (2009), Amblard et al. (2011), Planck Collaboration (2011b), and we vary them individually. Only one parameter is varied, by typically 2σ (for instance, Planck Collaboration 2011b), while the others keep their fiducial values. We consider $\alpha_{\rm sat} = 1.3$, $\log\, M_{\rm min} = 12$, $M_{\rm sat} = 10\, M_{\rm min}$ and $\sigma_{\log {\rm M}} = 0.65$. As a reminder, increasing (decreasing) $\alpha_{\rm sat}$ leads to a higher (lower) number of satellite galaxies. The value of $M_{\rm min}$ rules the mass at which a halo contains a central galaxy and $M_{\rm sat}$ is the average mass of a halo hosting satellite galaxies.

Fig. 15 displays the equilateral bispectrum at 350 μm and its components. We do not display the 1-galaxy shot noise term as it is independent of the HOD parameters. First, we see that the amplitude of each component of the bispectrum is very sensitive to the value of each HOD parameter. Indeed, the amplitude can increase/decrease by up to two orders of magnitude. For instance, a higher $\alpha_{\rm sat}$ leads to more power at all scales as it means more satellite galaxies as compared to a lower $\alpha_{\rm sat}$. The shape of the bispectrum terms only varies slightly with the HOD parameters. Furthermore, variations of different parameters induce similar changes on the equilateral bispectrum suggesting degeneracies between the HOD parameters. The degeneracy and how it could be broken are discussed in Paper II.

It is interesting to notice that $\alpha_{\rm sat}$ induces strong variations in the relative contributions of each term of the bispectrum. Fig. 16 shows each component of the bispectrum for two values of $\alpha_{\rm sat}$. For $\alpha_{\rm sat}$= 1.4 the 1-halo term dominates all the other contributions on nearly all angular scales. The bispectrum thus appears much more

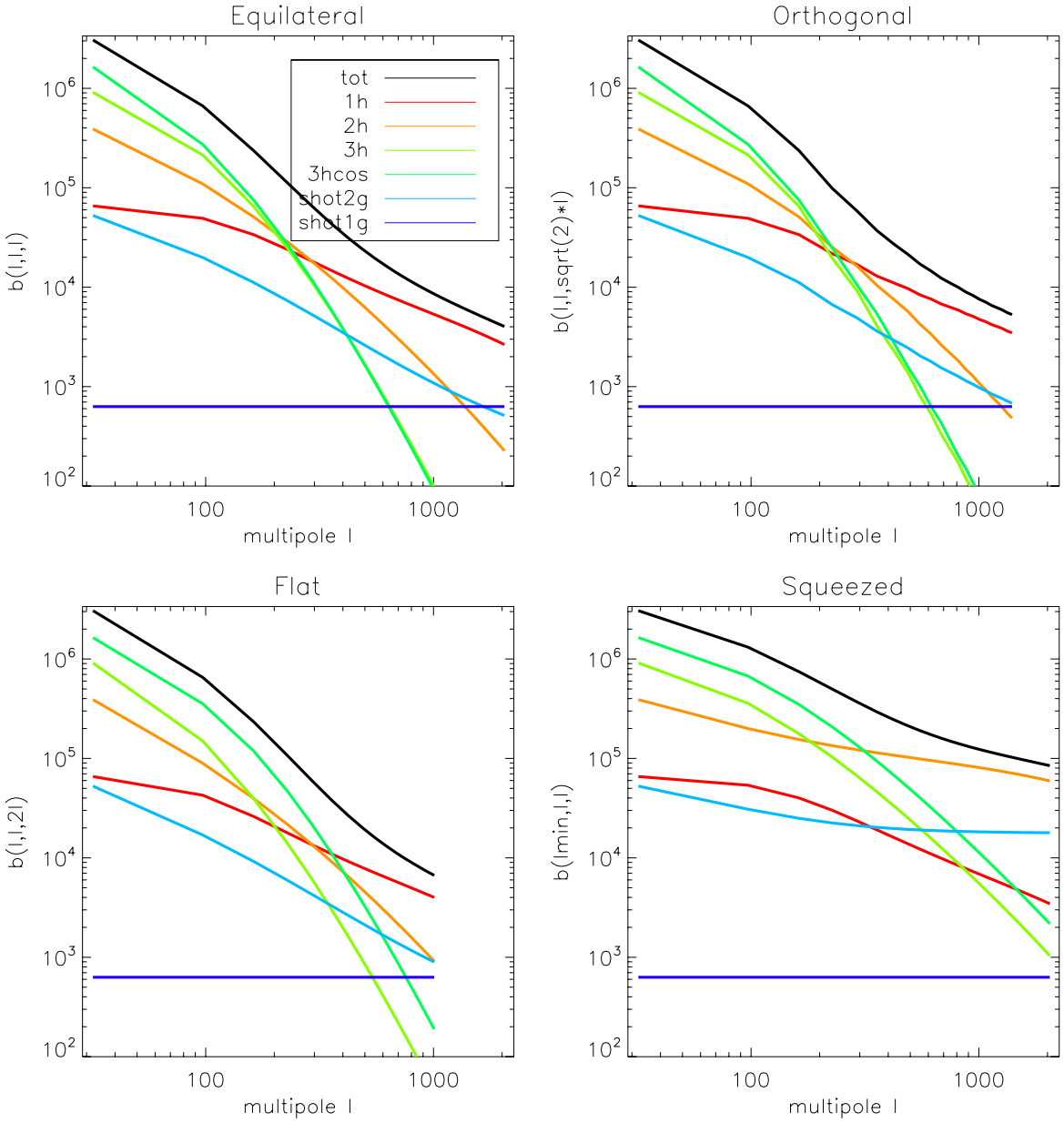

Figure 8. CIB bispectrum and its different terms in some particular configurations at 350 μm = 857 GHz. The squeezed configuration uses $\ell_{\rm min} = 32$.

sensitive to $\alpha_{\rm sat}$ than the power spectrum. This effect might help to alleviate the tension that exists today about the measured values of $\alpha_{\rm sat}$ (see Paper II).

5.3 Halos contribution

We can now focus on halo contribution to the 3D galaxy bispectrum. Most terms, except the 1-halo term, are not a simple integration over halo mass, but include cross-terms between haloes of different masses. The 3D galaxy bispectrum cannot be simply divided into a sum of contributions of different mass bins. We may focus on the dependence of the bispectrum on the halo mass upper cut-off. In Fig. 17, we display the total 3D galaxy bispectrum in the equilateral configuration, as a function of the halo mass upper cut-off, respectively, at $z = 0.1$ and 1.

At $z = 1$, we see that the equilateral bispectrum saturates for a cut-off at a few $10^{14}\,{\rm M}_\odot$, except at small scales where saturation is reached for $\sim 10^{15}\,{\rm M}_\odot$. So at this redshift, haloes with masses larger than a few $10^{15}\,{\rm M}_\odot$ contribute negligibly to the bispectrum, as there are too few of them. At $z = 0.1$, massive haloes contribute more importantly to the bispectrum, which saturates at a cut-off $\sim 10^{15}\,{\rm M}_\odot$ on small scales and at $\sim 4 \times 10^{15}\,{\rm M}_\odot$ on large scales. This reflects also the increase in number of massive haloes at low redshift.

We checked each of the bispectrum terms and found that the 1-halo term is the most sensitive to massive haloes, while the 3-halo terms (both 3h and 3hcos) are the least sensitive, saturating at a few $10^{14}\,{\rm M}_\odot$ even at low redshift. This behaviour is expected since the 3-halo terms involve the product of three halo mass functions. This penalizes massive haloes in the tail of the

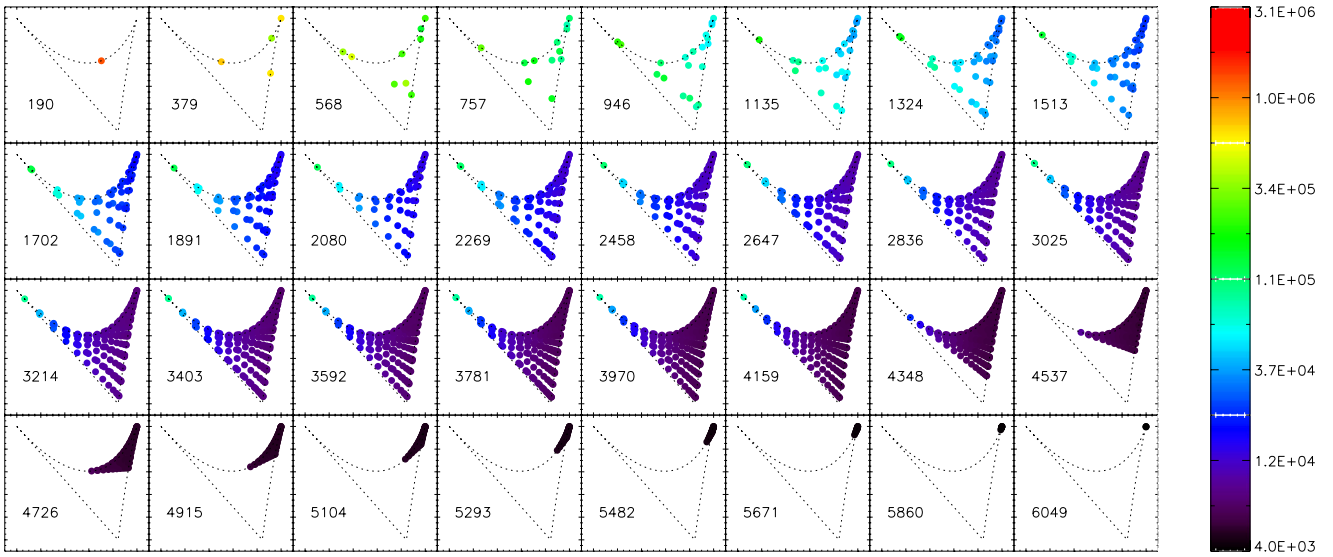

Figure 9. CIB bispectrum in the geometrical parametrization at 350 μm = 857 GHz. It exhibits a peak in the squeezed configurations (upper-left corner of each subplot).

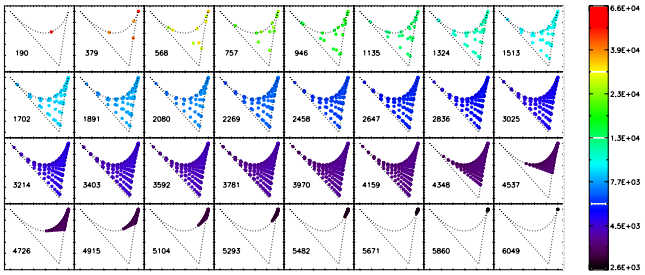

Figure 10. The 1-halo term of the CIB bispectrum in the geometrical parametrization at 350 μm = 857 GHz. It has strong scale dependence but few dependence on configuration.

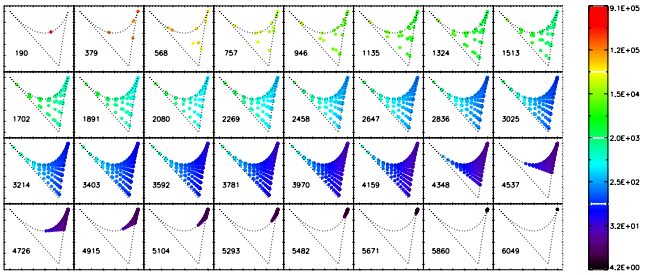

Figure 12. The 3-halo term of the CIB bispectrum in the geometrical parametrization at 350 μm = 857 GHz. It exhibits a peak in the squeezed configurations (upper-left corner of each subplot).

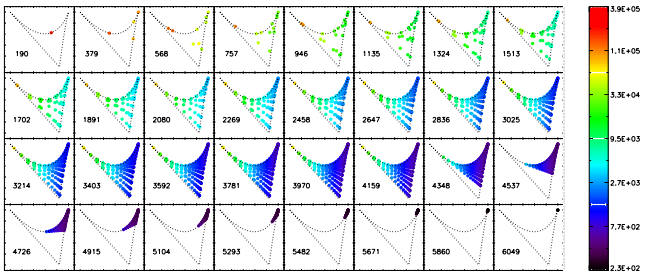

Figure 11. The 2-halo term of the CIB bispectrum in the geometrical parametrization at 350 μm = 857 GHz. It exhibits a peak in the squeezed configurations (upper-left corner of each subplot).

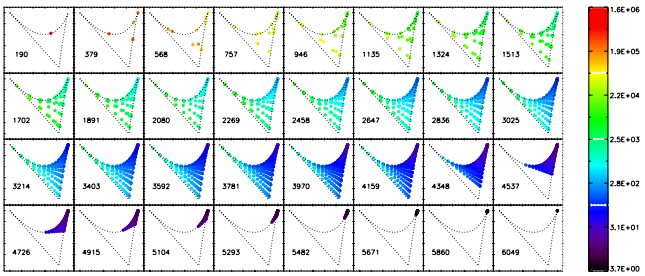

Figure 13. The 3-halo cos term of the CIB bispectrum in the geometrical parametrization at 350 μm = 857 GHz. It exhibits a peak in the squeezed configurations (upper-left corner of each subplot), and in the flat configurations (down corner of each subplot) for some perimeter bins.

mass function as there are few of them. On the contrary, the 1-halo term involves only one mass function, and it gives more weight to massive haloes containing more galaxies (see $N_{\rm gal}(M)^3$ weight).

The mass contribution to the angular bispectrum depends on the galaxy emissivities which give weights to each redshift. It hence depends on the specific galaxy evolution model chosen. This is discussed in detail in the companion paper (Paper II).

6 DISCUSSION

6.1 Comparison with empirical prescription

A simple empirical prescription for the CIB bispectrum based on its power spectrum was proposed by Lacasa et al. (2012). It reads

$$b_{123}^{\rm CIB} = \alpha \sqrt{C_{\ell_1}^{\rm CIB}\, C_{\ell_2}^{\rm CIB}\, C_{\ell_3}^{\rm CIB}}, \tag{54}$$

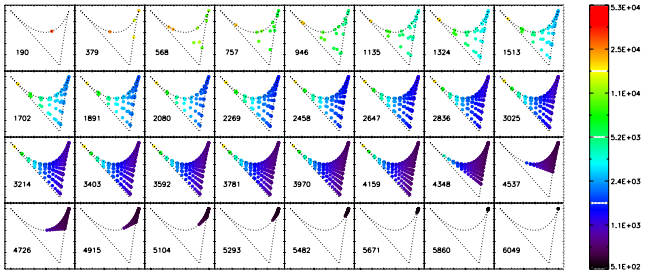

Figure 14. 2-galaxy shot noise term of the CIB bispectrum in the geometrical parametrization at 350 μm = 857 GHz. It exhibits a peak in the squeezed configurations (upper-left corner of each subplot).

where α is a proportionality constant that can be computed with the number counts of IR galaxies and their flux cut.

We compare this prescription with the CIB bispectrum obtained from the halo model theory. For the prescription, we used the CIB power spectrum predicted by the halo model with the same parameters as for the bispectrum, and the best-fitting value $\alpha = 2.25 \times 10^{-3}$ (compared to $\alpha = 3 \times 10^{-3}$ found by Lacasa et al. 2012 on simulations by Sehgal et al. 2010). Fig. 18 shows both bispectra at 350 μm (the prescription is in red and the halo model in black).

We see that the prescription reproduces reasonably well the shape of the bispectrum in equilateral, orthogonal and flat configurations. However, the empirical prescription shows an excess of power at intermediate multipoles $\ell \in [100, 1000]$ in the equilateral and isosceles orthogonal configurations. Finally, the prescription does not recover the bispectrum in the squeezed configuration, departing from

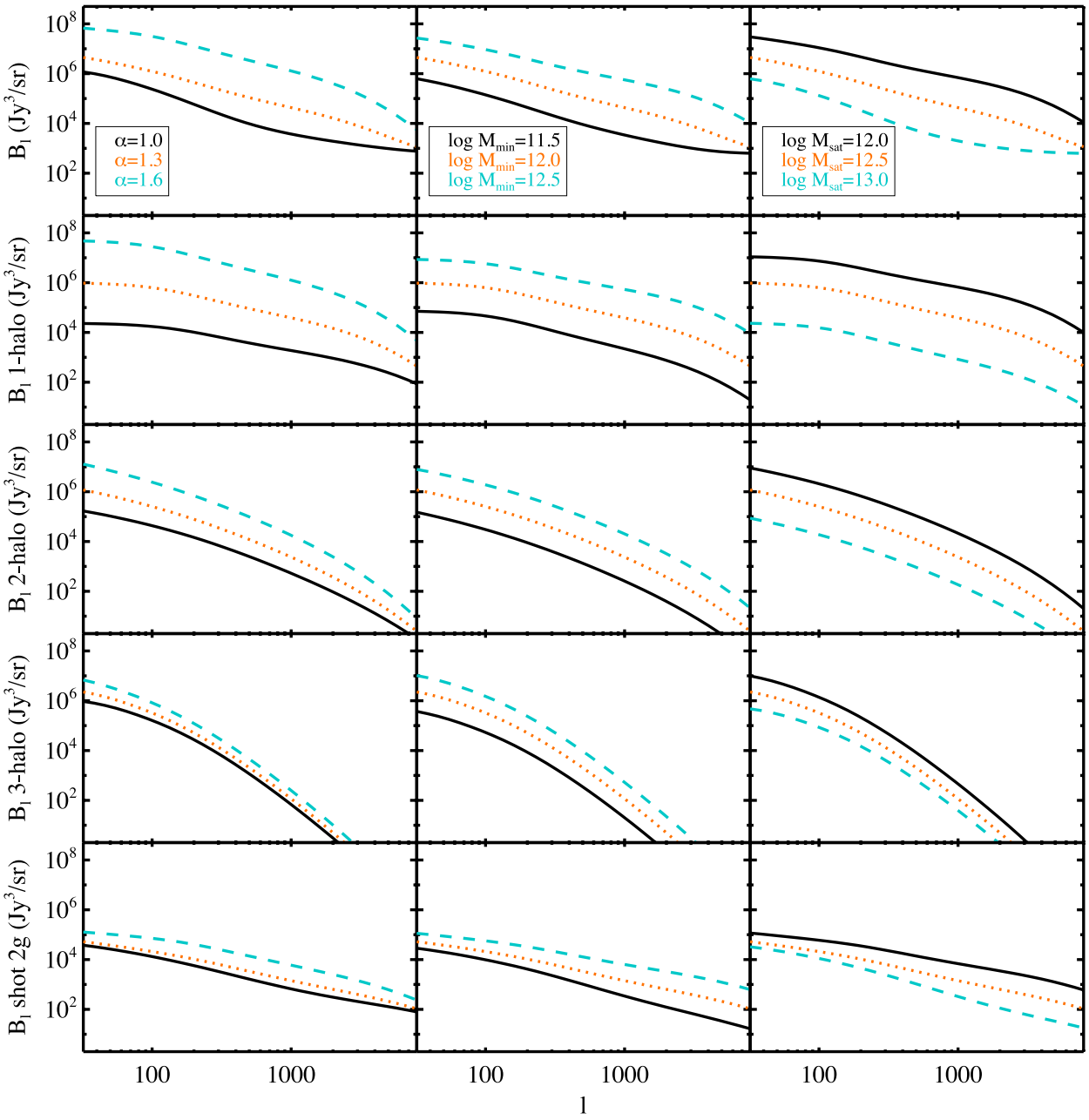

Figure 15. Equilateral bispectra at 350 μm for several sets of the HOD parameters. Only one parameter is varied while the others keep their fiducial values.

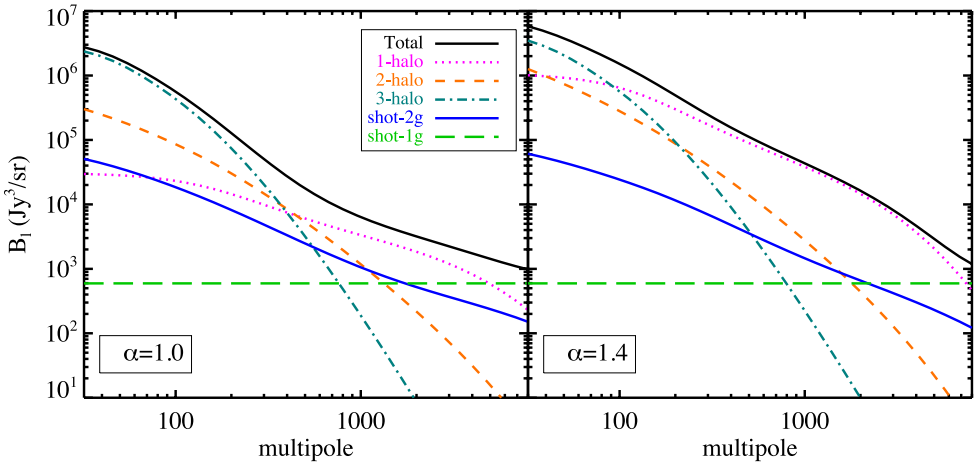

Figure 16. Equilateral bispectrum at 350 μm with α_{sat}= 1 and α_{sat}= 1.4. The other HOD parameters are fixed at their fiducial values.

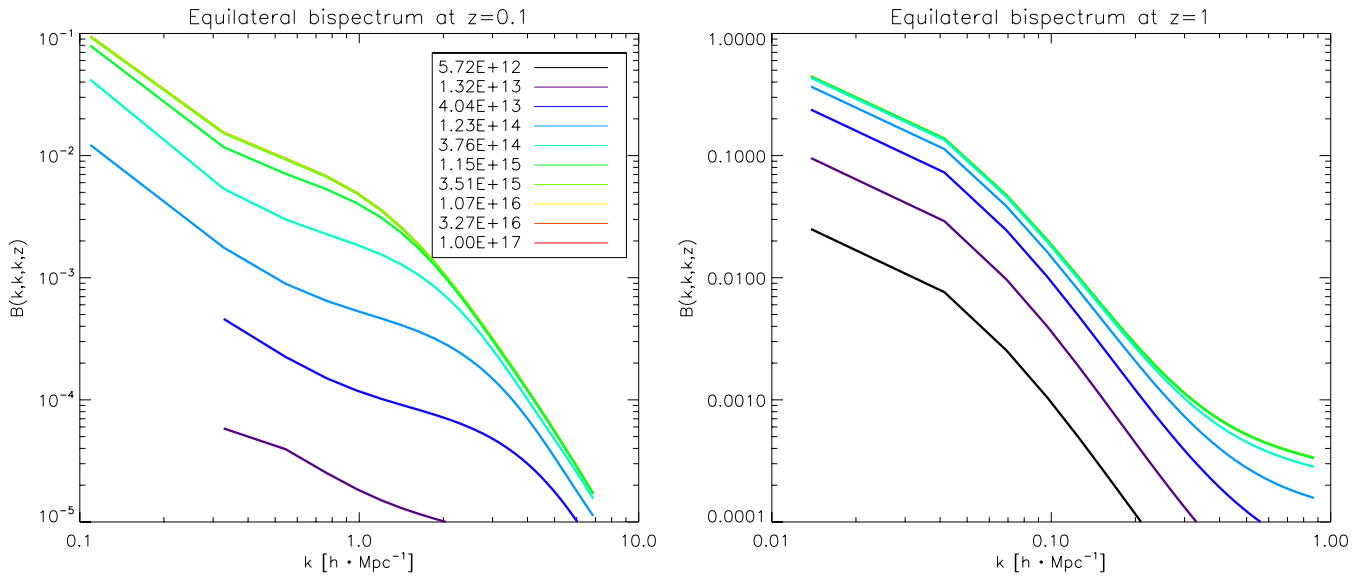

Figure 17. Total galaxy equilateral bispectrum at redshift 0.1 and 1, as a function of the halo mass upper cut-off from $M_{cut} = 5.7\,10^{12}\,M_\odot$ to $10^{17}\,M_\odot$.

the halo model at $\ell \sim 300$, i.e. when the 2-halo term begins to dominate the squeezed bispectrum (see Fig. 8).

The empirical prescription therefore gives a reasonable overall fit of the CIB bispectrum, for the considered galaxy emission model and HOD parameters. Furthermore, the prescription gives a separable template (i.e. $b_{123} = f(\ell_1)f(\ell_2)f(\ell_3)$ for some function f); it thus provides a convenient way to assess quickly the overall level of CIB NG present in a CMB map. This is useful in particular to assess the level of contamination of f_{NL} estimation (see Lacasa & Aghanim 2012). Nevertheless, the empirical prescription does not reproduce completely the theoretical bispectrum derived from the halo model. Additionally, in the companion paper (Paper II), we show that galaxy evolution models which are indistinguishable at the power spectrum level can produce distinguishable theoretical bispectra with the halo model. On the contrary, for different galaxy models the empirical prescription would give indistinguishable bispectra, as it is based on the power spectra. A full computation of the bispectrum using the halo model is thus necessary if one were to interpret a CIB NG measurement.

6.2 Comparison with radio sources and CMB bispectra

At microwave frequencies, several extragalactic signals other than the CIB are present. In particular, the CMB and radio sources which emit mostly at low frequencies, $\nu \leq 217$ GHz (Ade et al. (Planck Collaboration) 2013). We thus compare the bispectra of those extragalactic signals with the CIB theoretical bispectrum derived from the halo model.

Radio sources can be considered distributed randomly on the sky (Toffolatti et al. 1998), at least for the brightest sources. Hence, the extragalactic radio background has a white noise distribution entirely characterized by the number counts $\frac{dN}{dS}$. The expected power spectrum and bispectrum is

$$C_\ell^{RAD} = \int_0^{S_{cut}} S^2 \frac{dN}{dS}\, dS \tag{55}$$

$$b_{\ell_1\ell_2\ell_3}^{RAD} = \int_0^{S_{cut}} S^3 \frac{dN}{dS}\, dS \tag{56}$$

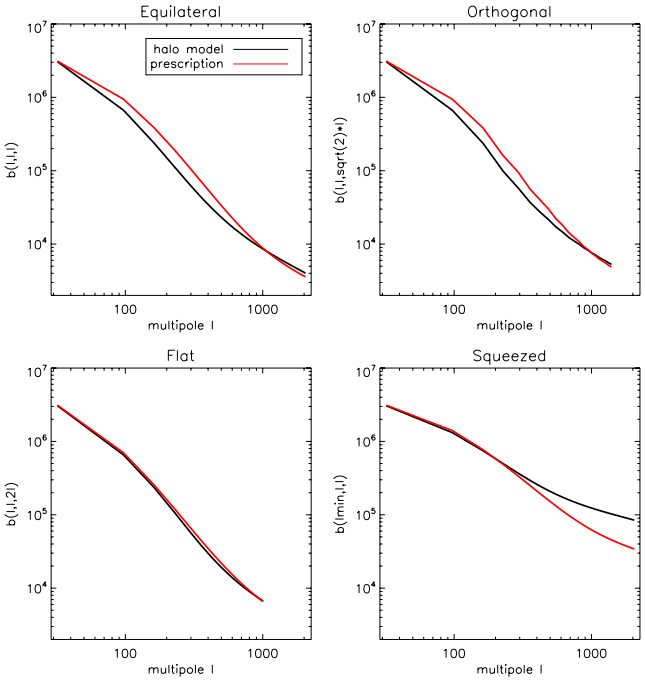

Figure 18. CIB bispectrum computed with the halo model (black line) and with the prescription (red line) at 350 μm.

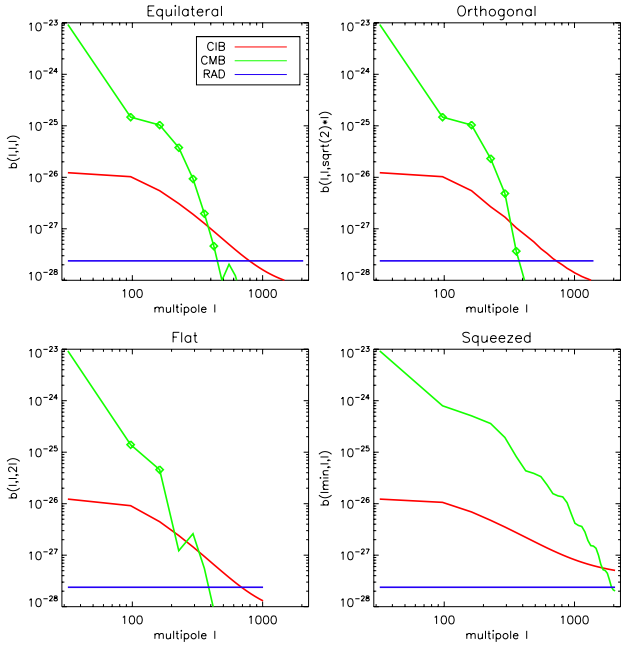

Figure 19. CIB (red line), radio (blue line) and CMB (green line) bispectra at 220 GHz in dimensionless units $\Delta T/T$. The latter is for $f_{\rm NL} = 1$, and we plot its absolute value as it is mostly negative, diamond indicate positive points.

in Jy^2 sr^{-1} and Jy^3 sr^{-2}, respectively, with $\frac{{\rm d}N}{{\rm d}S}$ in gal Jy^{-1} sr^{-1} and $S_{\rm cut}$ the flux cut. The radio bispectrum is hence flat, with neither scale nor geometrical dependence.

In the following, we use number counts from Tucci et al. (2011) and flux cuts from Planck Collaboration (2011a).

In the standard paradigm, inflation generates close to Gaussian perturbations which evolve to Gaussian-distributed temperature fluctuations of the CMB. In the last decade, interest has increased for the search of CMB NG (e.g. Komatsu et al. 2011), as it would be a signature of non-standard inflation (violating any of the following assumption: slow-roll single-field inflation with standard kinetic term and Bunch–Davies initial condition, see Bartolo et al. 2004, for a review), or any physical process generating the primordial perturbations, or of non-linear evolution (e.g. Pitrou, Uzan & Bernardeau 2010). With the recent Planck measurements, the CMB appears to be very close to a Gaussian field (Planck Collaboration XXIV 2013).

Among the many primordial NG shapes, the most studied is the 'local' type NG parametrized by a factor $f_{\rm NL}$, where the Bardeen potential takes the form (Komatsu & Spergel 2001)

$$\Phi(\boldsymbol{x}) = \Phi_{\rm G}(\boldsymbol{x}) + f_{\rm NL} \cdot (\Phi_{\rm G}(\boldsymbol{x})^2 - \langle \Phi_{\rm G}(\boldsymbol{x})^2 \rangle) \tag{57}$$

with $\Phi_{\rm G}$ being the Gaussian part of the potential. This primordial NG generates a CMB bispectrum of the form (Komatsu, Spergel & Wandelt 2005)

$$b^{\rm CMB}_{\ell_1\ell_2\ell_3} = f_{\rm NL} \int r^2 {\rm d}r \; \alpha_{\ell_1}(r)\, \beta_{\ell_2}(r)\, \beta_{\ell_3}(r) + {\rm perm.}, \tag{58}$$

with an integral along the line of sight, and filters

$$\alpha_\ell(r) = \frac{2}{\pi} \int k^2 {\rm d}k \; g_{{\rm T},\ell}(k)\, j_\ell(kr) \tag{59}$$

$$\beta_\ell(r) = \frac{2}{\pi} \int k^2 {\rm d}k \; P(k)\, g_{{\rm T},\ell}(k)\, j_\ell(kr) \tag{60}$$

where $g_{{\rm T},\ell}$ is the radiation transfer function, which can be computed with a Boltzmann code,[5] and $P(k) \propto k^{n_s - 4}$ is the primordial power spectrum, with a spectral index n_s.

The CMB physics, e.g. acoustic peaks and Silk damping, is encoded into its bispectrum thanks to the radiation transfer function.

We show, in Fig. 19, the bispectra of radio sources and the CMB for $f_{\rm NL} = 1$ together with the CIB bispectrum derived from the halo model, in units of relative temperature elevation $\Delta T/T$, at 1380 μm = 220 GHz. At this frequency, radio and IR point sources have comparable contributions, whereas radio sources dominate at lower frequencies and IR sources dominate at higher frequencies.

We see that the CMB dominates on large angular scales, but plummets at high multipoles and becomes negligible for $\ell \geq 500$, except in the squeezed configuration where the CMB dominates at all scales. Indeed for local type NG, the CMB bispectrum peaks strongly in the squeezed limit (Bucher, van Tent & Carvalho 2010). The CIB bispectrum also peaks on large angular scales, albeit less strongly than the CMB. It dominates the radio bispectrum up to $\ell \sim 700-800$ and becomes negligible afterwards except in the squeezed limit where it dominates the radio over the whole multipole range. Based on this simple comparison, it seems that detecting the CIB bispectrum can be possible above 220 GHz. At 220 GHz, if most of the CMB can be removed by a component separation method (which estimates a CMB map through multifrequency analysis, see Delabrouille & Cardoso 2007, for a review), the detection of the CIB bispectrum is possible at $\ell \leq 700$. Furthermore, the application of a lower flux cut would lower the level of the radio bispectrum and

[5] We used CAMB (Lewis, Challinor & Lasenby 2000).

uncover the CIB one. For instance, the CIB bispectrum has been detected in the South Pole Telescope (SPT) data in a multifrequency analysis using 95, 150 and 220 GHz channels (Crawford et al. 2013), and in the Planck data with a signal-to-noise ratio S/N = 5.8 at 217 GHz (Planck Collaboration XXX 2013).

7 CONCLUSION

We have presented a framework which allows us to predict galaxy clustering at all orders with the halo model including shot noise consistently. We have developed a new diagrammatic method which allows a clear representation and understanding of the different terms involved in the computation of the 3D galaxy polyspectrum. It further allows us to avoid cumbersome computation at high orders by replacing them with diagram drawings. This diagrammatic framework is adaptable to different galaxy-tracing signals and we apply it to the CIB. The latter being integrated over a large range of redshift, we show how the polyspectrum of the CIB anisotropies is projected on the celestial sphere. We further show how to account for the particular case of the shot noise terms.

This framework allows us to compute the CIB angular bispectra at any frequency. We have investigated how the different terms of the resulting CIB bispectrum depend on the scale and on the triangle configuration. We recover that the total bispectrum peaks in the squeezed limit as it is also the case for primordial NG of the local type. We discuss how the terms of the CIB bispectrum vary with the HOD parameters. We show that they vary similarly with respect to the different HOD parameters indicating degeneracies. Furthermore, we show that the bispectrum is much more sensitive to the variation of these parameters than the power spectrum.

We explore the halo mass contribution to each term of the 3D galaxy bispectrum, recovering that the 1-halo term gives more weight to massive haloes compared to the 2- and 3-halo terms. The halo mass contribution of the angular CIB bispectrum depends on the specific galaxy evolution model which is examined in Paper II.

Our predictions are finally compared to a previously proposed empirical prescription and to the bispectrum of radio galaxies and that of the CMB assuming a local-type primordial NG. First, we find an overall agreement with the prescription, although the halo model is needed for an accurate description of the bispectrum, in particular, in the squeezed configuration. Secondly, we show that the detection of the CIB bispectrum is possible at frequencies above 220 GHz, where the CIB bispectrum is contaminated; this detection has indeed been performed by SPT and Planck recent results.

This physically based model opens up the possibility to use, in the future, information present in NG measurement to constrain CIB models so as to extract a maximum of information of present and future surveys.

ACKNOWLEDGEMENTS

The authors wish to acknowledge O. Doré, G. P. Holder, G. Lagache, C. Porciani, S. Prunet, C. Schimd, E. Sefusatti, J. Tinker and A. Wetzel for useful discussions. We acknowledge M. Béthermin for computation of the redshift cut-off discussed in Section 4.4 and M. Langer for a thorough reading of the manuscript.

REFERENCES

Ade P. A. R. et al. (Planck Collaboration), 2013, A&A, 550, A133
Amblard A. et al., 2011, Nature, 470, 510
Argüeso F., González-Nuevo J., Toffolatti L., 2003, ApJ, 598, 86
Bartolo N., Komatsu E., Matarrese S., Riotto A., 2004, Phys. Rep., 402, 103
Berlind A. A. et al., 2003, ApJ, 593, 1
Bernardeau F., Colombi S., Gaztañaga E., Scoccimarro R., 2002, Phys. Rep., 367, 1
Béthermin M., Dole H., Lagache G., Le Borgne D., Penin A., 2011, A&A, 529, A4
Bucher M., van Tent B., Carvalho C. S., 2010, MNRAS, 407, 2193
Cooray A., Sheth R., 2002, Phys. Rep., 372, 1
Cooray A. et al., 2010, A&A, 518, L22
Coupon J. et al., 2012, A&A, 542, A5
Crawford T. M. et al., 2013, ApJ, arXiv:e-prints
Davis M., Peebles P. J. E., 1983, ApJ, 267, 465
De Lucia G., Blaizot J., 2007, MNRAS, 375, 2
Delabrouille J., Cardoso J., 2007, arXiv:e-prints
Fry J. N., 1984, ApJ, 279, 499
Fry J. N., Gaztanaga E., 1993, ApJ, 413, 447
Gil-Marín H., Wagner C., Fragkoudi F., Jimenez R., Verde L., 2012, J. Cosmol. Astropart. Phys., 2012, 047
González-Nuevo J., Toffolatti L., Argüeso F., 2005, ApJ, 621, 1
Jullo E. et al., 2012, ApJ, 750, 37
Knox L., Cooray A., Eisenstein D., Haiman Z., 2001, ApJ, 550, 7
Komatsu E., Spergel D. N., 2001, Phys. Rev., 63, 063002
Komatsu E. et al., 2003, ApJS, 148, 119
Komatsu E., Spergel D. N., Wandelt B. D., 2005, ApJ, 634, 14
Komatsu E. et al., 2011, ApJS, 192, 18
Kravtsov A. V., Berlind A. A., Wechsler R. H., Klypin A. A., Gottlöber S., Allgood B., Primack J. R., 2004, ApJ, 609, 35
Lacasa F., Aghanim N., 2012, arXiv:e-prints
Lacasa F., Aghanim N., Kunz M., Frommert M., 2012, MNRAS, 421, 1982
Lagache G., Puget J.-L., Dole H., 2005, ARA&A, 43, 727
Lagache G., Bavouzet N., Fernandez-Conde N., Ponthieu N., Rodet T., Dole H., Miville-Deschênes M.-A., Puget J.-L., 2007, ApJ, 665, L89
Le Fèvre O. et al., 2005, A&A, 439, 877
Lewis A., Challinor A., Lasenby A., 2000, ApJ, 538, 473
Madgwick D. S. et al., 2003, MNRAS, 344, 847
Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493
Neyman J., Scott E. L., 1952, ApJ, 116, 144
Peebles P. J. E., 1980, The Large-Scale Structure of the Universe. Princeton Univ. Press, Princeton, NJ
Pénin A., Doré O., Lagache G., Béthermin M., 2012, A&A, 537, A137
Pénin A., Lacasa F., Aghanim N., 2014, MNRAS, 439, 143 (Paper II)
Pitrou C., Uzan J.-P., Bernardeau F., 2010, J. Cosmol. Astropart. Phys., 7, 3
Planck Collaboration, 2011a, A&A, 536, A7
Planck Collaboration, 2011b, A&A, 536, A18
Planck Collaboration XXIV 2013, A&A, submitted
Planck Collaboration XXX 2013, A&A, submitted
Puget J.-L., Abergel A., Bernard J.-P., Boulanger F., Burton W. B., Desert F.-X., Hartmann D., 1996, A&A, 308, 5
Sehgal N., Bode P., Das S., Hernandez-Monteagudo C., Huffenberger K., Lin Y.-T., Ostriker J. P., Trac H., 2010, ApJ, 709, 920
Sheth R. K., Tormen G., 1999, MNRAS, 308, 119
Smith R. E., Watts P. I. R., Sheth R. K., 2006, MNRAS, 365, 214
Smith R. E., Sheth R. K., Scoccimarro R., 2008, Phys. Rev., 78, 023523
Thacker C. et al., 2013, ApJ, 768, 58
Tinker J. L., Wetzel A. R., 2010, ApJ, 719, 88
Tinker J. L., Wechsler R. H., Zheng Z., 2010, ApJ, 709, 67
Toffolatti L., Argueso Gomez F., de Zotti G., Mazzei P., Franceschini A., Danese L., Burigana C., 1998, MNRAS, 297, 117
Toffolatti L., De Zotti G., Argüeso F., Burigana C., 1999, in de Oliveira-Costa A., Tegmark M., eds, ASP Conf. Ser. Vol. 181, Microwave Foregrounds. Astron. Soc. Pac., San Francisco, p. 153
Tucci M., Toffolatti L., de Zotti G., Martínez-González E., 2011, A&A, 533, A57
Viero M. P. et al., 2009, ApJ, 707, 1766
Wang Y., Yang X., Mo H. J., van den Bosch F. C., Chu Y., 2004, MNRAS, 353, 287
Xia J.-Q., Negrello M., Lapi A., De Zotti G., Danese L., Viel M., 2012, MNRAS, 422, 1324

Zehavi I. et al. (SDSS Collaboration), 2004, ApJ, 608, 16
Zheng Z. et al., 2005, ApJ, 633, 791

APPENDIX A: DERIVATION OF THE GALAXY POWER SPECTRUM EQUATIONS

Using equation (3), the computation of the 2-point correlation function of $\delta_{\rm gal}$

$$\zeta_{\rm 2pt}(\boldsymbol{x}_1-\boldsymbol{x}_2)=\frac{\langle n_{\rm gal}(\boldsymbol{x}_1)\,n_{\rm gal}(\boldsymbol{x}_2)\rangle-\overline{n}^2_{\rm gal}}{\overline{n}^2_{\rm gal}}=\langle\delta_{\rm gal}(\boldsymbol{x}_1)\,\delta_{\rm gal}(\boldsymbol{x}_2)\rangle \tag{A1}$$

yields a double sum over halo and galaxy indexes $\sum_{i_1,i_2}\sum_{j_1,j_2}$ which can be split into three terms: $\sum_{i_1\neq i_2}\sum_{j_1\neq j_2}$ (2-halo term), $\sum_{i_1=i_2}\sum_{j_1\neq j_2}$ (1-halo 2-galaxy term) and $\sum_{i_1=i_2}\sum_{j_1=j_2}$ (1-halo 1-galaxy term). So that

$$\zeta_{\rm 2pt}(\boldsymbol{x}_1-\boldsymbol{x}_2)=\zeta^{\rm 2h}_{\rm 2pt}(\boldsymbol{x}_1-\boldsymbol{x}_2)+\zeta^{\rm 1h-2g}_{\rm 2pt}(\boldsymbol{x}_1-\boldsymbol{x}_2)+\zeta^{\rm 1h-1g}_{\rm 2pt}(\boldsymbol{x}_1-\boldsymbol{x}_2) \tag{A2}$$

with computations giving

$$\begin{aligned}\zeta^{\rm 2h}_{\rm 2pt}(\boldsymbol{x}_1-\boldsymbol{x}_2)=&\int{\rm d}M_1\,\frac{\langle N_{\rm gal}(M_1)\rangle\ \frac{{\rm d}n_{\rm h}}{{\rm d}M}\big|_{M_1}}{\overline{n}_{\rm gal}}\\&\times\int{\rm d}M_2\,\frac{\langle N_{\rm gal}(M_2)\rangle\ \frac{{\rm d}n_{\rm h}}{{\rm d}M}\big|_{M_2}}{\overline{n}_{\rm gal}}\\&\times\int{\rm d}^3\boldsymbol{x}'_{12}\,u(\boldsymbol{x}_1-\boldsymbol{x}'_1|M_1)\,u(\boldsymbol{x}_2-\boldsymbol{x}'_2|M_2)\\&\times\zeta^{\rm halo}_{\rm 2pt}(\boldsymbol{x}'_1-\boldsymbol{x}'_2|M_1,M_2),\end{aligned} \tag{A3}$$

$$\begin{aligned}\zeta^{\rm 1h-2g}_{\rm 2pt}(\boldsymbol{x}_1-\boldsymbol{x}_2)=&\int{\rm d}M\,\frac{{\rm d}n_{\rm h}}{{\rm d}M}\,\frac{\langle N_{\rm gal}(N_{\rm gal}-1)\rangle}{\overline{n}^2_{\rm gal}}\\&\times\int{\rm d}^3\boldsymbol{x}_{\rm h}\,u(\boldsymbol{x}_1-\boldsymbol{x}_{\rm h}|M)\,u(\boldsymbol{x}_2-\boldsymbol{x}_{\rm h}|M),\end{aligned} \tag{A4}$$

$$\zeta^{\rm 1h-1g}_{\rm 2pt}(\boldsymbol{x}_1-\boldsymbol{x}_2)=\frac{\delta^3(\boldsymbol{x}_1-\boldsymbol{x}_2)}{\overline{n}_{\rm gal}}, \tag{A5}$$

where $\frac{{\rm d}n_{\rm h}}{{\rm d}M}$ is the number of haloes with mass M per comoving volume a.k.a. the halo mass function, $u(\boldsymbol{x}|M)$ is the halo profile (with integral normalized to unity) and $\zeta^{\rm halo}_{\rm 2pt}(\boldsymbol{x}|M_1,M_2)$ is the halo correlation function conditioned to masses M_1 and M_2. In this paper, we use the Sheth & Tormen mass function (Sheth & Tormen 1999) and the associated bias functions, as it is the most recent one for which the second order bias is available.

At tree level, the halo correlation function takes the form (see Cooray & Sheth 2002, and Appendix C)

$$\zeta^{\rm halo}_{\rm 2pt}(\boldsymbol{x}|M_1,M_2)=b_1(M_1)\,b_1(M_2)\,\zeta^{\rm lin}_{\rm 2pt}(\boldsymbol{x}), \tag{A6}$$

where $b_1(M)$ is the first order halo bias and $\zeta^{\rm lin}$ is the dark matter correlation function at linear/first order in perturbation theory.

The correlation functions defined by equations (A3) and (A4) involve convolutions in real space, which become multiplications in Fourier space. Hence the galaxy power spectrum – defined by $\zeta_{\rm 2pt}(\boldsymbol{x})=\int\frac{{\rm d}^3\boldsymbol{k}}{(2\pi)^3}P(k)\,e^{i\boldsymbol{k}\cdot\boldsymbol{x}}$ – becomes

$$P_{\rm gal}(k)=P^{\rm 1h}_{\rm gal}(k)+P^{\rm 2h}_{\rm gal}(k)+P^{\rm shot}_{\rm gal}(k), \tag{A7}$$

where the shot noise contribution (corresponding to the 1h1g term) is examined in more detail in Section 4.3.

The 1-halo contribution is

$$P^{\rm 1h}_{\rm gal}(k)=\int{\rm d}M\,\frac{{\rm d}n_{\rm h}}{{\rm d}M}\,\frac{\langle N_{\rm gal}(M)(N_{\rm gal}(M)-1)\rangle}{\overline{n}^2_{\rm gal}}\,|u(k|M)|^2\,, \tag{A8}$$

And the 2-halo contribution

$$\begin{aligned}P^{\rm 2h}_{\rm gal}(k)=&\int{\rm d}M_1\,\frac{\langle N_{\rm gal}(M_1)\rangle\ \frac{{\rm d}n_{\rm h}}{{\rm d}M}\big|_{M_1}}{\overline{n}_{\rm gal}}\,u(k|M_1)\\&\times\int{\rm d}M_2\,\frac{\langle N_{\rm gal}(M_2)\rangle\ \frac{{\rm d}n_{\rm h}}{{\rm d}M}\big|_{M_2}}{\overline{n}_{\rm gal}}\,u(k|M_2)\,P_{\rm halo}(k|M_1,M_2)\\=&P_{\rm lin}(k)\left(\int{\rm d}M\,\frac{\langle N_{\rm gal}(M)\rangle\,\frac{{\rm d}n_{\rm h}}{{\rm d}M}}{\overline{n}_{\rm gal}}\,b_1(M)\,u(k|M)\right)^2\end{aligned} \tag{A9}$$

where as precedently, all redshift dependence are implicit to simplify notations.

APPENDIX B: DERIVATION OF THE GALAXY BISPECTRUM EQUATIONS

Similarly to the 2-point correlation function, using equation (3), the computation of the 3-point correlation function of $\delta_{\rm gal}$ can be split into six terms: $\sum_{i_1\neq i_2\neq i_3}\sum_{j_1\neq j_2\neq j_3}$ (3-halo term), $\sum_{i_1=i_2\neq i_3}\sum_{j_1\neq j_2\neq j_3}+$ perm. (2-halo 3-galaxy term), $\sum_{i_1=i_2\neq i_3}\sum_{j_1=j_2\neq j_3}+$ perm. (2-halo 2-galaxy term), $\sum_{i_1=i_2=i_3}\sum_{j_1\neq j_2\neq j_3}+$ perm. (1-halo 3-galaxy term), $\sum_{i_1=i_2=i_3}\sum_{j_1=j_2\neq j_3}+$ perm. (1-halo 2-galaxy term) and finally $\sum_{i_1=i_2=i_3}\sum_{j_1=j_2=j_3}+$ perm. (1-halo 1-galaxy term). So that

$$\begin{aligned}\zeta_{\rm 3pt}(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3)=&\zeta^{\rm 3h}_{\rm 3pt}(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3)+\zeta^{\rm 2h-3g}_{\rm 3pt}(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3)\\&+\zeta^{\rm 2h-2g}_{\rm 3pt}(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3)+\zeta^{\rm 1h-3g}_{\rm 3pt}(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3)\\&+\zeta^{\rm 1h-2g}_{\rm 3pt}(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3)+\zeta^{\rm 1h-1g}_{\rm 3pt}(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3),\end{aligned} \tag{B1}$$

with computations giving

$$\begin{aligned}\zeta^{\rm 3h}_{\rm 3pt}(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3)=&\int{\rm d}M_{123}\left[\frac{\langle N_{\rm gal}(M_i)\rangle\ \frac{{\rm d}n_{\rm h}}{{\rm d}M}\big|_{M_i}}{\overline{n}_{\rm gal}}\right]_{i=123}\\&\times\int{\rm d}^3\boldsymbol{x}'_{123}\left(u(\boldsymbol{x}_i-\boldsymbol{x}'_i|M_i)\right)_{i=123}\\&\times\zeta^{\rm halo}_{\rm 3pt}(\boldsymbol{x}'_{123}|M_{123}),\end{aligned} \tag{B2}$$

$$\begin{aligned}\zeta^{\rm 2h-3g}_{\rm 3pt}(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3)=&\int{\rm d}M_1\,\frac{\langle N_{\rm gal}(M_1)\left(N_{\rm gal}(M_1)-1\right)\rangle\ \frac{{\rm d}n_{\rm h}}{{\rm d}M}\big|_{M_1}}{\overline{n}^2_{\rm gal}}\\&\times\int{\rm d}M_3\,\frac{\langle N_{\rm gal}(M_3)\rangle\ \frac{{\rm d}n_{\rm h}}{{\rm d}M}\big|_{M_3}}{\overline{n}_{\rm gal}}\\&\times\int{\rm d}^3\boldsymbol{x}'_{13}\,u(\boldsymbol{x}_1-\boldsymbol{x}'_1|M_1)\,u(\boldsymbol{x}_2-\boldsymbol{x}'_1|M_1)\\&\times u(\boldsymbol{x}_3-\boldsymbol{x}'_3|M_3)\,\zeta^{\rm halo}_{\rm 2pt}(\boldsymbol{x}'_1-\boldsymbol{x}'_3|M_1M_3)\\&+{\rm perm.},\end{aligned} \tag{B3}$$

$$\zeta^{\rm 2h-2g}_{\rm 3pt}(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3)=\frac{\zeta^{\rm 2h-2g}_{\rm 2pt}(\boldsymbol{x}_1-\boldsymbol{x}_3)\,\delta^{(3)}(\boldsymbol{x}_1-\boldsymbol{x}_2)}{\overline{n}_{\rm gal}}+{\rm perm.} \tag{B4}$$

$$\zeta_{3\rm pt}^{\rm 1h-3g}(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3) = \int {\rm d}M\, \frac{\langle N_{\rm gal}\left(N_{\rm gal}-1\right)\left(N_{\rm gal}-2\right)\rangle}{\overline{n}_{\rm gal}^3}\frac{{\rm d}n_{\rm h}}{{\rm d}M} \times \int {\rm d}^3\boldsymbol{x}'\, u(\boldsymbol{x}_1-\boldsymbol{x}'|M)u(\boldsymbol{x}_2-\boldsymbol{x}'|M) \times u(\boldsymbol{x}_3-\boldsymbol{x}'|M), \quad \text{(B5)}$$

$$\zeta_{3\rm pt}^{\rm 1h-2g}(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3) = \frac{\zeta_{2\rm pt}^{\rm 1h-2g}(\boldsymbol{x}_1-\boldsymbol{x}_3)\,\delta^{(3)}(\boldsymbol{x}_1-\boldsymbol{x}_2)}{\overline{n}_{\rm gal}} + \text{perm.}, \quad \text{(B6)}$$

$$\zeta_{3\rm pt}^{\rm 1h-1g}(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3) = \frac{\delta^{(6)}(\boldsymbol{x}_1=\boldsymbol{x}_2=\boldsymbol{x}_3)}{\overline{n}_{\rm gal}^2}, \quad \text{(B7)}$$

where $\zeta_{3\rm pt}^{\rm halo}(\boldsymbol{x}_{123}|M_{123})$ is the halo 3-point correlation function conditioned to masses M_1 M_2 and M_3, and for which we take (see Appendix C)

$$\zeta_{3\rm pt}^{\rm halo}(\boldsymbol{x}_{123}|M_{123}) = b_1(M_1)\,b_1(M_2)\,b_1(M_3)\,\zeta_{3\rm pt}^{\rm DM}(\boldsymbol{x}_{123}) + b_1(M_1)\,b_1(M_2)\,b_2(M_3)\,\zeta_{2\rm pt}^{\rm lin}(\boldsymbol{x}_1-\boldsymbol{x}_3) \times \zeta_{2\rm pt}^{\rm lin}(\boldsymbol{x}_2-\boldsymbol{x}_3) + \text{perm.}, \quad \text{(B8)}$$

where b_2 is the second order halo bias.

The 3D galaxy bispectrum can be computed from equations (B2) to (B5) through Fourier transform

$$\zeta_{3\rm pt}(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3) = \int \frac{{\rm d}^3\boldsymbol{k}_{123}}{(2\pi)^9}\, B(k_1,k_2,k_3)\,(2\pi)^3\delta(\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3) \times e^{i(\boldsymbol{k}_1\cdot\boldsymbol{x}_1+\boldsymbol{k}_2\cdot\boldsymbol{x}_2+\boldsymbol{k}_3\cdot\boldsymbol{x}_3)}$$

giving

$$B_{\rm gal}(k_1,k_2,k_3) = B_{\rm gal}^{\rm 1h}(k_1,k_2,k_3) + B_{\rm gal}^{\rm 2h}(k_1,k_2,k_3) + B_{\rm gal}^{\rm 3h}(k_1,k_2,k_3) + B_{\rm gal}^{\rm shot2g}(k_1,k_2,k_3) + B_{\rm gal}^{\rm shot1g}(k_1,k_2,k_3), \quad \text{(B9)}$$

where we have the 1-halo term

$$B_{\rm gal}^{\rm 1h}(k_1,k_2,k_3) = \int {\rm d}M\, \frac{\langle N_{\rm gal}(N_{\rm gal}-1)(N_{\rm gal}-2)\rangle}{\overline{n}_{\rm gal}^3}\frac{{\rm d}n_{\rm h}}{{\rm d}M} \times u(\boldsymbol{k}_1|M)\,u(\boldsymbol{k}_2|M)\,u(\boldsymbol{k}_3|M). \quad \text{(B10)}$$

The 2-halo term

$$B_{\rm gal}^{\rm 2h}(k_1,k_2,k_3) = \int {\rm d}M_1 \frac{\langle N_{\rm gal}(N_{\rm gal}-1)\rangle\ \frac{{\rm d}n_{\rm h}}{{\rm d}M}\big|_{M_1}}{\overline{n}_{\rm gal}^2} \times \int {\rm d}M_3 \frac{\langle N_{\rm gal}(M_3)\rangle\ \frac{{\rm d}n_{\rm h}}{{\rm d}M}\big|_{M_3}}{\overline{n}_{\rm gal}} \times u(\boldsymbol{k}_1|M_1)u(\boldsymbol{k}_2|M_1)u(\boldsymbol{k}_3|M_3) \times P_{\rm halo}(k_3|M_1,M_3) + \text{perm.} \quad \text{(B11)}$$

$$= P_{\rm lin}(k_3)\int {\rm d}M_1 \frac{\langle N_{\rm gal}(N_{\rm gal}-1)\rangle\ \frac{{\rm d}n_{\rm h}}{{\rm d}M}\big|_{M_1}}{\overline{n}_{\rm gal}^2} \times b_1(M_1)\,u(\boldsymbol{k}_1|M_1)u(\boldsymbol{k}_2|M_1) \times \int {\rm d}M_3 \frac{\langle N_{\rm gal}(M_3)\rangle\ \frac{{\rm d}n_{\rm h}}{{\rm d}M}\big|_{M_3}}{\overline{n}_{\rm gal}} b_1(M_3)\,u(\boldsymbol{k}_3|M_3) + \text{perm.} \quad \text{(B12)}$$

And the 3-halo term

$$B_{\rm gal}^{\rm 3h}(k_1,k_2,k_3) = \int {\rm d}M_{123} \left(\frac{N_{\rm gal}(M_i)\ \frac{{\rm d}n_{\rm h}}{{\rm d}M}\big|_{M_i}}{\overline{n}_{\rm gal}}\right)_{i=123} |u(\boldsymbol{k}_i|M_i)|_{i=123} \times B_{\rm halo}(k_1,k_2,k_3|M_1,M_2,M_3) \quad \text{(B13)}$$

$$= B_{\rm lin}(k_{123}) \int {\rm d}M_{123} \left(\frac{N_{\rm gal}(M_i)\ \frac{{\rm d}n_{\rm h}}{{\rm d}M}\big|_{M_i}}{\overline{n}_{\rm gal}} \times b_1(M_i)\,u(\boldsymbol{k}_i|M_i)\right)_{i=123} + \int {\rm d}M_{123} \left(\frac{N_{\rm gal}(M_i)\ \frac{{\rm d}n_{\rm h}}{{\rm d}M}\big|_{M_i}}{\overline{n}_{\rm gal}}\,|u(\boldsymbol{k}_i|M_i)|\right)_{i=123} \times b_1(M_1)\,b_1(M_2)\,b_2(M_3)\,P_{\rm lin}(k_1)P_{\rm lin}(k_2) + \text{perm.}, \quad \text{(B14)}$$

with

$$B_{\rm DM}(k_1,k_2,k_3) = \underbrace{2\,f_{\rm NL}\left(\left(\frac{k_3\,k_H}{k_1\,k_2}\right)^2 P_{\rm lin}(k_1)\,P_{\rm lin}(k_2) + \text{perm.}\right)}_{\text{primordial NG}} + \underbrace{2\,F^s(\boldsymbol{k_1},\boldsymbol{k_2})\,P_{\rm lin}(k_1)\,P_{\rm lin}(k_2) + \text{perm.}}_{\text{gravity non}-\text{linearity at 2PT}} \quad \text{(B15)}$$

with

$$k_H^2 = 4\pi G\,\overline{\rho}\,a^2(t), \quad \text{(B16)}$$

which stems from the Poisson equation linking density contrast to Bardeen potential involved in the definition of $f_{\rm NL}$ (equation 57) and

$$F^s(\boldsymbol{k_i},\boldsymbol{k_j}) = \frac{5}{7} + \frac{1}{2}\cos(\theta_{ij})\left(\frac{k_i}{k_j}+\frac{k_j}{k_i}\right) + \frac{2}{7}\cos^2(\theta_{ij}), \quad \text{(B17)}$$

which stems from non-linear evolution at second order in perturbation theory (Fry 1984; Gil-Marín et al. 2012).

At the scales of interest in this paper $k_H \ll k_{1,2,3}$ so that we can neglect the primordial NG term (e.g. compared to the 2PT term in equation B15).

The equations above can be somewhat simplified if we introduce some notations (where we reintroduced the redshift dependence to show where it intervenes):

$$\mathcal{F}_1(k,z) = \int {\rm d}M \frac{\langle N_{\rm gal}(M)\rangle}{\overline{n}_{\rm gal}(z)}\frac{{\rm d}n_{\rm h}}{{\rm d}M}(M,z)\,b_1(M,z)\,u(k|M,z), \quad \text{(B18)}$$

$$\mathcal{F}_2(k,z) = \int {\rm d}M \frac{\langle N_{\rm gal}(M)\rangle}{\overline{n}_{\rm gal}(z)}\frac{{\rm d}n_{\rm h}}{{\rm d}M}(M,z)\,b_2(M,z)\,u(k|M,z), \quad \text{(B19)}$$

$$\mathcal{G}_1(k_1,k_2,z) = \int {\rm d}M \frac{\langle N_{\rm gal}(N_{\rm gal}-1)\rangle}{\overline{n}_{\rm gal}(z)^2}\frac{{\rm d}n_{\rm h}}{{\rm d}M}(M,z) \times b_1(M,z)\,u(k_1|M,z)\,u(k_2|M,z) \quad \text{(B20)}$$

and the F^s kernel can also be computed through the formula

$$F^s(\boldsymbol{k_\alpha}, \boldsymbol{k_\beta}) = \frac{2k_\gamma^4 - 5\left(k_\alpha^4 + k_\beta^4\right) + 3k_\gamma^2\left(k_\alpha^2 + k_\beta^2\right) + 10k_\alpha^2 k_\beta^2}{28k_\alpha^2 k_\beta^2}, \tag{B21}$$

where γ is the third index.

With these notations the 2-halo term takes the form

$$\begin{aligned} B_{\rm gal}^{\rm 2h}(k_1, k_2, k_3, z) = {} & \mathcal{G}_1(k_1, k_2, z)\, P_{\rm lin}(k_3, z)\, \mathcal{F}_1(k_3, z) \\ & + \mathcal{G}_1(k_1, k_3, z)\, P_{\rm lin}(k_2, z)\, \mathcal{F}_1(k_2, z) \\ & + \mathcal{G}_1(k_2, k_3, z)\, P_{\rm lin}(k_1, z)\, \mathcal{F}_1(k_1, z) \end{aligned} \tag{B22}$$

and the 3-halo term

$$\begin{aligned} B_{\rm gal}^{\rm 3h}(k_1, k_2, k_3, z) = {} & \mathcal{F}_1(k_1, z)\, \mathcal{F}_1(k_2, z)\, \mathcal{F}_1(k_3, z) \\ & \times \left[F^s(\boldsymbol{k_1}, \boldsymbol{k_2})\, P_{\rm lin}(k_1, z)\, P_{\rm lin}(k_2, z) + \text{perm.}\right] \\ & + \mathcal{F}_1(k_1, z)\, \mathcal{F}_1(k_2, z)\, \mathcal{F}_2(k_3, z) \\ & \times P_{\rm lin}(k_1, z)\, P_{\rm lin}(k_2, z) + \text{perm.} \end{aligned} \tag{B23}$$

Last, the shot noise terms are

$$B_{\rm gal}^{\rm 1h-1g}(k_1, k_2, k_3, z) = \frac{1}{\overline{n}_{\rm gal}^2(z)} \tag{B24}$$

$$B_{\rm gal}^{\rm 1h-2g}(k_1, k_2, k_3, z) = \frac{P_{\rm gal}^{\rm 1h-2g}(k_1) + P_{\rm gal}^{\rm 1h-2g}(k_2) + P_{\rm gal}^{\rm 1h-2g}(k_3)}{\overline{n}_{\rm gal}(z)} \tag{B25}$$

$$B_{\rm gal}^{\rm 2h-2g}(k_1, k_2, k_3, z) = \frac{P_{\rm gal}^{\rm 2h-2g}(k_1) + P_{\rm gal}^{\rm 2h-2g}(k_2) + P_{\rm gal}^{\rm 2h-2g}(k_3)}{\overline{n}_{\rm gal}(z)}. \tag{B26}$$

APPENDIX C: HALO CORRELATION FUNCTIONS

We assume that the halo density field follows the local bias scheme (Fry & Gaztanaga 1993)

$$\delta_{\rm h}(\boldsymbol{x}|M) = \sum_{n=1}^{+\infty} \frac{b_n(M)}{n!}\, \delta_{\rm DM}(\boldsymbol{x})^n \tag{C1}$$

where $\delta_{\rm DM}$ is the dark matter density field predicted through perturbation theory and $b_n(m)$ is the nth order bias

$$b_n(M) = \frac{1}{f(\nu)} \frac{\partial^n f(\nu)}{\partial \delta^n}. \tag{C2}$$

Because of the smallness of $\delta_{\rm DM}$, it is sufficient to develop the computation of halo correlation functions to tree level.

Hence the 2-point correlation function conditioned to mass M_1 and M_2 is

$$\begin{aligned} \zeta_{\rm 2pt}^{\rm hh}(\boldsymbol{x}_1 - \boldsymbol{x}_2|M_1, M_2) \equiv {} & \langle \delta_{\rm h}(\boldsymbol{x}_1|M_1)\, \delta_{\rm h}(\boldsymbol{x}_2|M_2)\rangle \\ = {} & b_1(M_1)\, b_1(M_2)\, \langle \delta_{\rm DM}(\boldsymbol{x}_1)\, \delta_{\rm DM}(\boldsymbol{x}_2)\rangle \\ & + \boldsymbol{o}\left(\delta^2\right) \\ \simeq {} & b_1(M_1)\, b_1(M_2)\, \zeta_{\rm 2pt}^{\rm lin}(\boldsymbol{x}_1 - \boldsymbol{x}_2). \end{aligned} \tag{C3}$$

Going to Fourier space gives the power spectrum

$$P_{hh}(k) = b_1(M_1)\, b_1(M_2)\, P_{\rm lin}(k). \tag{C4}$$

At tree level the 3-point correlation function conditioned to mass M_1 M_2 and M_3 is

$$\begin{aligned} \zeta_{\rm 3pt}^{\rm hhh}(\boldsymbol{x}_{123}|M_{123}) = {} & \left\langle \left(b_1(M_1)\delta_{\rm DM}(\boldsymbol{x}_1) + \frac{b_2(M_1)}{2!}\delta_{\rm DM}(\boldsymbol{x}_1)^2\right) \right. \\ & \times \left(b_1(M_2)\delta_{\rm DM}(\boldsymbol{x}_2) + \frac{b_2(M_2)}{2!}\delta_{\rm DM}(\boldsymbol{x}_2)^2\right) \\ & \left. \times \left(b_1(M_3)\delta_{\rm DM}(\boldsymbol{x}_3) + \frac{b_2(M_3)}{2!}\delta_{\rm DM}(\boldsymbol{x}_3)^2\right)\right\rangle \\ = {} & b_1(M_1)\, b_1(M_2)\, b_1(M_3)\, \zeta_{\rm 3pt}^{\rm DM}(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3) \\ & + \frac{b_2(M_1)}{2!}\, b_1(M_2)\, b_1(M_3) \\ & \times \langle \delta_{\rm lin}(\boldsymbol{x}_1)^2\, \delta_{\rm lin}(\boldsymbol{x}_2)\, \delta_{\rm lin}(\boldsymbol{x}_3)\rangle + 2 \text{ perm.} \end{aligned} \tag{C5}$$

$$\begin{aligned} = {} & b_1(M_1)\, b_1(M_2)\, b_1(M_3)\, \zeta_{\rm 3pt}^{\rm lin}(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3) \\ & + b_2(M_1)\, b_1(M_2)\, b_1(M_3) \\ & \times \left[\frac{\langle \delta_{\rm DM}^2\rangle}{2!}\, \zeta_{\rm 2pt}^{\rm lin}(\boldsymbol{x}_2 - \boldsymbol{x}_3) \right. \\ & \left. + \zeta_{\rm 2pt}^{\rm lin}(\boldsymbol{x}_1 - \boldsymbol{x}_2)\, \zeta_{\rm 2pt}^{\rm lin}(\boldsymbol{x}_1 - \boldsymbol{x}_3)\right] + 2 \text{ perm.}, \end{aligned} \tag{C6}$$

where we expanded the 4-point correlation function (at line C5) through Wick's theorem as the field is close to Gaussian, and where $\zeta_{\rm 3pt}^{\rm DM}(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3)$ contains two contributions: primordial NG which is observationally constrained to be small, and NG generated by the non-linearity of gravity at second order in perturbation theory.

Going to Fourier space gives the bispectrum (for $k_1, k_2, k_3 \neq 0$):

$$\begin{aligned} B_{\rm hhh}(k_{123}|M_{123}) = {} & b_1(M_1)\, b_1(M_2)\, b_1(M_3)\, B_{\rm DM}(k_1, k_2, k_3) \\ & + b_2(M_1)\, b_1(M_2)\, b_1(M_3)\, P_{\rm lin}(k_2)\, P_{\rm lin}(k_3) \\ & + 2 \text{ perm.} \end{aligned} \tag{C7}$$

APPENDIX D: POLYSPECTRA IN FOURIER SPACE AND ON THE SPHERE

Let a_k be the Fourier transform of a random field. For example, this may be the 3D galaxy density field, in which case $a_k = \delta_{\rm gal}(\boldsymbol{k}, z)$. The polyspectrum of order n is then defined via the connected correlation function in Fourier space

$$\langle a_{k_1} \ldots a_{k_n}\rangle_c. \tag{D1}$$

Under the assumption of statistical homogeneity and isotropy, this correlation function vanishes unless $\boldsymbol{k}_1 \ldots \boldsymbol{k}_n$ form a polygon. This polygon may be parametrized by the lengths of its sides $k_1 \ldots k_n$ and by diagonals $k_1^d \ldots k_m^d$ needed to fix the shape via a chosen triangulation. The correlation function of order n thus has $2n - 3$ degrees of freedom in 2D (i.e. $m = n - 3$), as can be seen in Fig. D1 at fourth (trispectrum) and nth order.

Correspondingly, the correlation function of order n has $3n - 6$ degrees of freedom in 3D ($m = 2n - 6$), as can be seen in Fig. D2 at fourth order (trispectrum).

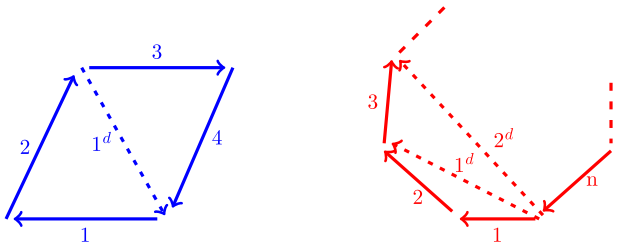

Figure D1. Parametrization of the 2D polyspectrum, with diagonals in dashed lines, at fourth (trispectrum) and nth order.

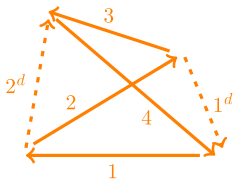

Figure D2. Parametrization of the 3D trispectrum with two diagonal degrees of freedom in dashed lines.

The polyspectrum of order n, $\mathcal{P}^{(n)}(k_1 \dots k_n, k_1^d \dots k_m^d)$, is then defined by

$$\langle a_{k_1} \cdots a_{k_n} \rangle_c = \int \frac{\mathrm{d}^D k_1^d}{(2\pi)^D} \cdots \frac{\mathrm{d}^D k_m^d}{(2\pi)^D} \, \mathcal{P}^{(n)}\left(k_1 \dots k_n, k_1^d \dots k_m^d\right) \times \prod_g (2\pi)^D \, \delta\left(\boldsymbol{k}_1(g) + \boldsymbol{k}_2(g) + \boldsymbol{k}_3(g)\right), \quad \text{(D2)}$$

where D is the dimension of the random field ($D = 3$ for the galaxy distribution), and g indexes the chosen triangulation of the polygon [e.g. for the 2D trispectrum $g = (1, 2)$ and the triangulation is $(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_1^d); (-\boldsymbol{k}_1^d, \boldsymbol{k}_3, \boldsymbol{k}_4)$].

Note that at orders $\geq$4, polyspectra have more degrees of freedom in 3D than in 2D. However, this does not happen at the bispectrum level which does not have diagonal degrees of freedom, as a triangle is flat and can be parametrized solely with its sides.

For some random fields (e.g. Gaussian or white noise), the polyspectra may not depend on diagonal degrees of freedom, such polyspectra are called diagonal independent and can be parametrized solely with the length of the sides $k_1 \dots k_n$. In this case, equation (D2) takes the simpler form

$$\langle a_{k_1} \dots a_{k_n} \rangle_c = (2\pi)^D \, \delta(\boldsymbol{k}_1 + \dots + \boldsymbol{k}_n) \times \mathcal{P}^{(n)}(k_{1\dots n}). \quad \text{(D3)}$$

The case of random fields on the sphere is similar to the 2D Fourier case, and can be defined simply with the replacements

$$\int \frac{\mathrm{d}^2 \boldsymbol{k}^d}{(2\pi)^2} \to \sum_{\ell^d m^d} \quad \text{(D4)}$$

$$(2\pi)^2 \delta(\boldsymbol{k}_1 + \boldsymbol{k}_2 + \boldsymbol{k}_3) \to G_{123}, \quad \text{(D5)}$$

the nth order polyspectrum has $2n - 3$ degrees of freedom and is defined through

$$\langle a_{\ell_1 m_1} \dots a_{\ell_n m_n} \rangle_c = \sum_{\substack{\ell_1^{\mathrm{d}} \dots \ell_{n-3}^{\mathrm{d}} \\ m_1^{\mathrm{d}} \dots \ell_{n-3}^{\mathrm{d}}}} \mathcal{P}^{(n)}\left(\ell_{1\dots n}, \ell_{1\dots(n-3)}^{\mathrm{d}}\right) \times \mathcal{G}_{1\dots n}\left(\ell_{1\dots(n-3)}^{\mathrm{d}}, m_{1\dots(n-3)}^{\mathrm{d}}\right), \quad \text{(D6)}$$

with

$$\mathcal{G}_{1\dots n}\left(\ell_{1\dots(n-3)}^{\mathrm{d}}, m_{1\dots(n-3)}^{\mathrm{d}}\right) = G_{1,2,1^d} \times G_{1^{d*},3,2^d} \times \cdots \times G_{(n-4)^{d*},n-2,(n-3)^d} \times G_{(n-3)^{d*},(n-1),n}. \quad \text{(D7)}$$

The polyspectrum is diagonal independent if $\mathcal{P}^{(n)}$ does not vary with $\ell_1^{\mathrm{d}} \dots \ell_{n-3}^{\mathrm{d}}$. In this case, equation (D6) takes a simpler form

$$\langle a_{\ell_1 m_1} \dots a_{\ell_n m_n} \rangle_c = \mathcal{P}^{(n)}(\ell_{1\dots n}) \times \mathcal{G}_{1,\dots,n} \quad \text{(D8)}$$

with

$$\mathcal{G}_{1,\dots,n} = \int \mathrm{d}^2\hat{n} \; Y_{1\dots n}(\hat{n}) = \sum_{\substack{\ell_{1\dots(n-3)}^{\mathrm{d}} \\ m_{1\dots(n-3)}^{\mathrm{d}}}} \mathcal{G}_{1\dots n}\left(\ell_{1\dots(n-3)}^{\mathrm{d}}, m_{1\dots(n-3)}^{\mathrm{d}}\right). \quad \text{(D9)}$$

APPENDIX E: PROJECTION OF 3D POLYSPECTRA ON THE SPHERE

Noting for simplicity $g(z) = \frac{\mathrm{dr}}{\mathrm{dz}} a(z)$, we have

$$a_{\ell m} = i^\ell \int \frac{\mathrm{d}^3 \boldsymbol{k}}{2\pi^2} \mathrm{d}z \, g(z) \, j_\ell(kr) \, Y_{\ell m}^*(\hat{k}) \, j_\nu(\boldsymbol{k}, z). \quad \text{(E1)}$$

Hence the n-order correlation function takes the form

$$\langle a_{\ell_1 m_1} \dots a_{\ell_n m_n} \rangle = i^{\ell_1 + \dots + \ell_n} \int \frac{\mathrm{d}^3 \boldsymbol{k}_{1\dots n}}{(2\pi^2)^n} \mathrm{d}z_{1\dots n} \left[g(z_i) \, j_{\ell_i}(k_i r_i) \times Y_{\ell_i m_i}^*(\hat{k}_i) \right]_{i=1\dots n} \times \underbrace{\left\langle \left[\mathrm{j}_\nu(\boldsymbol{k}_i, z_i) \right]_{i=1\dots n} \right\rangle}_{(2\pi)^3 \mathcal{P}_{\mathrm{j}}^{(n)}(\boldsymbol{k}_{1\dots n}, z_{1\dots n}) \, \delta^{(3)}(\boldsymbol{k}_1 + \dots + \boldsymbol{k}_n)} \quad \text{(E2)}$$

$$= \frac{(2\pi)^3}{(2\pi^2)^n} i^{\ell_1 + \dots + \ell_n} \int k_{1\dots n}^2 \mathrm{d}k_{1\dots n} \, \mathrm{d}z_{1\dots n} \times \left[g(z_i) \, j_{\ell_i}(k_i r_i) \right]_i \mathcal{P}_j^{(n)}(\boldsymbol{k}_{1\dots n}, z_{1\dots n}) \times \underbrace{\int \mathrm{d}^2\hat{k}_{1\dots n} \left[Y_{\ell_i m_i}^*(\hat{k}_i) \right]_i \delta^{(3)}(\boldsymbol{k}_1 + \dots + \boldsymbol{k}_n)}_{\equiv A^{(n)}(\ell_{1\dots n}, m_{1\dots n}, \boldsymbol{k}_{1\dots n})}, \quad \text{(E3)}$$

if $\mathcal{P}_j^{(n)}$ is diagonal independent, and with

$$A^{(n)}(\ell_{1\dots n}, m_{1\dots n}, \boldsymbol{k}_{1\dots n}) = \int \mathrm{d}^2\hat{k}_{1\dots n} \frac{\mathrm{d}^3 \boldsymbol{x}}{(2\pi)^3} \times \left[Y_{\ell_i m_i}^*(\hat{k}_i) \right]_i e^{i(\boldsymbol{k}_1 + \dots + \boldsymbol{k}_n) \cdot \boldsymbol{x}} \quad \text{(E4)}$$

$$= \frac{(4\pi)^n}{(2\pi)^3} \int \mathrm{d}^2\hat{k}_{1\dots n} \, \mathrm{d}^3 \boldsymbol{x} \times \sum_{\ell'_{1\dots n} m'_{1\dots n}} i^{\ell'_1 + \dots + \ell'_n} \left[Y_{\ell_i m_i}^*(\hat{k}_i) \times Y_{\ell'_i m'_i}(\hat{k}_i) \, j_{\ell'_i}(k_i x) \, Y_{\ell'_i m'_i}^*(\hat{x}) \right]_i \quad \text{(E5)}$$

$$= \frac{(4\pi)^n}{(2\pi)^3} i^{\ell_1+\cdots+\ell_n} \int x^2 \mathrm{d}x \left[j_{\ell_i}(k_i x) \right]_i \times \underbrace{\int \mathrm{d}^2\hat{n} \left[Y^*_{\ell_i m_i}(\hat{n}) \right]_{i=1\ldots n}}_{\equiv \mathcal{G}_{1\ldots n}}, \tag{E6}$$

where we introduced the Fourier form of the Dirac in line E4, the Rayleigh expansion of $e^{i\boldsymbol{k}_i\cdot\boldsymbol{x}}$ in line E5, used the orthonormality of the spherical harmonics and the definition of the generalized Gaunt coefficient in line E6 and the fact that it is a real number.

Hence the n-order (diagonal-independent) polyspectrum is

$$\mathcal{P}^{(n)}_{\rm IR}(\ell_1,\ldots,\ell_n) = \left(\frac{2}{\pi}\right)^n (-1)^{\ell_1+\cdots+\ell_n} \int k^2_{1\ldots n} \mathrm{d}k_{1\ldots n}\, \mathrm{d}z_{1\ldots n}\, x^2 \mathrm{d}x \times \left[g(z_i)\, j_{\ell_i}(k_i r_i) j_{\ell_i}(k_i x) \right]_i \mathcal{P}^{(n)}_j(\boldsymbol{k}_{1\ldots n}, z_{1\ldots n}), \tag{E7}$$

$$= \left(\frac{2}{\pi}\right)^n (-1)^{\ell_1+\cdots+\ell_n} \int k^2_{1\ldots n} \mathrm{d}k_{1\ldots n}\, \mathrm{d}z_{1\ldots n}\, x^2 \mathrm{d}x \times \left[a(z_i) \left.\frac{\mathrm{d}r}{\mathrm{d}z}\right|_{z_i} j_{\ell_i}(k_i r_i)\, j_{\ell_i}(k_i x)\, \overline{j}(\nu, z_i) \right]_i \times \mathcal{P}^{(n)}_{\rm gal}(\boldsymbol{k}_{1\ldots n}, z_{1\ldots n}) \tag{E8}$$

where for the last line, we used $\mathcal{P}^{(n)}_j = \overline{j}^n\, \mathcal{P}^{(n)}_{\rm gal}$ (valid for non-shot noise terms, with the flux-abundance independence assumption).

We can assume that, as a function of k_i, $\mathcal{P}^{(n)}_{\rm gal}$ varies slowly compared to the Bessel functions oscillations (the so-called Limber approximation). Then we have

$$\int k^2_{1\ldots n} \mathrm{d}k_{1\ldots n} \left[j_{\ell_i}(k_i r_i)\, j_{\ell_i}(k_i x) \right]_i \mathcal{P}^{(n)}_{\rm gal}(\boldsymbol{k}_{1\ldots n}, z_{1\ldots n}) \approx \mathcal{P}^{(n)}_{\rm gal}(k^*_{1\ldots n}, z_{1\ldots n}) \underbrace{\int k^2_{1\ldots n} \mathrm{d}k_{1\ldots n} \left[j_{\ell_i}(k_i r_i)\, j_{\ell_i}(k_i x) \right]_i}_{= \left[\frac{\pi}{2x^2}\, \delta(r_i - x) \right]_{i=1\ldots n}}, \tag{E9}$$

where $k^\star = (\ell + 1/2)/r$ is the peak of the Bessel function.

And equation (E8) simplifies to

$$\mathcal{P}^{(n)}_{\rm IR}(\ell_1,\ldots,\ell_n) = \int \frac{r^2 \mathrm{d}r}{r^{2n}} a^n(z)\, \overline{j}^n(\nu, z)\, \mathcal{P}^{(n)}_{\rm gal}(k^*_{1\ldots n}, z), \tag{E10}$$

with $k^\star = (\ell + 1/2)/r$, $r = r(z)$, and because of parity invariance $\ell_1 + \cdots + \ell_n$ is even (otherwise $\mathcal{P}^{(n)}$ is zero).

In particular at order 3, we get the bispectrum

$$b_{\ell_1\ell_2\ell_3} = \left(\frac{2}{\pi}\right)^3 (-1)^{\ell_1+\ell_2+\ell_3} \int k^2_{123} \mathrm{d}k_{123}\, \mathrm{d}z_{123}\, x^2 \mathrm{d}x \left[a(z_i) \left.\frac{\mathrm{d}r}{\mathrm{d}z}\right|_{z_i} \overline{j}(\nu, z_i)\, j_{\ell_i}(k_i r_i)\, j_{\ell_i}(k_i x) \right]_{i=123} \times B_{\rm gal}(k_{123}, z_{123}). \tag{E11}$$

And Limber approximation simplifies it to

$$b_{\ell_1\ell_2\ell_3} = \int \frac{\mathrm{d}z}{r^4} \frac{\mathrm{d}r}{\mathrm{d}z}\, a^3(z)\, \overline{j}^3(\nu, z)\, B_{\rm gal}(k^*_{123}, z). \tag{E12}$$

Again, except for shot noise terms which are treated in Section 4.3.

This paper has been typeset from a TEX/LATEX file prepared by the author.