

# Combining cluster number counts and galaxy clustering

To cite this article: Fabien Lacasa and Rogerio Rosenfeld JCAP08(2016)005

View the [article online](#) for updates and enhancements.

## Related content

- [THE CLUSTERING OF MASSIVE GALAXIES AT  \$z \sim 0.5\$  FROM THE FIRST SEMESTER OF BOSS DATA](#)  
Martin White, M. Blanton, A. Bolton et al.
- [SIMULATIONS OF WIDE-FIELD WEAK-LENSING SURVEYS. II. COVARIANCE MATRIX OF REAL-SPACE CORRELATION FUNCTIONS](#)  
Masanori Sato, Masahiro Takada, Takashi Hamana et al.
- [THE WEAK LENSING SIGNAL AND THE CLUSTERING OF BOSS GALAXIES. II. ASTROPHYSICAL AND COSMOLOGICAL CONSTRAINTS](#)  
Surhud More, Hironao Miyatake, Rachel Mandelbaum et al.

## Recent citations

- [Fast and easy super-sample covariance of large-scale structure observables](#)  
Fabien Lacasa and Julien Grain
- [The squeezed matter bispectrum covariance with responses](#)  
Alexandre Barreira
- [Cosmological constraints from a joint analysis of cosmic microwave background and spectroscopic tracers of the large-scale structure](#)  
Cyrille Doux *et al*



**IOP Astronomy** ebooks

Part of your publishing universe and your first choice for astronomy, astrophysics, solar physics and planetary science ebooks.

[iopscience.org/books/aas](http://iopscience.org/books/aas)

# Combining cluster number counts and galaxy clustering

**Fabien Lacasa and Rogerio Rosenfeld**

ICTP South American Institute for Fundamental Research,  
Instituto de Física Teórica, Universidade Estadual Paulista,  
São Paulo, Brazil

E-mail: [fabien@ift.unesp.br](mailto:fabien@ift.unesp.br), [rosenfel@ift.unesp.br](mailto:rosenfel@ift.unesp.br)

Received April 18, 2016

Revised June 23, 2016

Accepted July 18, 2016

Published August 2, 2016

**Abstract.** The abundance of clusters and the clustering of galaxies are two of the important cosmological probes for current and future large scale surveys of galaxies, such as the Dark Energy Survey. In order to combine them one has to account for the fact that they are not independent quantities, since they probe the same density field. It is important to develop a good understanding of their correlation in order to extract parameter constraints. We present a detailed modelling of the joint covariance matrix between cluster number counts and the galaxy angular power spectrum. We employ the framework of the halo model complemented by a Halo Occupation Distribution model (HOD). We demonstrate the importance of accounting for non-Gaussianity to produce accurate covariance predictions. Indeed, we show that the non-Gaussian covariance becomes dominant at small scales, low redshifts or high cluster masses. We discuss in particular the case of the super-sample covariance (SSC), including the effects of galaxy shot-noise, halo second order bias and non-local bias. We demonstrate that the SSC obeys mathematical inequalities and positivity. Using the joint covariance matrix and a Fisher matrix methodology, we examine the prospects of combining these two probes to constrain cosmological and HOD parameters. We find that the combination indeed results in noticeably better constraints, with improvements of order 20% on cosmological parameters compared to the best single probe, and even greater improvement on HOD parameters, with reduction of error bars by a factor 1.4-4.8. This happens in particular because the cross-covariance introduces a synergy between the probes on small scales. We conclude that accounting for non-Gaussian effects is required for the joint analysis of these observables in galaxy surveys.

**Keywords:** cosmological parameters from LSS, galaxy clusters, power spectrum

**ArXiv ePrint:** [1603.00918](https://arxiv.org/abs/1603.00918)

---

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Notations, inputs and conventions</b>	<b>3</b>
<b>3</b>	<b>Cluster number counts and its covariance</b>	<b>5</b>
<b>4</b>	<b>Galaxy angular power spectrum</b>	<b>7</b>
<b>5</b>	<b>Cross-covariance between cluster counts and galaxy power spectrum</b>	<b>9</b>
<b>6</b>	<b>Super-sample covariance</b>	<b>14</b>
<b>7</b>	<b>Covariance of the galaxy angular power spectrum</b>	<b>16</b>
<b>8</b>	<b>Joint covariance and consequences for parameter constraints</b>	<b>17</b>
8.1	Joint covariance	18
8.2	Fisher matrix constraints on cosmological parameters	20
8.3	Fisher constraints on HOD parameters	22
<b>9</b>	<b>Conclusion</b>	<b>24</b>
<b>A</b>	<b>3D halo-galaxy-galaxy bispectrum in the Halo Model</b>	<b>25</b>
<b>B</b>	<b>From the 3D hgg bispectrum to the cross-covariance</b>	<b>27</b>
B.1	Angle-independent bispectrum	28
B.2	Angle dependent bispectrum	30
<b>C</b>	<b>SSC inequalities</b>	<b>33</b>
<b>D</b>	<b>Multipole binning</b>	<b>35</b>

---

## 1 Introduction

The large scale structure of the universe has a long history of being a useful probe of both global cosmological parameters and the formation of structures in our universe. This is the motivation for carrying out large galaxy surveys, such as the Dark Energy Survey (DES<sup>1</sup>), which will use different observational probes, such as weak gravitational lensing, correlations in the distribution of galaxies, statistics of galaxy clusters, and type Ia supernovae, in order to improve our understanding of background cosmology and the growth of structures. For example the large scale distribution of matter is sensitive to the presence and specific properties of dark energy, see e.g. [2] for a review. On smaller scales however, the clustering is dependent on galaxy formation models and halo physics.

---

<sup>1</sup><http://www.darkenergysurvey.org>.

Different observables have different sensitivities to cosmological parameters, and combining them can help to break degeneracies present when only single probes are used. However these probes are in principle not independent. Thus in order to estimate parameter constraints and their accuracy from these large surveys, it is necessary to have a careful understanding of the statistical properties of the probes used, including their cross-covariance.

In this work we will be interested in the combination of two probes, namely the galaxy angular clustering and the abundances of clusters. The main motivation comes from DES-like surveys where these are two of the most important observables to be used to estimate cosmological parameters. Indeed cluster counts can be used to test cosmological models through both volume effects and the growth rate [3–5], e.g. different dark energy models affect differently the formation of dark matter haloes [6, 7]. Likewise, measurements of the galaxy correlation function can be used to estimate cosmological parameters and the formation of structures, see e.g. [8] for a recent example with SDSS observations.

Cluster counts and galaxy-galaxy correlations both probe the same density field, whose non-linearity couples different Fourier modes. Besides, some of the galaxies probed will be living in the same halos probed by the number counts. Through both these effects, we expect the two probes to be importantly correlated, in particular on small scales, where non-Gaussian effects are more relevant. Thus in order to determine the statistical power of combining cluster counts and 2-point galaxy clustering, we need to model the covariance of these observables as well as their cross-covariance. We will do so in the context of the halo model framework to describe the non-linear density field (see, e.g., [9] and references therein). The auto-covariance of cluster counts and of the 2-point galaxy correlation function are described by 2-point and 4-point correlation functions respectively, whereas the cross-covariance is related to a 3-point correlation function.

The necessary ingredients for the modelling are the halo mass function, the halo density profile and a model to describe how galaxies populate dark matter halos. We will use the so-called Tinker mass function [10], the Navarro-Frenk-White (NFW) density profile [11] and a halo occupation distribution (HOD) model from [12, 13]. In addition to the usual cosmological parameters, we will study the sensitivity of these probes to the dark energy equation of state (assuming a flat  $\Lambda$ CDM model) and to HOD parameters.

We build on previous studies in the literature. For example the covariance of the matter power spectrum was predicted with the halo model by Cooray and Hu [14] and compared favourably with simulations. Takada and Hu [15] emphasized the importance of the so-called super-sample covariance in the matter power spectrum, due to long wavelength modes outside the volume of the survey. Their role in simulations was discussed in [16]. In this article we will derive and discuss the super-sample covariance of our observables. Cluster counts have already been shown to be a powerful probe for cosmological constraints, even more so when combined with other probes. In particular Takada and Bridle [17] (see also [18]) showed that it mitigates the super-sample covariance of the cosmic shear power spectrum. This was later generalized by [19] to 3D  $n$ -point correlation functions with a joint likelihood. Cluster counts were also combined with the cluster power spectrum in [20]. Moreover, Hutsi showed that the cluster-galaxy correlation contains a significant amount of information for cosmological and HOD constraints [21]. Finally for photometric surveys, a framework was developed to analyse projected angular probes of weak lensing and the galaxy distribution in [22].

In this article, we will develop a unified framework to treat the cluster number counts and the galaxy-galaxy angular power spectrum in the full-sky case. We pay particular attention to the non-linear effects at small scales, to a careful modelling of the cross-covariance of these probes, and deriving and discussing the super-sample covariance for our observables.

Our goal can be easily stated as follows. We will consider grouping the data from cluster counts and angular power spectrum into a single data vector  $X$ :

$$X = \begin{pmatrix} N_{\text{cl}} \\ C_{\ell}^{\text{gal}} \end{pmatrix} \quad (1.1)$$

where  $N_{\text{cl}}$  is the number of clusters in a given mass and redshift bin and  $C_{\ell}^{\text{gal}}$  is the angular power spectrum of galaxies in a given redshift bin (the dependence on mass and redshift are kept implicit for simplicity). We want to model the joint covariance matrix

$$\text{Cov}(\hat{X}, \hat{X}) = \begin{pmatrix} \text{Cov}(\hat{N}_{\text{cl}}, \hat{N}_{\text{cl}}) & \text{Cov}(\hat{C}_{\ell}^{\text{gal}}, \hat{N}_{\text{cl}}) \\ \text{Cov}(\hat{N}_{\text{cl}}, \hat{C}_{\ell}^{\text{gal}}) & \text{Cov}(\hat{C}_{\ell}^{\text{gal}}, \hat{C}_{\ell}^{\text{gal}}) \end{pmatrix} \quad (1.2)$$

This covariance matrix will be used to study the statistical power of these observables in estimating parameters, using a Fisher matrix approach. In particular we will investigate the relevance of accounting for the cross-covariance represented by the off-diagonal terms in the covariance matrix eq. (1.2).

The cross-covariance is related to a bispectrum involving a halo-galaxy-galaxy correlator. This bispectrum is modelled in detail, with contributions arising from perturbation theory, the halo model and HOD. In particular we include the effects of second order halo bias as well as non-local bias and galaxy shot noise. We are able to show explicitly how the super-sample covariance appears naturally in the bispectrum, in a framework that incorporates it in a systematic way.

This article is organized as follows. In the next section we set the notations and conventions that will be used throughout the article. In section 3 we describe the cluster number counts and its covariance. Next in section 4 we model the galaxy angular power spectrum. In section 5 we describe the modelling of the cross-covariance between cluster counts and galaxy angular power spectrum. In section 6 we propose a unified framework to treat the super-sample covariance among different probes. In section 7 we discuss the covariance of the galaxy angular power spectrum. In section 8 we present the joint covariance and perform a Fisher matrix analysis of the constraints on cosmological and HOD parameters, respectively on large and small scales. Finally, we conclude in section 9. Several appendices are presented with details of the computations and proofs of assertions from the main text.

## 2 Notations, inputs and conventions

As stated in the Introduction, we will base our analysis on the halo model framework. In this framework, all galaxies are located in halos and one can write the number density of halos as:

$$n_h(\mathbf{x}) = \sum_i \delta^3(\mathbf{x} - \mathbf{x}_i) \quad (2.1)$$

where the index  $i$  denotes halos, with center located at  $\mathbf{x}_i$  and mass  $M_i$ .

In particular the halo mass function, giving the comoving number density of halos per unit mass is given by the ensemble average [9]:

$$\frac{dn_h}{dM} = \langle \mathcal{N}_{cl}(\mathbf{x}|M, z) \rangle, \quad (2.2)$$

where

$$\mathcal{N}_{cl}(\mathbf{x}|M, z) = \sum_i \delta^3(\mathbf{x} - \mathbf{x}_i) \delta(M - M_i). \quad (2.3)$$

Galaxies inhabit the dark matter halos and their number density can thus be written as:

$$n_{\text{gal}}(\mathbf{x}) = \sum_i \sum_{j=1}^{N_{\text{gal}}(i)} \delta^3(\mathbf{x} - \mathbf{x}_j), \quad (2.4)$$

where the index  $j$  refers to galaxies in a given halo  $i$  with  $N_{\text{gal}}(i)$  galaxies. We further assume that within a halo, the galaxies are distributed independently following the normalized halo density profile

$$u(\mathbf{x} - \mathbf{x}_i|M_i) = \frac{1}{M_i} \rho(\mathbf{x} - \mathbf{x}_i|M_i). \quad (2.5)$$

We will adopt the spherically symmetric NFW profile [11]

$$u(r|M_i) = \frac{1}{M} \frac{\rho_s}{r/r_s(1+r/r_s)^2}, \quad (2.6)$$

with  $r_s$  the scale radius and  $\rho_s$  is the scale density. They are specified as in, e.g. [23, 24].

The normalization enforces in particular that the Fourier transform  $u(k|M, z)$  obeys

$$u(k|M, z) \rightarrow 1 \quad \text{when} \quad k \rightarrow 0. \quad (2.7)$$

For the mass function, we use the one from Tinker et al. [10], which is a fit to the number density of halos obtained from simulations in cold dark matter scenarios. The halo bias relating fluctuations in the halo density field  $\delta_h$  and those of the dark matter field  $\delta_m$  are parametrized via the local bias ansatz:

$$\delta_h = \sum_i b_i(M, z) (\delta_m^i - \langle \delta_m^i \rangle), \quad (2.8)$$

where  $\langle \delta_h \rangle = 0$  by construction and the biases are consistently related to the chosen mass function. We will use both the first [25] and second order bias [26] in this work, neglecting higher order terms.

Galaxies are introduced using a HOD model which specifies how galaxies populate halos. In this model, the probability distribution of galaxies is a function of the halo mass only and is composed of two populations: a central galaxy drawn from a binomial distribution, and satellite galaxies drawn from a Poisson distribution conditioned to the presence of a central galaxy (we will not consider the issue of sample selection here). More specifically, one writes the average number of galaxies in a halo as

$$\langle N_g \rangle = \langle N_{\text{cen}} \rangle + \langle N_{\text{sat}} \rangle \quad (2.9)$$

and we will adopt [27–29]

$$\langle N_{\text{cen}} \rangle = \frac{1}{2} \left[ 1 + \text{erf} \left( \frac{\log M - \log M_{\text{min}}}{\sigma_{\log M}} \right) \right] \quad (2.10)$$

and

$$\langle N_{\text{sat}} \rangle = \frac{1}{2} \left[ 1 + \text{erf} \left( \frac{\log M - \log 2M_{\text{min}}}{\sigma_{\log M}} \right) \right] \left( \frac{M}{M_{\text{sat}}} \right)^{\alpha_{\text{sat}}} \quad (2.11)$$

This parametrization has 4 parameters: a mass threshold above which a halo has a large probability of containing a central galaxy ( $M_{\text{min}}$ ), the width of the transition of the central probability ( $\sigma_{\log M}$ ), the typical mass above which a halo contains satellite galaxies ( $M_{\text{sat}}$ ) and the index of the power law for the number of satellites at large halo masses ( $\alpha_{\text{sat}}$ ).

Finally, we use the following mathematical notations. Estimators of a given quantity are noted with an overhat, and overbar denote the corresponding average. We also use the comoving volume per solid angle, defined as

$$dV = r^2(z) \frac{dr}{dz} dz, \quad (2.12)$$

as integration element for simplicity, with  $r(z)$  being the comoving distance.

We may also denote only with indices the multiplication of some quantities, e.g.:

$$dV_{12} = dV_1 dV_2 = r^2(z_1) \frac{dr}{dz} \Big|_{z_1} r^2(z_2) \frac{dr}{dz} \Big|_{z_2} dz_1 dz_2. \quad (2.13)$$

Otherwise stated, we will adopt a flat  $\Lambda$ CDM model with the cosmological parameters  $\Omega_b h^2 = 0.022$ ,  $\Omega_{\text{cdm}} h^2 = 0.12$ ,  $h = 0.67$ ,  $n_s = 0.962$ ,  $\sigma_8 = 0.834$  and use the parametrization of Eisenstein and Hu [30] to generate the linear matter power spectrum. For HOD parameters we use  $\alpha_{\text{sat}} = 1.3$ ,  $M_{\text{min}} = 10^{12.2} M_{\odot}$  and  $M_{\text{sat}} = 10^{13.2} M_{\odot}$  and  $\sigma_{\log M} = 0.2$  [28].

### 3 Cluster number counts and its covariance

Cluster number counts is an important cosmological observable that is sensitive to both the background evolution as well as the growth of perturbations. Cluster counts are also independent of the cluster density profile and HOD modelling, depending exclusively on the cluster mass function.

The estimator of the binned cluster number counts per unit of solid angle is simply the number of clusters in a given mass and redshift bin divided by the survey area, which we take to be the full sky from now,  $\Omega_S = 4\pi$ . It can be written as [9]:

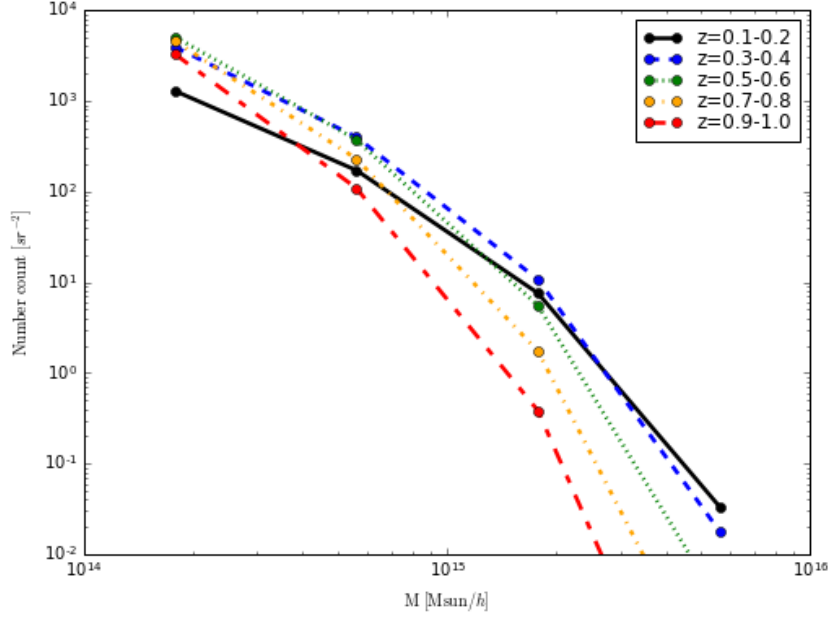
$$\hat{N}_{\text{cl}}(i_M, i_z) = \frac{1}{4\pi} \int d^2 \hat{n} dM dz \sum_j \delta(M - M_j) \delta^2(\hat{n} - \hat{n}_j) \delta(z - z_j), \quad (3.1)$$

where the mass and redshift integrals run over the bins defined by  $i_M$  and  $i_z$  and the index  $j$  runs over all halos in the universe with mass  $M_j$  and centered at  $\vec{x}_j = r(z_j) \hat{n}_j$ . We assume that all clusters can be detected without fake detections and that both their mass and redshift can be measured perfectly. The effects of the spread in the mass-observable relation for the cluster mass, the purity and completeness of the cluster catalogue and of the photometric redshift errors will be considered in future works.

Eq. (3.1) can be rewritten as:

$$\hat{N}_{\text{cl}}(i_M, i_z) = \frac{1}{4\pi} \int d^2 \hat{n} dM dV \mathcal{N}_{\text{cl}}(\mathbf{x} = r \hat{n} | M, z) \quad (3.2)$$





**Figure 1.** Cluster number counts in five representative redshift bins. The mass bins have width  $\Delta \log M = 0.5$  and the points are drawn at the start of the mass bin.

and hence the average of the cluster number count is given by

$$\bar{N}_{\text{cl}}(i_M, i_z) = \langle \hat{N}_{\text{cl}}(i_M, i_z) \rangle = \int dM dV \frac{dn_h}{dM} \quad (3.3)$$

where the mass and redshift integrals implicitly run over the corresponding bins. The resulting cluster counts are shown in figure 1.

Eq. (3.2) can be rewritten in harmonic space as

$$\hat{N}_{\text{cl}}(i_M, i_z) = \int \frac{d^3 \mathbf{k}}{(2\pi)^3} dM dV j_0(kr) \mathcal{N}_{\text{cl}}(\mathbf{k}|M, z). \quad (3.4)$$

From eq. (3.4), the covariance of cluster number counts is given by

$$\text{Cov} \left( \hat{N}_{\text{cl}}(i_M, i_z), \hat{N}_{\text{cl}}(j_M, j_z) \right) = \int \frac{d^3 \mathbf{k}}{(2\pi)^3} dM_{12} dV_{12} j_0(kr_1) j_0(kr_2) \left. \frac{dn_h}{dM} \right|_{M_1, z_1} \left. \frac{dn_h}{dM} \right|_{M_2, z_2} P_{\text{cl}}(k|M_{12}, z_{12})$$

where the cluster power spectrum

$$P_{\text{cl}}(k|M_{12}, z_{12}) = (2\pi)^3 \delta^3(\mathbf{k}_1 - \mathbf{k}_2) \langle \delta_h(\mathbf{k}_1|M_1, z_1) \delta_h(\mathbf{k}_2|M_2, z_2) \rangle \quad (3.5)$$

has contributions from the 1- and 2-halo terms and we have introduced the halo density contrast  $\delta_h$  such that

$$\mathcal{N}_{\text{cl}}(\mathbf{k}|M, z) = \left. \frac{dn_h}{dM} \right|_{M, z} \delta_h(\mathbf{k}|M, z). \quad (3.6)$$

The 1-halo term contribution is well known [9] and it represents a Poissonian shot-noise error.



The 2-halo contribution to the number counts covariance can be written as:

$$\begin{aligned} \text{Cov}_{2h} \left( \hat{N}_{\text{cl}}(i_M, i_z), \hat{N}_{\text{cl}}(j_M, j_z) \right) &= \int dM_{12} dV_{12} \left. \frac{dn_h}{dM} \right|_{M_1, z_1} \left. \frac{dn_h}{dM} \right|_{M_2, z_2} b_1(M_1, z_1) b_1(M_2, z_2) \\ &\times \int \frac{d^3 \mathbf{k}}{(2\pi)^3} j_0(kr_1) j_0(kr_2) P_{\text{lin}}(k|z_{12}) \end{aligned} \quad (3.7)$$

where  $P_{\text{lin}}(k|z_{12})$  is the linear matter power spectrum and we have assumed that halos are linearly biased with respect to matter.

This equation can be rewritten as:

$$\text{Cov}_{2h} \left( \hat{N}_{\text{cl}}(i_M, i_z), \hat{N}_{\text{cl}}(j_M, j_z) \right) = \int dV_{12} \frac{\partial n_h}{\partial \delta_b}(i_M, z_1) \frac{\partial n_h}{\partial \delta_b}(j_M, z_2) \sigma_{\text{proj}}^2(z_1, z_2) \quad (3.8)$$

with

$$\frac{\partial n_h}{\partial \delta_b}(i_M, z) \equiv \int_{M \in \text{bin}(i_M)} dM \frac{dn_h}{dM} b_1(M, z), \quad (3.9)$$

and

$$\sigma_{\text{proj}}^2(z_1, z_2) \equiv \int \frac{d^3 \mathbf{k}}{(2\pi)^3} j_0(kr_1) j_0(kr_2) P_{\text{lin}}(k|z_{12}). \quad (3.10)$$

The significance of this form of writing the result and its relation to sample variance will become clear in the discussion presented in section 6.

This 2-halo contribution is off-diagonal: it couples different mass bins and different redshift bins. Figure 2 shows the normalised cluster counts covariance, in absolute value, obtained with the following set up: three mass bins of width  $\Delta \log M = 0.5$  between  $10^{14} M_\odot$  and  $10^{15.5} M_\odot$  and nine redshift bins of width  $\Delta z = 0.1$  between  $z = 0.1$  and 1. We plotted the absolute value of the covariance for clarity, as the cross-covariance between neighbouring redshift bins is negative due to the behaviour of  $\sigma_{\text{proj}}^2(z_1, z_2)$  (shown in figure 6 of section 6). The covariance shows significant correlation between the two lowest mass bins, and the off-diagonality increases at lower redshift, to the left, as expected from the larger nonlinear effects.

## 4 Galaxy angular power spectrum

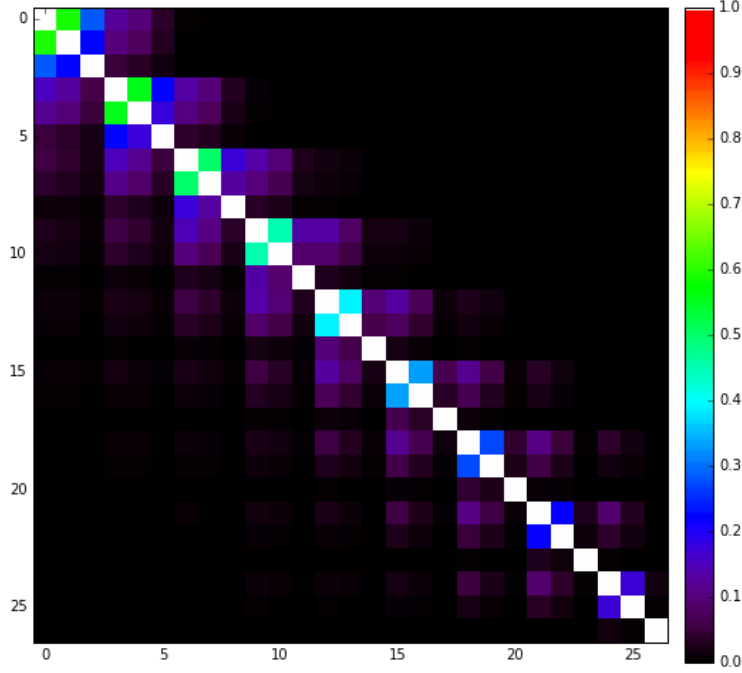
The projected galaxy density contrast<sup>2</sup> in the direction  $\hat{n}$  in a redshift bin  $i_z$ ,  $\delta_{\text{gal}}(\hat{n}, i_z)$  is:

$$\delta_{\text{gal}}(\hat{n}, i_z) = \int_{i_z} dV \frac{\bar{n}_{\text{gal}}(z)}{\Delta N_{\text{gal}}(i_z)} \delta_{\text{gal}}(\mathbf{x} = r\hat{n}, z) \quad (4.1)$$

where we denote as  $\bar{n}_{\text{gal}}(z)$  the 3D average galaxy density at redshift  $z$  modelled by

$$\bar{n}_{\text{gal}}(z) = \int_{i_z} dV dM \frac{dn_h}{dM} \langle N_g \rangle, \quad (4.2)$$

<sup>2</sup>Note that we normalise the contrast by the theoretical average  $\bar{n}_{\text{gal}}$ , instead of the often used empirical average over the survey. Indeed we note that normalising by the empirical average yields an ill-defined statistical quantity, as a ratio of estimators, which is in particular biased. Normalising by the theoretical average from the model is perfectly feasible too and yields constraints identical to those using the empirical absolute difference  $n_{\text{gal}} - \bar{n}_{\text{gal}}$ . We leave for future works the question of dealing with all the problems of the empirical normalisation.



**Figure 2.** Cluster number counts covariance for 3 mass bins and 9 redshift bins. See text for details.

and where the number of galaxies per steradian in the redshift bin  $i_z$  is given by

$$\Delta N_{\text{gal}}(i_z) = \int_{i_z} dV \bar{n}_{\text{gal}}(z). \quad (4.3)$$

After spherical harmonic decomposition, the harmonic coefficients are:

$$a_{\ell m}^{\text{gal}}(i_z) = i^\ell \int dV \frac{\bar{n}_{\text{gal}}(z)}{\Delta N_{\text{gal}}(i_z)} \int \frac{d^3 \mathbf{k}}{2\pi^2} j_\ell(kr) \delta_{\text{gal}}(\mathbf{k}, z) Y_{\ell m}^*(\hat{k}). \quad (4.4)$$

The galaxy power spectrum estimator is then:

$$\hat{C}_\ell^{\text{gal}}(i_z, j_z) = \frac{1}{2\ell + 1} \sum_m a_{\ell m}^{\text{gal}}(i_z) \left( a_{\ell m}^{\text{gal}}(j_z) \right)^* \quad (4.5)$$

$$= 4\pi \int dV_{12} \frac{\bar{n}_{\text{gal}}(z_1) \bar{n}_{\text{gal}}(z_2)}{\Delta N_{\text{gal}}(i_z) \Delta N_{\text{gal}}(j_z)} \int \frac{d^3 \mathbf{k}_{12}}{(2\pi)^6} j_\ell(k_1 r_1) j_\ell(k_2 r_2) \\ \times P_\ell(\hat{k}_1 \cdot \hat{k}_2) \delta_{\text{gal}}(\mathbf{k}_1, z_1) \delta_{\text{gal}}^*(\mathbf{k}_2, z_2) \quad (4.6)$$

It's average is

$$\bar{C}_\ell^{\text{gal}}(i_z, j_z) = \frac{2}{\pi} \int dV_{12} \frac{\bar{n}_{\text{gal}}(z_1) \bar{n}_{\text{gal}}(z_2)}{\Delta N_{\text{gal}}(i_z) \Delta N_{\text{gal}}(j_z)} \int k^2 dk j_\ell(kr_1) j_\ell(kr_2) P_{\text{gal}}(k|z_{12}) \\ \simeq \frac{\delta_{i_z, j_z}}{\Delta N_{\text{gal}}(i_z)^2} \int dV \bar{n}_{\text{gal}}(z)^2 P_{\text{gal}}(k_\ell|z) \quad (4.7)$$

where in the last step we have used Limber's approximation, taking the power spectrum at the maximum of the Bessel functions,  $k_\ell = \frac{\ell+1/2}{r(z)}$ , and performing the integration of the

Bessel functions that yields a Dirac delta function in redshift (for more details on the Limber approximation for polyspectra at any order, see appendix E of [33]). The 3-d galaxy power spectrum is defined in the usual manner

$$P_{\text{gal}}(k|z_{12}) = (2\pi)^3 \delta^3(\mathbf{k} - \mathbf{k}') \langle \delta_{\text{gal}}(\mathbf{k}, z_1) \delta_{\text{gal}}^*(\mathbf{k}', z_2) \rangle. \quad (4.8)$$

We can use the Halo Model to write the galaxy power spectrum as a sum of a 2-halo, a 1-halo and a shot-noise term [9]:

$$P_{\text{gal}}(k|z_{12}) = P_{\text{gal}}^{2\text{h}}(k|z_{12}) + P_{\text{gal}}^{1\text{h}}(k|z_{12}) + P_{\text{gal}}^{\text{shot}}(k|z_{12}) \quad (4.9)$$

with

$$P_{\text{gal}}^{2\text{h}}(k|z_{12}) = b_1^{\text{gal,eff}}(k, z_1) b_1^{\text{gal,eff}}(k, z_2) P_{\text{DM}}(k|z_{12}) \quad (4.10)$$

$$P_{\text{gal}}^{1\text{h}}(k|z_{12}) = \frac{\delta_{z_1, z_2}}{\bar{n}_{\text{gal}}(z_1)^2} \int dM \frac{dn_h}{dM} \langle N_{\text{gal}}(N_{\text{gal}} - 1)(M) \rangle u(k|M, z_1)^2 \quad (4.11)$$

$$P_{\text{gal}}^{\text{shot}}(k|z_{12}) = \frac{\delta_{z_1, z_2}}{\bar{n}_{\text{gal}}(z_1)} \quad (4.12)$$

where  $b_1^{\text{gal,eff}}$  is the (first-order) effective galaxy bias:

$$b_1^{\text{gal,eff}}(k, z) = \int dM \left. \frac{dn_h}{dM} \right|_{M, z} \frac{\langle N_{\text{gal}}(M) \rangle}{\bar{n}_{\text{gal}}(z)} u(k|M) b_1(M, z). \quad (4.13)$$

Note that on large scales,  $u(k|M) \rightarrow 1$  and the effective bias goes to a constant, the usual galaxy bias:  $b_1^{\text{gal}}(z) = \int dM \frac{dn_h}{dM} \frac{\langle N_{\text{gal}}(M) \rangle}{\bar{n}_{\text{gal}}(z)} b_1(M, z)$ .

The resulting galaxy power spectrum and its different terms are shown in figure 3. Cosmological and HOD parameters used for this computation are given in section 2. The galaxy density ranges between  $2.6 \times 10^5$  and  $2.5 \times 10^6$  galaxies per steradian. Consequently, the shot-noise term reaches at best 8% of the total spectrum on these scales. On the other hand, the 1-halo term becomes fairly important at low redshift, dominating at the (relatively) small scales, as expected.

## 5 Cross-covariance between cluster counts and galaxy power spectrum

We now turn to the modelling of the cross-covariance between the number counts studied in section 3 and the galaxy spectrum studied in section 4. It is indeed the main goal of this study to estimate the impact of including this cross-covariance in the extraction of cosmological and astrophysical parameters from large galaxy surveys.

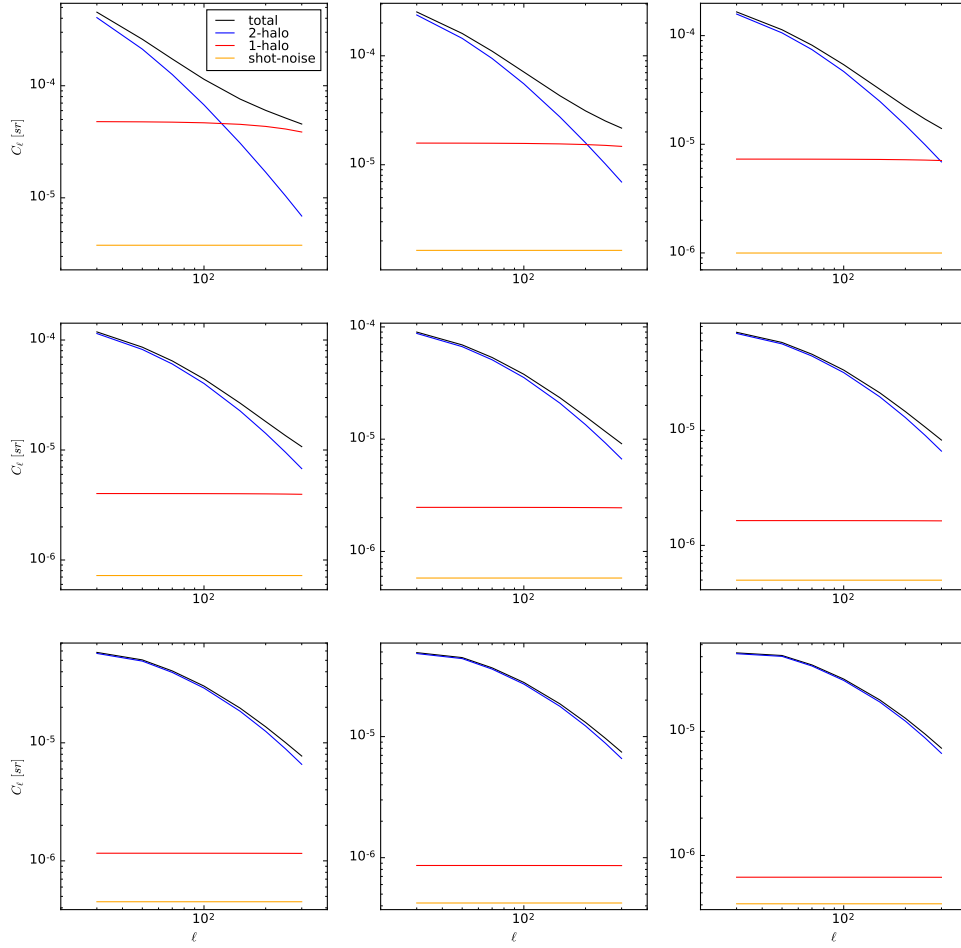
Following eq. (3.4) and eq. (4.6), the covariance between the number counts and the galaxy power spectrum estimators is:

$$\begin{aligned} \text{Cov} \left( \hat{N}_{\text{cl}}(i_M, i_z), \hat{C}_{\ell}^{\text{gal}}(j_z, k_z) \right) &= 4\pi \int \frac{d^3 \mathbf{k}_{123}}{(2\pi)^9} dM_1 dV_{123} \frac{\bar{n}_{\text{gal}}(z_2) \bar{n}_{\text{gal}}(z_3)}{\Delta N_{\text{gal}}(j_z)^2} P_{\ell}(\hat{k}_2 \cdot \hat{k}_3) \\ &\quad \times j_0(k_1 r_1) j_{\ell}(k_2 r_2) j_{\ell}(k_3 r_3) \langle \mathcal{N}_{\text{cl}}(\mathbf{k}_1|M_1, z_1) \delta_{\text{gal}}(\mathbf{k}_2, z_2) \delta_{\text{gal}}^*(\mathbf{k}_3, z_3) \rangle_c. \end{aligned} \quad (5.1)$$

Therefore the cross covariance is related to a 3-point function involving one halo/cluster and two galaxies, the so-called bispectrum defined by:

$$\langle \delta_h(\mathbf{k}_1|M_1, z_1) \delta_g(\mathbf{k}_2, z_2) \delta_g(\mathbf{k}_3, z_3) \rangle_c = (2\pi)^3 \delta^3(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) B_{\text{hgg}}(k_{123}|M_1, z_{123}), \quad (5.2)$$

where we used the halo density contrast  $\delta_h$  defined in eq. (3.6).

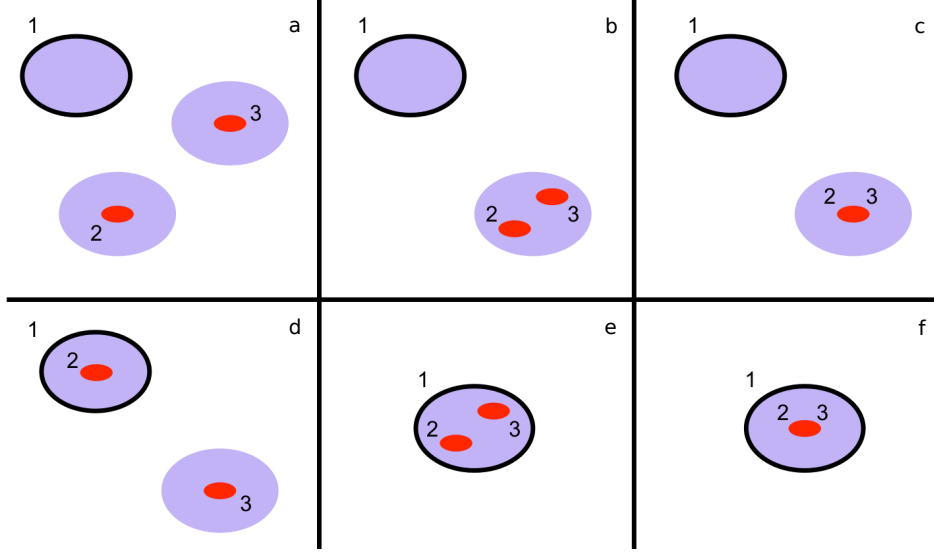


**Figure 3.** Galaxy angular power spectrum and its different terms on large scales  $\ell = 30 - 300$ . Each subplot represent a redshift bin of width  $\Delta z = 0.1$ , starting at  $z = 0.1 - 0.2$  at the upper left and ending at  $z = 0.9 - 1$  at the bottom right. The inputs used for the computations are spelled out in section 2.

As done for the galaxy power spectrum, we model this halo-galaxy-galaxy bispectrum using the halo model coupled with a Halo Occupation Distribution. There are six terms involved and they can be easily visualised with simple diagrams, as first proposed by Lacasa et al. [29], which are shown in figure 4.

The first three diagrams (upper row) can couple number counts and galaxy spectrum in different redshift bins, and we will later identify with super-sample covariance. They are described as:

- a) the 3h term (upper left diagram) quantifies how halo counts and the 2-halo part of the galaxy spectrum fluctuate together with large scale structure. It involves the halo bispectrum, which has three contributions: the dark matter bispectrum from standard eulerian second order perturbation theory [32] (hereafter referred as the 2PT term), second order halo bias (b2 term), and non-local halo bias (s2 term).
- b) the 2h-1h2g term (upper central diagram) quantifies how halo counts and the 1-halo part of the galaxy spectrum fluctuate together with large scale structure.



**Figure 4.** Diagrams for the halo-galaxy-galaxy bispectrum. From left to right and top to bottom: 3h, 2h-1h2g, 2h-1h1g, 2h-2h, 1h2g, 1h1g.

- c) the 2h-1h1g term (upper right diagram) quantifies how halo counts and the shot-noise part of the galaxy spectrum fluctuate together with large scale structure.

The three diagrams in the bottom row contribute only when the number count and the galaxy spectrum are in the same redshift bin and are described as:

- d) the 2h-2h term (bottom left diagram) quantifies how the halos in that bin of mass source the 2-halo part of the galaxy power spectrum.  
 e) the 1h2g term (bottom central diagram) quantifies how the halos in that bin of mass source the 1-halo part of the galaxy power spectrum.  
 f) the 1h1g term (bottom right diagram) quantifies how the halos in that bin of mass source the shot-noise part of the galaxy power spectrum.

Therefore we can write:

$$B_{\text{hgg}}(k_{123}, z_{123} | M_1, z_{123}) = B^{3h} + B^{2h-2h} + B^{2h-1h2g} + B^{2h-1h1g} + B^{1h2g} + B^{1h1g} \quad (5.3)$$

with

$$B^{3h} = B^{2PT} + B^{b2} + B^{s2}. \quad (5.4)$$

The expressions for each of these terms are given in appendix A.

We show in appendix B that using the Limber approximation in  $k_2$  and  $k_3$  but not in  $k_1$  the cross-covariance reduces to

$$\begin{aligned} \text{Cov} \left( \hat{N}_{\text{cl}}(i_M, i_z), \hat{C}_\ell(j_z, k_z) \right) &= \frac{\delta_{j_z, k_z}}{4\pi} \int \frac{dM_1 dV_{12}}{\Delta N_{\text{gal}}(j_z)^2} \frac{dn_h}{dM} \Big|_{M_1, z_1} \bar{n}_{\text{gal}}(z_2)^2 \\ &\times \frac{2}{\pi} \int k_1^2 dk_1 B_{\text{hgg}}(k_1, k_\ell, k_\ell | M_1, z_{122}) j_0(k_1 r_1) j_0(k_1 r_2) \end{aligned} \quad (5.5)$$

and the different cross covariance terms are:

$$\text{Cov}_{1\text{h}1\text{g}} \left( \hat{N}_{\text{cl}}(i_M, i_z), \hat{C}_\ell^{\text{gal}}(j_z) \right) \approx \frac{\delta_{i_z, j_z}}{4\pi} \int dV_1 \frac{n_{\text{gal}}(z_1 | i_M)}{\Delta N_{\text{gal}}(i_z)^2} \quad (5.6)$$

$$\text{Cov}_{1\text{h}2\text{g}} \left( \hat{N}_{\text{cl}}(i_M, i_z), \hat{C}_\ell^{\text{gal}}(j_z) \right) = \frac{\delta_{i_z, j_z}}{4\pi} \int dV_1 \frac{\bar{n}_{\text{gal}}(z_1)^2 P_{\text{gal}}^{\text{1h}}(k_\ell, z_1 | i_M)}{\Delta N_{\text{gal}}(i_z)^2} \quad (5.7)$$

$$\begin{aligned} \text{Cov}_{2\text{h}-2\text{h}} \left( \hat{N}_{\text{cl}}(i_M, i_z), \hat{C}_\ell^{\text{gal}}(j_z) \right) &= \frac{\delta_{i_z, j_z}}{4\pi} \int dV_1 2 b_1^{\text{gal,eff}}(k_\ell, z_1 | i_M) b_1^{\text{gal,eff}}(k_\ell, z_1) \\ &\times \frac{\bar{n}_{\text{gal}}(z_1)^2}{\Delta N_{\text{gal}}(i_z)^2} P_{\text{DM}}(k_\ell, z_1) \end{aligned} \quad (5.8)$$

$$\begin{aligned} \text{Cov}_{2\text{h}-1\text{h}1\text{g}} \left( \hat{N}_{\text{cl}}(i_M, i_z), \hat{C}_\ell^{\text{gal}}(j_z) \right) &= \int \frac{dV_{12}}{\Delta N_{\text{gal}}(j_z)^2} \frac{\partial n_h}{\partial \delta_b}(i_M, z_1) \\ &\times \bar{n}_{\text{gal}}(z_2) b_1^{\text{gal,eff}}(z_2) \sigma_{\text{proj}}^2(z_1, z_2) \end{aligned} \quad (5.9)$$

$$\begin{aligned} \text{Cov}_{2\text{h}-1\text{h}2\text{g}} \left( \hat{N}_{\text{cl}}(i_M, i_z), \hat{C}_\ell^{\text{gal}}(j_z) \right) &= \int \frac{dV_{12}}{\Delta N_{\text{gal}}(j_z)^2} \frac{\partial n_h}{\partial \delta_b}(i_M, z_1) \\ &\times \bar{n}_{\text{gal}}(z_2) b_1^{\text{gal,eff}-2}(k_\ell, z_2) \sigma_{\text{proj}}^2(z_1, z_2) \end{aligned} \quad (5.10)$$

$$\text{Cov}_{\text{s}2} \left( \hat{N}_{\text{cl}}(i_M, i_z), \hat{C}_\ell^{\text{gal}}(j_z) \right) = 0 \quad (5.11)$$

$$\begin{aligned} \text{Cov}_{\text{b}2} \left( \hat{N}_{\text{cl}}(i_M, i_z), \hat{C}_\ell^{\text{gal}}(j_z) \right) &= \int \frac{dV_{12}}{\Delta N_{\text{gal}}(j_z)^2} 2 \bar{n}_{\text{gal}}(z_2)^2 b_1^{\text{gal,eff}}(k_\ell, z_2) b_2^{\text{gal,eff}}(k_\ell, z_2) \\ &\times \frac{\partial n_h}{\partial \delta_b}(i_M, z_1) P_{\text{DM}}(k_\ell | z_2) \sigma_{\text{proj}}^2(z_1, z_2) \end{aligned} \quad (5.12)$$

$$\begin{aligned} \text{Cov}_{2\text{PT}} \left( \hat{N}_{\text{cl}}(i_M, i_z), \hat{C}_\ell^{\text{gal}}(j_z) \right) &= \int \frac{dV_{12}}{\Delta N_{\text{gal}}(j_z)^2} 4 F_{\text{avg}} \bar{n}_{\text{gal}}(z_2)^2 b_1^{\text{gal,eff}}(k_\ell, z_2)^2 \\ &\times \frac{\partial n_h}{\partial \delta_b}(i_M, z_1) P_{\text{DM}}(k_\ell | z_2) \sigma_{\text{proj}}^2(z_1, z_2), \end{aligned} \quad (5.13)$$

where we re-used notations of the previous sections, and introduced the following additional short-hand notations (notice that some of the integrals are on a given mass bin):

$$n_{\text{gal}}(z | i_M) = \int_{M \in \text{bin}(i_M)} dM \frac{dn_h}{dM} \langle N_{\text{gal}}(M) \rangle \quad (5.14)$$

$$P_{\text{gal}}^{\text{1h}}(k, z | i_M) = \frac{1}{\bar{n}_{\text{gal}}^2(z)} \int_{M \in \text{bin}(i_M)} dM \frac{dn_h}{dM} \langle N_{\text{gal}}(N_{\text{gal}} - 1)(M) \rangle u^2(k | M, z) \quad (5.15)$$

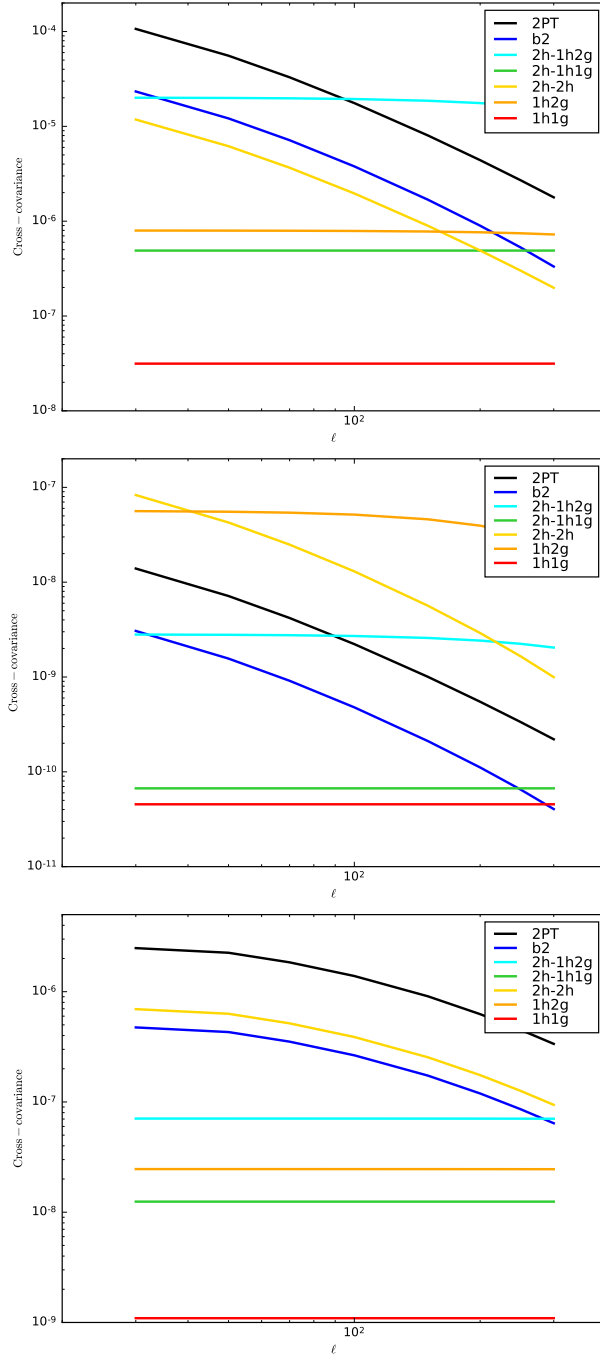
$$b_{1,2}^{\text{gal,eff}}(k, z | i_M) = \frac{1}{\bar{n}_{\text{gal}}(z)} \int_{M \in \text{bin}(i_M)} dM \frac{dn_h}{dM} \langle N_{\text{gal}}(M) \rangle b_{1,2}(M, z) u(k | M, z) \quad (5.16)$$

$$b_1^{\text{gal,eff}-2}(k, z) = \frac{1}{\bar{n}_{\text{gal}}(z)} \int dM \frac{dn_h}{dM} \langle N_{\text{gal}}(N_{\text{gal}} - 1)(M) \rangle b_1(M, z) u(k | M, z)^2 \quad (5.17)$$

$$F_{\text{avg}} = \frac{17}{21} \quad (5.18)$$

The resulting cross-covariance is shown in figure 5 as a function of multipole  $\ell$ , with three plots for different cases of mass bin and redshift bin. In these plots we take  $i_z = j_z$  for simplicity and because this is where the cross-covariance is maximal.

We see from figure 5 that different terms become important depending on scale, mass and redshift. In particular HOD terms are far from negligible, being dominant in some



**Figure 5.** Cross-covariance between cluster counts and the galaxy spectrum, depending on the multipole of the spectrum. *Top:* low redshift low mass case,  $z \in [0.1, 0.2]$   $\log M \in [14, 14.5]$ . *Center:* low redshift high mass case,  $z \in [0.1, 0.2]$   $\log M \in [15, 15.5]$ . *Bottom:* high redshift low mass case,  $z \in [0.8, 0.9]$   $\log M \in [14, 14.5]$ .



cases, while the 2PT term can become negligible in some cases. We note that the 2PT term is the only term that one would have considered from a model with only linear bias and perturbation theory. We thus conclude that it is critical to use a fully non-linear model, such as the halo model combined with HOD that we have been using here, in order to estimate accurately the cross-covariance.

## 6 Super-sample covariance

We can unify the equations (5.9), (5.10), (5.12), (5.13) for the 2h-1h1g, 2h-1h2g, b2 and 2PT terms involving the super-sample covariance, into a single equation:

$$\begin{aligned} \text{Cov}_{\text{SSC}} \left( \hat{N}_{\text{cl}}(i_M, i_z), \hat{C}_{\ell}^{\text{gal}}(j_z) \right) &\equiv \text{Cov}_{2\text{h}-1\text{h}1\text{g}} + \text{Cov}_{2\text{h}-1\text{h}2\text{g}} + \text{Cov}_{\text{b}2} + \text{Cov}_{2\text{PT}} \\ &= \int dV_{12} \frac{\bar{n}_{\text{gal}}(z_2)^2}{\Delta N_{\text{gal}}(j_z)^2} \frac{\partial n_h}{\partial \delta_b}(i_M, z_1) \frac{\partial P_{\text{gal}}(k_{\ell}|z_2)}{\partial \delta_b} \sigma_{\text{proj}}^2(z_1, z_2) \end{aligned} \quad (6.1)$$

$$(6.2)$$

if we define

$$\begin{aligned} \frac{\partial P_{\text{gal}}(k|z)}{\partial \delta_b} &\equiv \underbrace{4F_{\text{avg}}}_{=\frac{68}{21}} b_1^{\text{gal,eff}}(k, z)^2 P_{\text{DM}}(k|z) + b_1^{\text{gal,eff}-2}(k, z) / \bar{n}_{\text{gal}}(z) \\ &\quad + 2b_1^{\text{gal,eff}}(k, z) b_2^{\text{gal,eff}}(k, z) P_{\text{DM}}(k|z) + b_1^{\text{gal,eff}}(z) / \bar{n}_{\text{gal}}(z). \end{aligned} \quad (6.3)$$

These terms correspond to the diagrams in the top row of figure 4 and represent how cluster counts and the galaxy power spectrum fluctuate together due to the large scale structure.

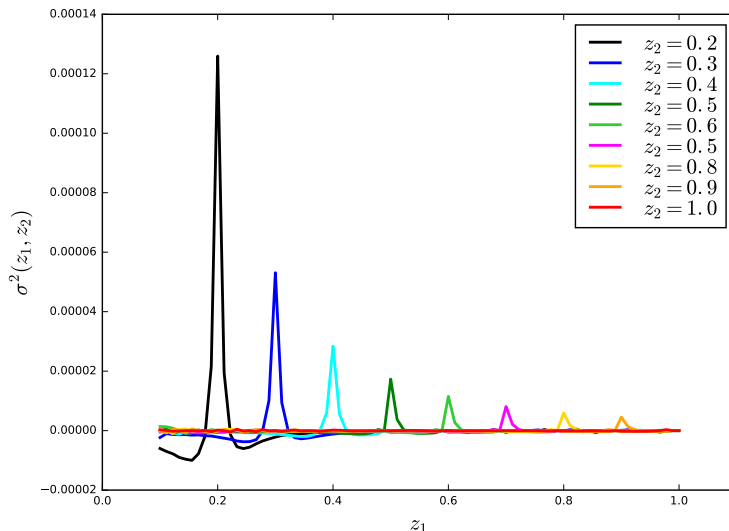
By analogy with previous literature results for the 3D matter power spectrum [15], eq. (6.3) can be interpreted as the reaction of the galaxy power spectrum to a change of background density and hence the notation  $\frac{\partial P_{\text{gal}}}{\partial \delta_b}$ . We also interpret  $\frac{\partial n_h}{\partial \delta_b}$ , defined in eq. (3.9) as the reaction of the cluster number counts to a change of background density.

The first two terms in eq. (6.3) correspond to the expression found in [15]. However, in contrast to [15] we have here derived our equations without any Taylor expansion. Furthermore, we have derived here the contributions from second order bias (third term), galaxy shot-noise (fourth term), and we have shown that non-local tidal effects give a zero contribution (see appendix B for details). In our framework, eq. (6.2) can be interpreted physically as the fact that both cluster counts and the galaxy spectrum react to a change of background, so their covariance is related to the covariance of the background density  $\sigma_{\text{proj}}^2(z_1, z_2)$ . Indeed:

$$\sigma_{\text{proj}}^2(z_1, z_2) = \int \frac{d^3 \mathbf{k}}{(2\pi)^3} j_0(kr_1) j_0(kr_2) P_{\text{DM}}(k|z_{12}) \quad (6.4)$$

$$= \langle \delta_{\text{avg}}(z_1) \delta_{\text{avg}}(z_2) \rangle_c \quad \text{with} \quad \delta_{\text{avg}}(z) = \int \frac{d^2 \hat{n}}{4\pi} \delta(r\hat{n}, z). \quad (6.5)$$

This source of covariance has been dubbed “super-sample covariance” in the literature [15, 16], as it gives a non-diagonal covariance due to modes larger than the survey. In our case the survey is the whole sphere and the Bessel function  $j_0$  indicates that the contributing Fourier modes are those associated with the monopole. Eq. (3.8) for the 2-halo term of the covariance of cluster counts can also be interpreted as a super-sample covariance



**Figure 6.** Covariance of the background density:  $\sigma^2_{\text{proj}}(z_1, z_2)$  for different redshifts.

contribution, though it is more often more simply called “sample variance” in the cluster literature [3].

As a generalisation, it is natural to expect that the covariance of two observables  $\mathcal{O}_1$  and  $\mathcal{O}_2$  will have a contribution of the form:

$$\text{Cov}_{\text{SSC}}(\mathcal{O}_1, \mathcal{O}_2) = \int dV_{12} \frac{\partial \mathcal{O}_1}{\partial \delta_b}(z_1) \frac{\partial \mathcal{O}_2}{\partial \delta_b}(z_2) \sigma^2_{\text{proj}}(z_1, z_2) \quad (6.6)$$

Our framework also extends the work of [15] (see their appendix A) since we find that the SSC covariance of projected observables involves an integral over *two* redshifts instead of a single redshift. If we were able to use Limber’s approximation in  $k_1$  in equation (5.5), then  $\sigma^2_{\text{proj}}(z_1, z_2)$  would collapse to a Dirac delta function of redshifts and we would indeed end up with a single redshift integral. However, as is evident from figure 6, this covariance does have a redshift width.

Physically, this means that the background density of two neighbouring slices of the past light cone are correlated. As the redshift separation increases, this correlation decreases with oscillations, asymptoting to zero. This behaviour is the reason why neighbouring redshift bins have some anti-correlation in figure 2. We note that this behaviour would be even more important for finer redshift bins, e.g. for analysis of a spectroscopic or semi-spectroscopic survey.

Finally in appendix C, we show that super-sample covariance obeys the following mathematical inequality:

$$\text{Cov}_{\text{SSC}}(\mathcal{O}_1, \mathcal{O}_2) \leq \sqrt{\text{Cov}_{\text{SSC}}(\mathcal{O}_1, \mathcal{O}_1) \times \text{Cov}_{\text{SSC}}(\mathcal{O}_2, \mathcal{O}_2)} \quad (6.7)$$

for any observables  $\mathcal{O}_1$  and  $\mathcal{O}_2$ .

As an example, for number counts this implies

$$\begin{aligned} \text{Cov}_{2h}(N_{\text{cl}}(i_M, i_z), N_{\text{cl}}(j_M, j_z)) &\leq \left( \text{Cov}_{2h}(N_{\text{cl}}(i_M, i_z), N_{\text{cl}}(i_M, i_z)) \right. \\ &\quad \times \left. \text{Cov}_{2h}(N_{\text{cl}}(j_M, j_z), N_{\text{cl}}(j_M, j_z)) \right)^{1/2} \end{aligned} \quad (6.8)$$

for any bins of redshifts and mass  $i_z, j_z, i_M, j_M$ .

In our case, we found inequality eq. (6.7) to be saturated by some terms of the joint covariance matrix which will be described in section 8, and could even be sometimes violated at the percent level due to numerical errors. As such, enforcing inequality eq. (6.7) a posteriori would help to regularize the covariance matrix in cases where it was highly degenerate (low redshifts, small angular scales).

More generally we show in appendix C that super-sample covariance is positive. That is, even in the absence of the other covariance terms, the covariance matrix is certain to have positive eigenvalues.

## 7 Covariance of the galaxy angular power spectrum

Let us recall the equation for the galaxy power spectrum estimator eq. (4.6):

$$\begin{aligned} \hat{C}_\ell^{\text{gal}}(i_z) &= 4\pi \int dV_{12} \frac{\bar{n}_{\text{gal}}(z_1) \bar{n}_{\text{gal}}(z_2)}{\Delta N_{\text{gal}}(i_z)^2} \int \frac{d^3 \mathbf{k}_{12}}{(2\pi)^6} j_\ell(k_1 r_1) j_\ell(k_2 r_2) \\ &\quad \times P_\ell(\hat{k}_1 \cdot \hat{k}_2) \delta_{\text{gal}}(\mathbf{k}_1, z_1) \delta_{\text{gal}}^*(\mathbf{k}_2, z_2) \end{aligned} \quad (7.1)$$

Taking the covariance of this estimator will involve the quantity

$$\langle \delta_{\text{gal}}(\mathbf{k}_1, z_1) \delta_{\text{gal}}^*(\mathbf{k}_2, z_2) \delta_{\text{gal}}(\mathbf{k}_3, z_3) \delta_{\text{gal}}^*(\mathbf{k}_4, z_4) \rangle \quad (7.2)$$

The unconnected part of this average yields the usual Gaussian  $C_\ell$  covariance:

$$\text{Cov} \left( \hat{C}_\ell^{\text{gal}}(i_z), \hat{C}_{\ell'}^{\text{gal}}(j_z) \right) = \frac{2 C_\ell^{\text{gal}}(i_z, j_z)^2}{2\ell + 1} \delta_{\ell, \ell'} \quad (7.3)$$

$$= \frac{2 C_\ell^{\text{gal}}(i_z)^2}{2\ell + 1} \delta_{\ell, \ell'} \delta_{i_z, j_z} \quad (7.4)$$

where the second equality is a consequence of Limber's approximation forcing  $i_z = j_z$ .

The connected part of eq. (7.2) is

$$\begin{aligned} \langle \delta_{\text{gal}}(\mathbf{k}_1, z_1) \delta_{\text{gal}}^*(\mathbf{k}_2, z_2) \delta_{\text{gal}}(\mathbf{k}_3, z_3) \delta_{\text{gal}}^*(\mathbf{k}_4, z_4) \rangle_c &= (2\pi)^3 \delta^{(3)}(\mathbf{k}_1 - \mathbf{k}_2 + \mathbf{k}_3 - \mathbf{k}_4) \\ &\quad \times T_{\text{gal}}(\mathbf{k}_1, -\mathbf{k}_2, \mathbf{k}_3, -\mathbf{k}_4) \end{aligned} \quad (7.5)$$

where  $T_{\text{gal}}$  is the galaxy 3D trispectrum. Such term yields in principle the power spectrum covariance [34]

$$\text{Cov} \left( \hat{C}_\ell^{\text{gal}}(i_z), \hat{C}_{\ell'}^{\text{gal}}(j_z) \right) = \frac{1}{4\pi} T_{\ell\ell'}^{\ell'\ell'}(\ell_{\text{diag}} = 0 | i_z, j_z) \quad (7.6)$$

where  $T_{\ell_1 \ell_2}^{\ell_3 \ell_4}(\ell_{\text{diag}})$  is the 2D (projected) trispectrum (see the derivation of eq. (2.80) in chapter 2 of [34] for details).

The galaxy 3D trispectrum is a complicated beast with the contribution from many terms: there are 14 different diagrams, given e.g. in figure 7 of [33], with non-trivial permutations and some diagrams splitting into contributions from perturbation theory, both at second and third order, from halo quadratic and cubic local bias as well as from non-local bias effects. A complete census of these terms is beyond the scope of this article and left to a future work. Furthermore, there is no analytical solution known to the authors for the 2D projection of a 3D trispectrum with a general momentum dependence. This projection is known to us only for the diagonal-independent case, i.e.  $T_{\text{gal}}$  depending only on the moduli

$(k_1, k_2, k_3, k_4)$ . In that case, the projection is given in appendix E of [33] and recast here for convenience for the case of interest:

$$T_{\ell\ell}^{\ell'\ell'}(\ell_{\text{diag}}|i_z, j_z) = \delta_{i_z, j_z} \int dV \frac{\bar{n}_{\text{gal}}(z)^4}{\Delta N_{\text{gal}}(i_z)^4} T_{\text{gal}}(k_\ell, k_\ell, k_{\ell'}, k_{\ell'}|z) \quad (7.7)$$

This is the case of the 1-halo term of the galaxy trispectrum, that is when the four points of the correlation function hit four different galaxies inside the same halo:

$$T_{\text{gal}}^{1h}(k_1, k_2, k_3, k_4|z) = \int dM \frac{dn_h}{dM} \frac{\langle N(N-1)(N-2)(N-3) \rangle (M)}{\bar{n}_{\text{gal}}^4} \times u(k_1|M, z) u(k_2|M, z) u(k_3|M, z) u(k_4|M, z). \quad (7.8)$$

We also found this to be the case of 1-halo shot-noise terms, that is when two or more points of the correlation function hit the same galaxies, and all galaxies are in the same halo. However we found these shot-noise terms to be negligible in all our tests.

An important covariance contribution which does not fit the previous criterion (diagonal-independence) is the super-sample covariance (SSC) discussed in section 6. By analogy with the SSC of the cross-covariance and following the discussion of section 6, we take:

$$\text{Cov}_{\text{SSC}} \left( C_\ell^{\text{gal}}(i_z), C_{\ell'}^{\text{gal}}(j_z) \right) = \int dV_{12} \frac{\bar{n}_{\text{gal}}(z_1)^2 \bar{n}_{\text{gal}}(z_2)^2}{\Delta N_{\text{gal}}(i_z)^2 \Delta N_{\text{gal}}(j_z)^2} \frac{\partial P_{\text{gal}}(k_\ell, z_1)}{\partial \delta_b} \frac{\partial P_{\text{gal}}(k_{\ell'}, z_2)}{\partial \delta_b} \sigma_{\text{proj}}^2(z_1, z_2) \quad (7.9)$$

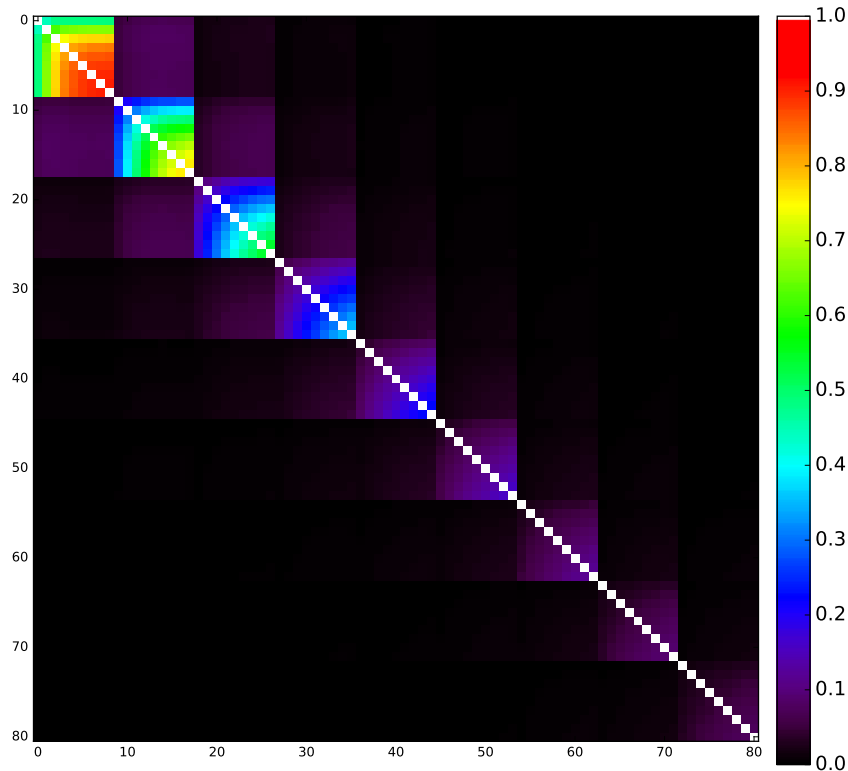
where all terms have been defined previously. We leave a rigorous derivation of this equation for a future work.

In realistic data analysis, multipoles of the power spectrum are binned together. The binning and its effect on the covariance is discussed in detail in appendix D. This aspect is critical to produce realistic covariance matrices. Indeed, the Gaussian diagonal error bars decrease as  $1/\sqrt{\Delta\ell}$  with binning, whereas other terms like, in particular SSC, decrease much slower because of their large off-diagonal contribution.

Figure 7 shows the resulting correlation matrix (in absolute value) of the binned galaxy spectrum on large scales  $\ell = 30 - 300$  with 9 bins of width  $\Delta\ell = 30$  and 9 redshift bins of width  $\Delta z = 0.1$  starting at  $z = 0.1$ . Small scales will also be considered in the next section. We see that the covariance matrix of the galaxy spectrum is highly off-diagonal at low redshift, and still has  $\sim 10\%$  off-diagonal elements in the highest redshift bin. Due to super-sample covariance, there is also some negative covariance between the neighbouring redshift bins at low redshift, which is plotted in absolute value in figure 7 and reaches  $\sim 10\%$  for  $0.1 < z < 0.3$ .

## 8 Joint covariance and consequences for parameter constraints

This section presents the main results of the article, showing the outcome of our modelling of the joint covariance of cluster counts and the galaxy angular power spectrum developed in the previous sections. The different contributions to the covariance matrix are analysed and the dominance of non-gaussian effects at low redshifts and small scales is emphasized. A Fisher matrix study is used to evaluate the importance of combining these observables in the determination of cosmological and HOD parameters using different binnings in the angular scale to reflect large scales (mostly sensitive to cosmological parameters) or small scales (mostly sensitive to HOD parameters).



**Figure 7.** Covariance of the galaxy angular power spectrum for nine multipole bins and nine redshift bins. Similar to figure 2 but for the galaxy spectrum instead of cluster counts. See text for details.

### 8.1 Joint covariance

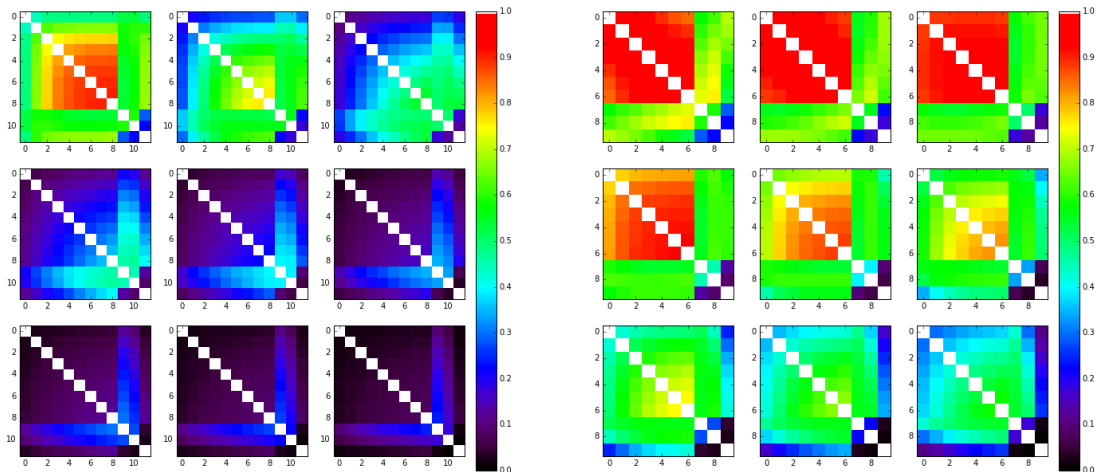
As before, we consider nine redshift bins of width  $\Delta z = 0.1$  between  $z_{\min} = 0.1$  and  $z_{\max} = 1$ , and three logarithmic mass bins for the cluster counts of width  $\Delta \log M = 0.5$  between  $M_{\min} = 10^{14} M_{\odot}$  and  $M_{\max} = 10^{15.5} M_{\odot}$ .

The angular multipoles are binned in order to follow the more realistic procedure performed on real data. The issue and the effects of binning are explained in more details in appendix D. We consider two case studies for the binning scheme:

- a “cosmological” case focusing on large angular scales, with nine regular multipole bins between  $\ell_{\min} = 30$  and  $\ell_{\max} = 300$  (9 bins of width  $\Delta \ell = 30$ );
- a “HOD” case focusing on smaller angular scales, with seven regular multipole bins between  $\ell_{\min} = 300$  and  $\ell_{\max} = 1000$  (7 bins of width  $\Delta \ell = 100$ ).

As implied by the names, the first case is expected to be representative of a cosmological constraint by focusing on linear scales, while the second is expected to be representative of a constraint on the Halo Occupation Distribution parameters by focusing on small non-linear scales. It is not, however, the purpose of this article to study rigorously the separation of the linear and non-linear regimes. The inclusion of external priors and a detailed investigation of a cosmological study are left for future work.

The joint correlation matrix of the galaxy power spectrum and the cluster counts can be seen in figure 8, with the cosmological case on the left panel and the HOD case on the right panel.



**Figure 8.** *Left:* joint correlation matrix of  $(C_\ell^{\text{gal}}, N_{\text{cl}})$  in the cosmological case; *Right:* in the HOD case. Each plot is composed of nine subplots corresponding to the nine redshift bins ( $z=0.1-0.2$  to  $z=0.9-1$ ). In each subplot, the upper left block (respectively  $9 \times 9$  in the cosmological case and  $7 \times 7$  in the HOD case) is the auto-covariance of  $C_\ell^{\text{gal}}$ , the  $3 \times 3$  bottom right block is the auto-covariance of  $N_{\text{cl}}$ , and finally the cross-covariance is given in the rectangular blocks.

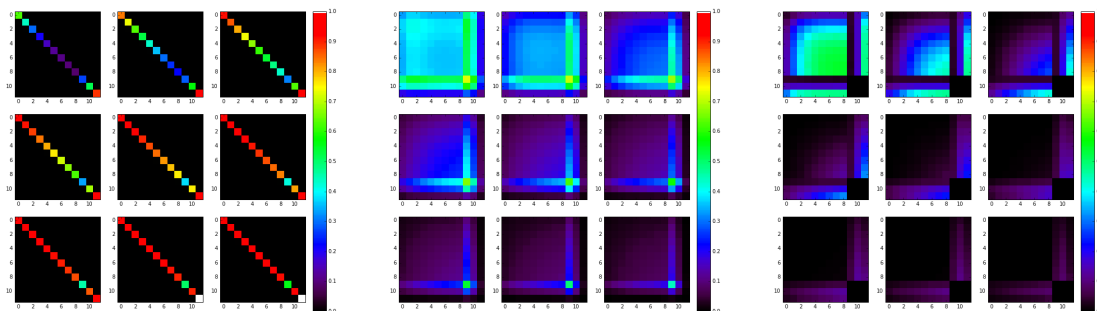
This figure is one of the main results of this article. One can clearly see the importance of non-Gaussian off-diagonal terms at low redshift. As expected, their significance decreases at higher redshifts, as perturbations become smaller and linear theory becomes a better approximation. Note however that even at redshift  $z = 0.9 - 1$  and large angular scales (left panel), the off-diagonal elements reach  $\geq 20\%$  in the cross-covariance, showing how important it is to account for non-Gaussianity.

Furthermore, one sees that non-Gaussianity becomes dramatic at small angular scales (right panel), to the point that it reaches total and irremediable domination at low redshift. Properly accounting for these effects is thus absolutely critical to analyse the small scale power spectrum, as they saturate the information content of the galaxy power spectrum alone.

To make more evident the different contributions to the covariance matrix arising from standard terms (Gaussian for  $C_\ell^{\text{gal}}$  and Poissonian for  $N_{\text{cl}}$ ), super-sample covariance and other non-Gaussian terms, we show in figure 9 their separate contributions in the cosmological (large scales) case. We see that standard covariance terms only contribute to the diagonal of the covariance matrix, as expected, and that they dominate at high redshift. As redshift decreases and the universe gets more structured, their relative contribution decreases and they end up being important only for  $C_\ell^{\text{gal}}$  on the largest scales and for  $N_{\text{cl}}$  at the highest mass bin.

Super-sample covariance gives a relatively smooth contribution across the elements of the covariance matrix, being higher at low redshift though still reaching 10-20% at high redshift. For the cross-covariance, it correlates the cluster counts at the lowest mass with all the galaxy spectrum, though more importantly with the smallest scales.

The other non-Gaussian terms have a more complex behaviour, with strong dependence on scale, redshift and mass. It is mostly negligible at high redshifts, especially in the auto-covariances, but become very important at small redshifts. It increases strongly with multipole as the non-linear part of the galaxy spectrum becomes important. For the cross-covariance, it correlates mostly the cluster counts at the highest masses with the galaxy spectrum on small scales, even reaching large scales in the lowest redshifts.



**Figure 9.** Separate contributions to the joint correlation matrix in the cosmological case: standard terms (left panel), super-sample covariance (middle panel) and other non-Gaussian terms (right panel). All matrix elements are normalised to the *total* variance, so that the sum of these three matrices gives the left panel of figure 8.

case / probe(s)	$C_{\ell}^{\text{gal}}$	$N_{\text{cl}}$	joint	independent
cosmological	141.09	162.18	194.35	214.96
HOD	161.55	171.63	235.70	237.61

**Table 1.** Signal-to-noise ratios of the different probes and their combination, in the two case studies.

One measure of the importance of the probes is the so-called signal-to-noise ratio (SNR). The SNR of an observable  $\mathcal{O}$  with covariance matrix  $\mathcal{C}$  is computed through the standard equation:

$$(\text{SNR})^2 = \mathcal{O}^T \cdot \mathcal{C}^{-1} \cdot \mathcal{O}. \quad (8.1)$$

Using our computed covariance matrices in the cosmological and HOD cases, we show in table 1 the SNR for each probe and the joint combination.

The two probes have comparable SNR, and as expected the combined analysis yields a larger value, being even higher in the hypothetical independent case than in the case where the cross-covariance is properly accounted for. In the next subsections we present a more detailed analysis of this comparison.

## 8.2 Fisher matrix constraints on cosmological parameters

In this subsection, we compare the strength of cosmological parameter constraints that can be reached by using the galaxy angular power spectrum, the cluster number counts, and their combination. We also consider the hypothetical case where the galaxy spectrum and the number counts are independent, in order to isolate the particular effect of properly including their cross-covariance. We emphasize that these constraints are not meant to be completely realistic, as we did not include a number of experimental effects in the modelling developed here (in particular photometric redshift errors, purity and completeness of cluster detections, and scatter in the cluster mass determination). That said, the choices of redshift binning, mass binning and mass range that we made are sufficiently conservative to mitigate some experimental errors, so that we expect the inclusion of experimental effects to be a correction to our results that will not change qualitatively our conclusions.

The strength of parameter constraints are evaluated by a standard Fisher matrix formalism. Specifically for a vectorized observable  $\mathcal{O}$  and model parameters  $\theta_i$ , the element



observable/parameter	$\sigma_8$	$\Omega_m h^2$	$w_{\text{DE}}$
$C_\ell^{\text{gal}}$	10.1 %	8.67 %	36.1 %
$N_{\text{cl}}$	1.17 %	4.10 %	7.48 %
joint	0.90 %	3.08 %	6.16 %
independent	0.86 %	2.96 %	5.87 %

**Table 2.**  $1\sigma$  marginalised error bars on cosmological parameters in the cosmological case study.

$(i, j)$  of the Fisher matrix is given by:

$$F_{i,j} = \left( \frac{\partial \mathcal{O}}{\partial \theta_i} \right)^T \cdot \mathcal{C}^{-1} \cdot \frac{\partial \mathcal{O}}{\partial \theta_j} \quad (8.2)$$

and  $F^{-1}$  is the covariance matrix of the Gaussian likelihood of parameters.

In this article, the parameter of interest are both cosmological ( $\sigma_8, \Omega_m h^2, w_{\text{DE}}$ ) and HOD parameters ( $\alpha_{\text{sat}}, M_{\text{min}}, M_{\text{sat}}$ ), with fixed  $\sigma_{\log M} = 0.2$ . For this subsection we are concentrating on cosmological parameters by marginalising over HOD parameters without priors. In the framework of Fisher matrices, through properties of Gaussian multivariate probabilities, this corresponds simply to extracting the appropriate block of the inverse Fisher matrix. Given

$$F^{-1} = \begin{pmatrix} F_{\text{c,c}}^{-1} & F_{\text{c,H}}^{-1} \\ F_{\text{H,c}}^{-1} & F_{\text{H,H}}^{-1} \end{pmatrix} \quad (8.3)$$

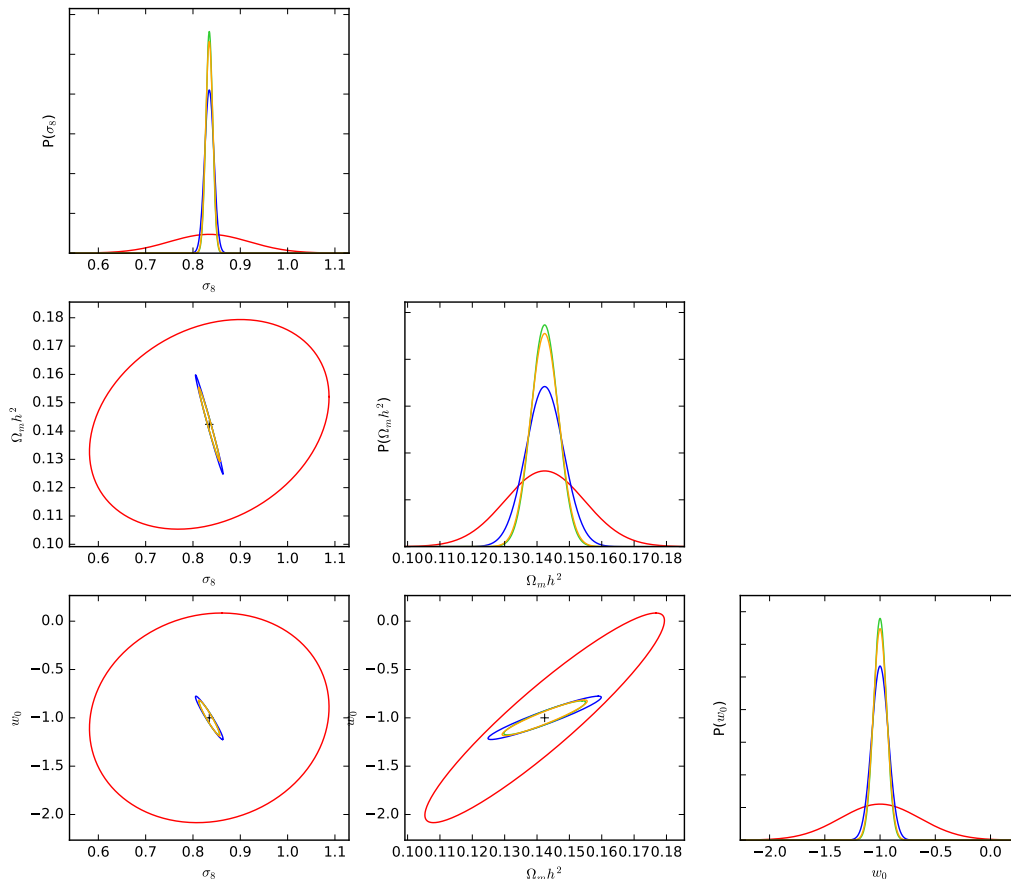
where c stands for cosmology and H for HOD,  $F_{\text{c,c}}^{-1}$  is the covariance matrix of cosmological parameters after HOD marginalisation.<sup>3</sup>

With this setup, figure 10 shows the constraint on cosmological parameters in the cosmological case explained in section 8.1. In this figure, the red lines give the Fisher ellipses and 1-dimensional probability distribution function (pdf) for the constraints coming from the galaxy power spectrum alone, the blue lines are for the constraints from the cluster number counts alone, orange is for the joint analysis, and finally green corresponds to the hypothetical case where the two probes were independent (i.e. setting the cross-covariance to zero).

Table 2 summarizes the marginalised error bars on each parameter that may be achieved in each case.

From figure 10 and table 2, one sees that constraints from the galaxy angular power spectrum are weaker than cluster counts constraints for the case studied. We emphasize that this is partly due to the choice of angular scales considered and that a study being able to reliably push the cosmological analysis to smaller scales would find better galaxy power spectrum constraints. In our case though, the difference of constraints between the galaxy power spectrum and cluster counts is partly due to the small difference in SNR (see table 1), the degeneracy with HOD parameters, and the higher sensitivity of cluster counts to cosmology. In the case of  $\sigma_8$ , for instance, the galaxy constraints are weakened by a factor larger than 6 due to the degeneracy with galaxy bias, while it is well-known that cluster abundances are exceptionally sensitive to  $\sigma_8$ .

<sup>3</sup>It turns out to be important to marginalise over HOD parameters, instead of fixing them to their fiducial values, as that can significantly increase the cosmological error bars. An important example is  $\sigma_8$  for which the error bars are increased by a factor larger than 6 with  $C_\ell^{\text{gal}}$  in the cosmological study case, due to its high degeneracy with galaxy bias.



**Figure 10.** 1-dimensional pdfs and  $3\sigma$  2-dimensional contours for cosmological parameter constraints in the cosmological study case. Red lines give the constraints from  $C_\ell^{\text{gal}}$ , blue from  $N_{\text{cl}}$ , orange from the joint analysis, green from the hypothetical independent case.

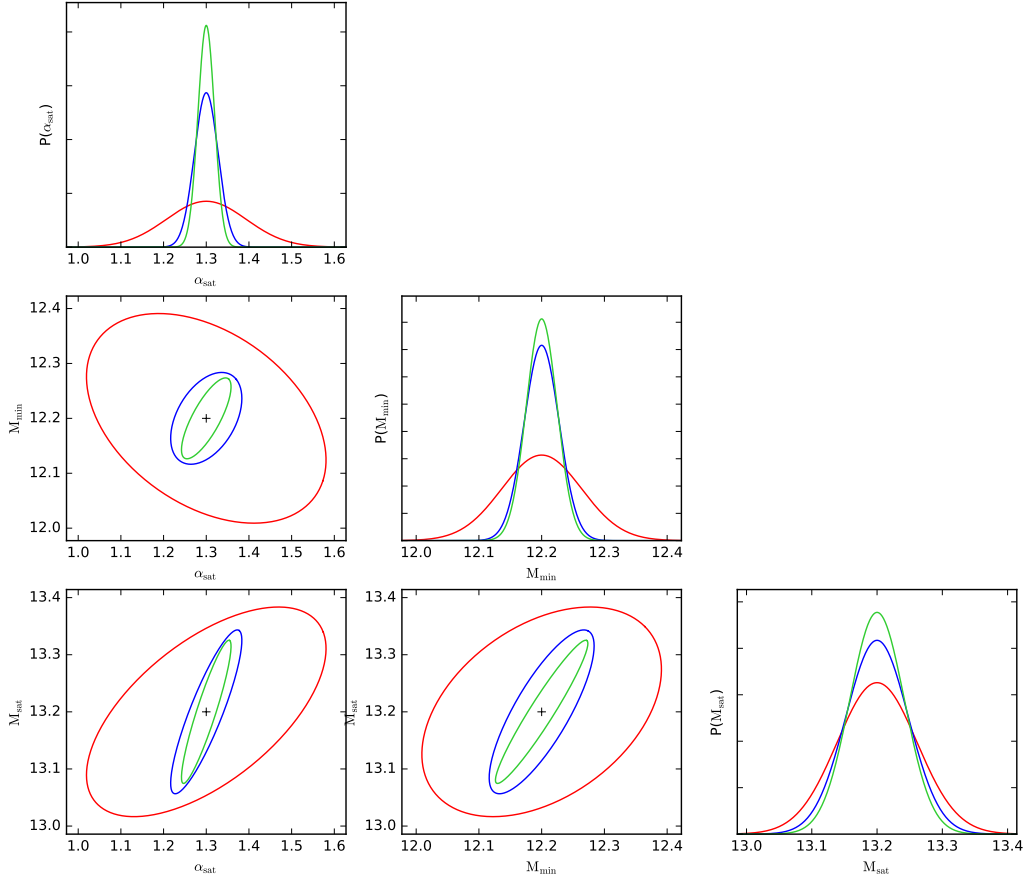
Due to the higher sensitivity of cluster counts to cosmology as compared to the galaxy angular power spectrum, the joint cosmological constraints are close to the cluster counts constraints, as can be seen in figure 10. An improvement of 17-25% is still to be noted. Additionally, we see that the joint constraints are slightly worse than in the hypothetical independent case, so the inclusion of the cross-covariance slightly degrades the constraints.

### 8.3 Fisher constraints on HOD parameters

This subsection mirrors section 8.2, but focusing on HOD parameters instead of cosmological ones. We follow the same Fisher equations, and HOD constraints are achieved after marginalisation over cosmology, which corresponds to extracting the block  $F_{\text{H,H}}^{-1}$  in eq. (8.3).

Contrary to section 8.2 however, number counts alone do not produce any parameter constraints, as by definition cluster counts do not depend on HOD. Number counts however help in a combination with the galaxy angular power spectrum in two ways: first they reduce the error on cosmology, which helps when marginalising over it, and second they break the degeneracy between the mass function and HOD parameters, by fixing the former.

Figure 11 shows the forecasted constraints on HOD parameters, in the HOD case study. Red lines show the constraints from the galaxy power spectrum alone, green lines for the joint analysis and blue corresponds to the hypothetical independent case.



**Figure 11.** 1-dimensional pdfs and  $3\sigma$  2-dimensional contours for HOD parameter constraints in the HOD study case. Red lines give the constraints from  $C_\ell^{\text{gal}}$ , green from the joint analysis, blue from the hypothetical independent case.

observable / parameter	$\alpha_{\text{sat}}$	$\log M_{\text{min}}$	$\log M_{\text{sat}}$
$C_\ell^{\text{gal}}$	7.20 %	0.52 %	0.46 %
joint	1.49 %	0.20 %	0.32 %
independent	2.14 %	0.23 %	0.36 %

**Table 3.**  $1\sigma$  marginalised error bars on HOD parameters in the HOD case study.

A quantitative assessment of the constraints is given in table 3 which lists the marginalised  $1\sigma$  error bars on each HOD parameter in the different cases.

From figure 11 and table 3, one sees that the galaxy HOD constraints are clearly improved by including the information from cluster counts. The improvement is the most spectacular for  $\alpha_{\text{sat}}$  where the errors bars are decreased by a factor larger than 4. This can be understood through the following reasoning. Since  $\alpha_{\text{sat}}$  governs the number of galaxies in massive haloes and massive haloes are rare, there can be important variations in their number due to cosmic variance. The effect of  $\alpha_{\text{sat}}$  on the galaxy power spectrum is thus highly degenerate with the mass function. Hence including cluster counts measurement breaks this degeneracy and allows for a much stronger  $\alpha_{\text{sat}}$  constraint.

Furthermore, the parameter constraints are this time improved in the combined realistic case compared to the naive independent hypothesis, despite the joint probe having a lower SNR (see table 1). This is the opposite of the results on the cosmological constraints in the cosmology case, as reported in section 8.2.<sup>4</sup> This is evidence for a synergy of the probes and the breaking of degeneracies. Indeed, in the independent hypothesis, the increased HOD constraints, compared to  $C_{\ell}^{\text{gal}}$  alone, come only from the improvement of cosmological constraints which are then marginalised over.

## 9 Conclusion

With the advent of large surveys such as the Dark Energy Survey, which provides us with a rich number of different observables, comes the issue of combining probes to obtain stronger constraints. In this article, we developed the framework to combine two particular probes of galaxy surveys: the cluster number counts and the galaxy angular power spectrum. We worked in full sky and did not consider a number of experimental effects such as photometric redshift errors of galaxies and clusters, purity and completeness of cluster detections, and scatter in the cluster mass determination. However, we chose in our analysis conservative values for the width of redshift and mass bins that should mitigate the impact of these experimental limitations. We thus expect the inclusion of these effects to be a correction to our results, and hence it should not change qualitatively our conclusions.

We used the halo model and Halo Occupation Distribution to model both the cluster counts, the galaxy spectrum, their respective auto-covariances, and for the first time the cross-covariance between these two observables, accounting for the important non-Gaussian contributions to the covariances. In particular we took into account second order halo bias, galaxy shot noise, and showed that halo non-local bias gives a null contribution to the cross-covariance. We also discussed in details the case of the super-sample covariance and showed that it obeys general mathematical inequalities whatever the probes considered.

We presented the joint covariance matrix of these observables for two cases focussing either on large or small scales. In both cases we pointed out the importance of non-Gaussian contribution to the covariance matrix, which become irretrievably dominant at low redshifts and/or small scales. A careful modeling of non-linearities is thus essential if one wants to extract useful information in this regime. Furthermore we showed that a fully non-linear model is required to reproduce the cross-covariance in the full range needed, as perturbation theory alone fails on small scales, low redshifts and/or high cluster masses. A non-perturbative model is thus essential to combine these two probes.

We then used our predicted covariance matrix to analyse the consequences of combining both observables, showing first the increase of signal-to-noise with respect to the single probe case. Next we conducted a Fisher matrix analysis in the two cases: first concentrating on large scales to constrain cosmological parameters while marginalising over HOD parameters, and second focussing on small scales to constrain HOD parameters while marginalising over cosmological parameters. For cosmology, compared to the best single-probe constraints, improvements of the order of 20% can be expected by including the combination. For HOD, although it is not constrained by cluster counts alone, the inclusion of cluster counts breaks

---

<sup>4</sup>In fact we checked that the improvement scenario (i.e. constraints being better in the realistic case than in the hypothetical independent case) is the most frequent. For instance, HOD constraints are also improved in the cosmological study case and the same goes for cosmological constraints in the HOD study case. However we refrain from reporting those results as they are unrealistic for actual data analysis.

the degeneracy between the mass function and HOD. This results in a dramatic improvement of the bounds on HOD parameters, reducing the error bars by a factor of 1.4 to 4.8 depending on the parameter.

These results demonstrate the critical importance of combining these observables in deriving constraints on cosmology and HOD.

Finally, compared to an hypothetical independent case, we showed that the consistent inclusion of the cross-correlation of the probes can lead to a few percent degradation in the cosmological constraints, but it leads to a large improvement of the HOD constraints, up to 50% in the case of  $\alpha_{\text{sat}}$ . This shows conclusively the importance of accounting for the cross-covariance in the combination of these probes.

## A 3D halo-galaxy-galaxy bispectrum in the Halo Model

The different terms of the halo-galaxy-galaxy bispectrum can be derived from eq. (2.1) & eq. (2.4) through lengthy and error-prone calculations. Alternatively, [33] have shown that the galaxy polyspectra equations can be derived via an elegant diagrammatic formalism with ‘Feynman’-like rules. Here we follow this approach and extend it to cross-polyspectra with halo statistics in given mass bins.

For the halo-galaxy-galaxy bispectrum, the first step is to draw all the diagrams involving one halo for the cluster count, labeled “1”, and two possibly merged galaxies for the galaxy spectrum, labeled “2” and “3”. Merging occurs for shot-noise diagrams. These diagrams are shown in figure 4.

Then a diagram contributes as a term to the hgg bispectrum given by:

- a prefactor  $1/(\bar{n}_{\text{gal}}(z_2) \bar{n}_{\text{gal}}(z_3))$
- a possible Kronecker of redshifts if some halos coincide, e.g.  $\delta_{z_1, z_2}$  if the first galaxy (labeled “2”) is in the halo of the number count (labeled “1”).
- integrals over the halo mass  $\int dM \frac{dn_h}{dM}$  for the halos containing the galaxies, except if that is the halo of the number count (since that one has a fixed mass  $M_1$ ).
- average number of uplet of galaxy given by the HOD, e.g.  $\langle N_{\text{gal}}(M) \rangle$  for a single galaxy in a halo,  $\langle N(N-1) \rangle$  for a pair.
- for each galaxy, halo profile  $u(k|M)$  at the mass of the host halo, with  $k$  being the norm of the sum of Fourier wavevectors over the galaxy labels, i.e. if the galaxy has label “2”  $k = k_2$ , but for a shot-noise diagram  $k = |\mathbf{k}_2 + \mathbf{k}_3| = k_1$
- the halo polyspectrum of the considered halos, conditioned to their mass and redshift
- a possible permutation of indices to account for symmetries. e.g. in the 2h-2h diagram of figure 4, we have put galaxy “2” in the halo “1” of counts, but there is the symmetric diagram where galaxy “3” is in the halo “1”.

Note that this diagrammatic formalism and these rules allow to compute any polyspectrum of halos and galaxies, though we just need the hgg bispectrum in this article. We note that the complexity increases very quickly with the polyspectrum order considered, with e.g. the hggg trispectrum requiring 14 different terms, so that this diagrammatic becomes increasingly useful with order to avoid lengthy and cumbersome calculations.

Applying this diagrammatic formalism to the hgg bispectrum, we get 6 different diagrams/terms with the following expressions:

$$\begin{aligned}
 B^{3h}(k_{123}, z_{123}|M_1, z_1) &= \frac{1}{\bar{n}_{\text{gal}}(z_2) \bar{n}_{\text{gal}}(z_3)} \int dM_{23} \left. \frac{d^2 n_h}{dM} \right|_{M_2, z_2} \left. \frac{d^2 n_h}{dM} \right|_{M_3, z_3} \\
 &\times \langle N_{\text{gal}}(M_2) \rangle \langle N_{\text{gal}}(M_3) \rangle u(k_2|M_2) u(k_3|M_3) \\
 &\times B_{\text{halo}}(k_{123}|M_{123}, z_{123}) \quad (\text{A.1})
 \end{aligned}$$

$$\begin{aligned}
 B^{2h-2h}(k_{123}, z_{123}|M_1, z_1) &= \frac{\delta_{z_1, z_2}}{\bar{n}_{\text{gal}}(z_2) \bar{n}_{\text{gal}}(z_3)} \int dM_3 \left. \frac{d^2 n_h}{dM} \right|_{M_3, z_3} \langle N_{\text{gal}}(M_1) \rangle \langle N_{\text{gal}}(M_3) \rangle \\
 &\times u(k_2|M_1) u(k_3|M_3) P_{\text{halo}}(k_3|M_{13}, z_{13}) + \text{perm. } (2 \leftrightarrow 3) \quad (\text{A.2})
 \end{aligned}$$

$$\begin{aligned}
 B^{2h-1h2g}(k_{123}, z_{123}|M_1, z_1) &= \frac{\delta_{z_2, z_3}}{\bar{n}_{\text{gal}}(z_2) \bar{n}_{\text{gal}}(z_3)} \int dM \frac{d^2 n_h}{dM} \langle N_{\text{gal}}(N_{\text{gal}} - 1)(M) \rangle \\
 &\times u(k_2|M) u(k_3|M) P_{\text{halo}}(k_1|M_1, M, z_1, z_2) \quad (\text{A.3})
 \end{aligned}$$

$$\begin{aligned}
 B^{2h-1h1g}(k_{123}, z_{123}|M_1, z_1) &= \frac{\delta_{z_2, z_3}}{\bar{n}_{\text{gal}}(z_2) \bar{n}_{\text{gal}}(z_3)} \int dM \frac{d^2 n_h}{dM} \langle N_{\text{gal}}(M) \rangle \\
 &\times u(k_1|M) P_{\text{halo}}(k_1|M_1, M, z_1, z_2) \quad (\text{A.4})
 \end{aligned}$$

$$\begin{aligned}
 B^{1h2g}(k_{123}, z_{123}|M_1, z_1) &= \frac{\delta_{z_1, z_2, z_3}}{\bar{n}_{\text{gal}}(z_2) \bar{n}_{\text{gal}}(z_3)} \langle N_{\text{gal}}(N_{\text{gal}} - 1)(M_1) \rangle u(k_2|M_1) u(k_3|M_1) \\
 &\quad (\text{A.5})
 \end{aligned}$$

$$\begin{aligned}
 B^{1h1g}(k_{123}, z_{123}|M_1, z_1) &= \frac{\delta_{z_1, z_2, z_3}}{\bar{n}_{\text{gal}}(z_2) \bar{n}_{\text{gal}}(z_3)} \langle N_{\text{gal}}(M_1) \rangle u(k_1|M_1) \quad (\text{A.6})
 \end{aligned}$$

where  $\delta_{z_1, z_2, z_3} = \delta_{z_1, z_2} \delta_{z_2, z_3}$ .

Note that the  $2h-1h1g$  and  $1h1g$  terms are shot-noise terms and are thus irrelevant for the covariance with the real-space ACF (as shot-noise, corresponding to  $\theta = 0$ , is explicitly not considered), but they are relevant for the covariance with the power spectrum.

Eq. (A.1) involves the halo bispectrum, which we compute at tree-level using the following equation for the halo density:

$$\delta_h(\mathbf{x}|M, z) = b_1(M, z) \delta_{\text{DM}}(\mathbf{x}, z) + \frac{b_2(M, z)}{2!} (\delta_{\text{DM}}(\mathbf{x}, z)^2 - \langle \delta_{\text{DM}}^2 \rangle) + b_{s2}(M, z) (s^2(\mathbf{x}, z) - \langle s^2 \rangle) \quad (\text{A.7})$$

where  $s^2$  is the square of the tidal tensor and we have [31]:

$$s^2(\mathbf{k}, z) = \int \frac{d\mathbf{k}'}{(2\pi)^3} S_2(\mathbf{k}', \mathbf{k} - \mathbf{k}') \delta_{\text{DM}}(\mathbf{k}', z) \delta_{\text{DM}}(\mathbf{k} - \mathbf{k}', z) \quad \text{with} \quad S_2(\mathbf{k}_1, \mathbf{k}_2) = \frac{(\mathbf{k}_1 \cdot \mathbf{k}_2)^2}{k_1^2 k_2^2} - \frac{1}{3} \quad (\text{A.8})$$

The resulting halo bispectrum has the expression:

$$\begin{aligned}
 B_{\text{halo}}(k_{123}|M_{123}, z_{123}) &= b_1(M_1, z_1) b_1(M_2, z_2) b_1(M_3, z_3) B_{\text{DM}}(k_{123}, z_{123}) \\
 &+ [b_2(M_1, z_1) b_1(M_2, z_2) b_1(M_3, z_3) \\
 &\times P_{\text{DM}}(k_2|z_{12}) P_{\text{DM}}(k_3|z_{13}) + 2 \text{ perm.}] \\
 &+ [2 b_{s2}(M_1, z_1) b_1(M_2, z_2) b_1(M_3, z_3) \\
 &\times S_2(\mathbf{k}_2, \mathbf{k}_3) P_{\text{DM}}(k_2|z_{12}) P_{\text{DM}}(k_3|z_{13}) + 2 \text{ perm.}] \quad (\text{A.9})
 \end{aligned}$$

where the dark matter bispectrum is given at second-order in perturbation theory:

$$B_{\text{DM}}(k_{123}, z_{123}) = 2F^{2\text{PT}}(\mathbf{k}_1, \mathbf{k}_2) P_{\text{lin}}(k_1|z_1, z_3) P_{\text{lin}}(k_2|z_2, z_3) + 2 \text{ perm.} \quad (\text{A.10})$$

(adapted from [32] for different redshifts) with

$$F^{2\text{PT}}(\mathbf{k}_\alpha, \mathbf{k}_\beta) = \frac{5}{7} + \frac{1}{2} \cos \theta_{\alpha\beta} \left( \frac{k_\alpha}{k_\beta} + \frac{k_\beta}{k_\alpha} \right) + \frac{2}{7} \cos^2 \theta_{\alpha\beta}. \quad (\text{A.11})$$

## B From the 3D hgg bispectrum to the cross-covariance

From eq. (5.1), the definition of the halo density contrast and the halo-galaxy-galaxy (hgg) bispectrum comes:

$$\begin{aligned} \text{Cov} \left( N_{\text{cl}}(i_M, i_z), C_\ell^{\text{gal}}(j_z) \right) &= 4\pi \int dM_1 dV_{123} \frac{\bar{n}_{\text{gal}}(z_2) \bar{n}_{\text{gal}}(z_3)}{\Delta N_{\text{gal}}(j_z)^2} \frac{dn_h}{dM} \Big|_{M_1, z_1} \\ &\quad \times \int \frac{d^3 \mathbf{k}_{123}}{(2\pi)^9} j_0(k_1 r_1) j_\ell(k_2 r_2) j_\ell(k_3 r_3) P_\ell(\hat{k}_2 \cdot \hat{k}_3) \\ &\quad \times (2\pi)^3 \delta^3(\mathbf{k}_1 + \mathbf{k}_2 - \mathbf{k}_3) B_{\text{hgg}}(\mathbf{k}_{123}|M_1, z_{123}) \end{aligned} \quad (\text{B.1})$$

We note that this can also be written as:

$$\text{Cov} \left( N_{\text{cl}}(i_M, i_z), C_\ell^{\text{gal}}(j_z) \right) = 4\pi \int dM_1 dV_{123} \frac{dn_h}{dM} \Big|_{M_1, z_1} \frac{\bar{n}_{\text{gal}}(z_2) \bar{n}_{\text{gal}}(z_3)}{\Delta N_{\text{gal}}(j_z)^2} b_{0\ell\ell}^{\text{hgg}}(M_1, z_{123}) \quad (\text{B.2})$$

where  $b^{\text{hgg}}$  is the halo-galaxy-galaxy angular bispectrum per comoving volumes and unit mass. This bispectrum is involved in the squeezed configuration  $\ell_1 \ll \ell_2, \ell_3$ , in fact the most squeezed possible:  $\ell_1 = 0$ . Hence we will have contributions in the integral only from triplets of Fourier modes satisfying  $k_1 \ll k_2 \approx k_3$ .

Now the thirteen integrals of eq. (B.1) are not numerically tractable and thus need to be simplified analytically. To shorten further equations, we are going to first concentrate on the simplification of  $b_{0\ell\ell}^{\text{hgg}}(M_1, z_{123})$ :

$$\begin{aligned} b_{0\ell\ell}^{\text{hgg}}(M_1, z_{123}) &= \int \frac{d^3 \mathbf{k}_{123}}{(2\pi)^9} j_0(k_1 r_1) j_\ell(k_2 r_2) j_\ell(k_3 r_3) P_\ell(\hat{k}_2 \cdot \hat{k}_3) \\ &\quad \times (2\pi)^3 \delta^3(\mathbf{k}_1 + \mathbf{k}_2 - \mathbf{k}_3) B_{\text{hgg}}(k_{123}|M_1, z_{123}) \end{aligned} \quad (\text{B.3})$$

We are first going to complexify this equation by writing the Legendre polynomial as sum of spherical harmonics with the addition theorem, introduce the Fourier representation of the dirac, and expand the resulting Fourier modes in terms of Bessel functions and spherical harmonics:

$$\begin{aligned} b_{0\ell\ell}^{\text{hgg}}(M_1, z_{123}) &= \int \frac{d^3 \mathbf{k}_{123}}{(2\pi)^9} j_0(k_1 r_1) j_\ell(k_2 r_2) j_\ell(k_3 r_3) \frac{4\pi}{2\ell+1} \sum_m Y_{\ell m}^*(\hat{k}_2) Y_{\ell m}(\hat{k}_3) \\ &\quad \times B_{\text{hgg}}(\mathbf{k}_{123}|M_1, z_{123}) \int d^3 \mathbf{x} (4\pi)^3 \sum_{1,2,3} i^{\ell_1+\ell_2-\ell_3} j_{\ell_1}(k_1 x) j_{\ell_2}(k_2 x) j_{\ell_3}(k_3 x) \\ &\quad \times Y_1(\hat{k}_1) Y_2(\hat{k}_2) Y_3^*(\hat{k}_3) Y_1^*(\hat{x}) Y_2^*(\hat{x}) Y_3(\hat{x}) \end{aligned} \quad (\text{B.4})$$

which sure looks better.



Now we are faced with an alternative depending on the bispectrum term considered: either the bispectrum term only depends on the moduli  $(k_1, k_2, k_3)$  (case of most bispectrum terms), or it also depends on the angles between the Fourier vectors  $(\hat{k}_1 \cdot \hat{k}_2, \hat{k}_1 \cdot \hat{k}_3, \hat{k}_2 \cdot \hat{k}_3)$  (case of the 2PT and s2 terms). We are going to treat these cases in separate subsections.

### B.1 Angle-independent bispectrum

We first note that the derivation below is essentially a copy of the derivation of appendix E of [33] in the particular case of the bispectrum in the squeezed limit  $0\ell\ell$ , except that we will not apply a Limber approximation on  $k_1$ .

In the angle-independent case,  $B_{\text{hgg}}(\mathbf{k}_{123}|M_1, z_{123}) = B_{\text{hgg}}(k_{123}|M_1, z_{123})$ . Inputting this in eq. (B.4), we see that the integrals over the unit vectors  $\hat{k}_1, \hat{k}_2, \hat{k}_3, \hat{x}$  can be trivially performed as they only involve spherical harmonics. Specifically the integrals over  $\hat{k}_1, \hat{k}_2, \hat{k}_3$  force  $(\ell_1, m_1) = (0, 0)$ ,  $(\ell_2, m_2) = (\ell, m)$  and  $(\ell_3, m_3) = (\ell, m)$ . Taking care that  $Y_{00} = 1/\sqrt{4\pi}$ , the integral over  $\hat{x}, \hat{k}_1$  and the sum over  $m$  are trivial and we get:

$$b_{0\ell\ell}^{\text{hgg}}(M_1, z_{123}) = \frac{\bar{n}_{\text{gal}}(z_2) \bar{n}_{\text{gal}}(z_3)}{\Delta N_{\text{gal}}(j_z)^2} \left( \frac{1}{2\pi^2} \right)^3 \int k_{123}^2 dk_{123} 4\pi x^2 dx j_0(k_1 r_1) j_\ell(k_2 r_2) j_\ell(k_3 r_3) \\ \times j_0(k_1 x) j_\ell(k_2 x) j_\ell(k_3 x) B_{\text{hgg}}(k_{123}|M_1, z_{123}) \quad (\text{B.5})$$

We now apply the Limber approximation on  $k_2$  and  $k_3$ , that is to assume that the bispectrum varies slowly with these arguments compared to the oscillations of the corresponding Bessel functions. We can thus take out the bispectrum of these integrals, evaluating it at the Bessel peak  $(\ell + 1/2)/r_{2/3}$ , and use the identity:

$$\int k^2 dk j_\ell(kr) j_\ell(kx) = \frac{\pi}{2r^2} \delta(r - x) \quad (\text{B.6})$$

This greatly simplifies the bispectrum:

$$b_{0\ell\ell}^{\text{hgg}}(M_1, z_{123}) = \frac{\bar{n}_{\text{gal}}(z_2) \bar{n}_{\text{gal}}(z_3)}{\Delta N_{\text{gal}}(j_z)^2} \frac{1}{(2\pi)^3} \int k_1^2 dk_1 x^2 dx j_0(k_1 r_1) j_0(k_1 x) B_{\text{hgg}}(k_1, k_\ell, k_\ell|M_1, z_{123}) \\ \times \frac{\delta(r_2 - x) \delta(r_3 - x)}{r_2^2 r_3^2} \quad (\text{B.7})$$

where  $k_\ell = (\ell + 1/2)/r_2$  (since  $r_2 = r_3 = x$ ).

Inputting this in eq. (B.2) for the covariance gives:

$$\text{Cov} \left( N_{\text{cl}}(i_M, i_z), C_\ell^{\text{gal}}(j_z) \right) = 4\pi \int dM_1 dV_{123} dV_x \left. \frac{dn_h}{dM} \right|_{M_1, z_1} \frac{\bar{n}_{\text{gal}}(z_2) \bar{n}_{\text{gal}}(z_3)}{\Delta N_{\text{gal}}(j_z)^2} \frac{\delta(r_2 - x) \delta(r_3 - x)}{r_2^2 r_3^2} \\ \times \frac{1}{(2\pi)^3} \int k^2 dk j_0(kr_1) j_0(kx) B_{\text{hgg}}(k, k_\ell, k_\ell|M_1, z_{123}) \quad (\text{B.8})$$

$$= \int dM_1 dV_1 dV_x \left. \frac{dn_h}{dM} \right|_{M_1, z_1} \frac{\bar{n}_{\text{gal}}(z_x)^2}{\Delta N_{\text{gal}}(j_z)^2} \\ \times \frac{1}{2\pi^2} \int k^2 dk j_0(kr_1) j_0(kx) B_{\text{hgg}}(k, k_\ell, k_\ell|M_1, z_1, z_x, z_x) \quad (\text{B.9})$$

where  $x$  has been considered has the comoving distance to a redshift  $z_x$ ,  $k_\ell = (\ell + 1/2)/x$  and the integral over  $z_x$  implicitly runs over the redshift bin  $j_z$ .

This equation allows to compute the cross-covariance from the 2h-1h1g, and 2h-1h2g bispectrum terms (eq. (A.4)–(A.3)), giving explicitly:

$$\begin{aligned} \text{Cov}_{2\text{h}-1\text{h}1\text{g}} \left( N_{\text{cl}}(i_M, i_z), C_\ell^{\text{gal}}(j_z) \right) &= \frac{1}{\Delta N_{\text{gal}}(j_z)^2} \int dV_{12} \frac{\partial n_h}{\partial \delta_b}(i_M, z_1) \\ &\times \bar{n}_{\text{gal}}(z_2) b_1^{\text{gal,eff}}(z_2) \sigma_{\text{proj}}^2(z_1, z_2) \end{aligned} \quad (\text{B.10})$$

$$\begin{aligned} \text{Cov}_{2\text{h}-1\text{h}2\text{g}} \left( N_{\text{cl}}(i_M, i_z), C_\ell^{\text{gal}}(j_z) \right) &= \frac{1}{\Delta N_{\text{gal}}(j_z)^2} \int dV_{12} \frac{\partial n_h}{\partial \delta_b}(i_M, z_1) \\ &\times \bar{n}_{\text{gal}}(z_2) b_1^{\text{gal,eff}-2}(k_\ell, z_2) \sigma_{\text{proj}}^2(z_1, z_2) \end{aligned} \quad (\text{B.11})$$

where in the 2h-1h1g term, we approximated  $u(k_1|M) \approx 1$ , as  $k_1$  gets aliased into the monopole. That is, no considered halo takes up the whole sky. This seems a fair approximation as none of our observables is sensitive to the local group.

Eq. (B.9) also allow to compute the cross-covariance from the b2 subterm of the 3h bispectrum. Recalling

$$B_{\text{halo}}^{\text{b2}}(k_{123}|M_{123}, z_{123}) = b_2(M_1, z_1) b_1(M_2, z_2) b_1(M_3, z_3) P_{\text{DM}}(k_2|z_{12}) P_{\text{DM}}(k_3|z_{13}) + 2 \text{perm.} \quad (\text{B.12})$$

and given that  $k_1 \ll k_2, k_3$  and that  $P(k)$  is decreasing over the range of interest, we can assume that  $P(k_2)P(k_3) \ll P(k_1)P(k_2), P(k_1)P(k_3)$  and neglect the corresponding term. This yields:

$$\begin{aligned} \text{Cov}_{\text{b2}} \left( N_{\text{cl}}(i_M, i_z), C_\ell^{\text{gal}}(j_z) \right) &= \frac{2}{\Delta N_{\text{gal}}(j_z)^2} \int dV_{12} \frac{\partial n_h}{\partial \delta_b}(i_M, z_1) \bar{n}_{\text{gal}}(z_2)^2 \\ &\times b_2^{\text{gal,eff}}(k_\ell, z_2) b_1^{\text{gal,eff}}(k_\ell, z_2) P_{\text{DM}}(k_\ell|z_2) \sigma_{\text{proj}}^2(z_1, z_2) \end{aligned} \quad (\text{B.13})$$

Some bispectrum terms do not depend on  $k_1$  (2h-2h and 1h2g terms), or only through  $u(k_1|M, z)$  (1h1g term) that we approximate as 1 as argued previously. In that case, the integral over  $k$  in eq. (B.9) can be performed analytically and yields:

$$\text{Cov} \left( N_{\text{cl}}(i_M, i_z), C_\ell^{\text{gal}}(j_z) \right) = \delta_{i_z, j_z} \int dM_1 dV \left. \frac{dn_h}{dM} \right|_{M_1, z} \frac{\bar{n}_{\text{gal}}(z)^2}{\Delta N_{\text{gal}}(j_z)^2} \frac{1}{4\pi} B_{\text{hgg}}(k, k_\ell, k_\ell|M_1, z) \quad (\text{B.14})$$

This equation allows to compute the cross-covariance from the 2h-2h, 1h2g and 1h1g bispectrum terms (eq. (A.2)–(A.5)–(A.6)), giving:

$$\text{Cov}_{1\text{h}1\text{g}} \left( N_{\text{cl}}(i_M, i_z), C_\ell^{\text{gal}}(j_z) \right) = \frac{\delta_{i_z, j_z}}{4\pi \Delta N_{\text{gal}}(j_z)^2} \int dV n_{\text{gal}}(z|i_M) \quad (\text{B.15})$$

$$\text{Cov}_{1\text{h}2\text{g}} \left( N_{\text{cl}}(i_M, i_z), C_\ell^{\text{gal}}(j_z) \right) = \frac{\delta_{i_z, j_z}}{4\pi \Delta N_{\text{gal}}(j_z)^2} \int dV \bar{n}_{\text{gal}}(z)^2 P_{\text{gal}}^{\text{1h}}(k_\ell, z|i_M) \quad (\text{B.16})$$

$$\begin{aligned} \text{Cov}_{2\text{h}-2\text{h}} \left( N_{\text{cl}}(i_M, i_z), C_\ell^{\text{gal}}(j_z) \right) &= \frac{\delta_{i_z, j_z}}{4\pi \Delta N_{\text{gal}}(j_z)^2} \int dV 2 \bar{n}_{\text{gal}}(z)^2 b_1^{\text{gal,eff}}(k_\ell, z) \\ &\times b_1^{\text{gal,eff}}(k_\ell, z|i_M) P_{\text{DM}}(k_\ell, z) \end{aligned} \quad (\text{B.17})$$

with

$$n_{\text{gal}}(z|i_M) = \int_{M \in \text{bin}(i_M)} dM \frac{dn_h}{dM} \langle N_{\text{gal}}(M) \rangle \quad (\text{B.18})$$

$$P_{\text{gal}}^{\text{1h}}(k, z|i_M) = \frac{1}{\bar{n}_{\text{gal}}^2(z)} \int_{M \in \text{bin}(i_M)} dM \frac{dn_h}{dM} \langle N_{\text{gal}}(N_{\text{gal}} - 1)(M) \rangle u^2(k|M, z) \quad (\text{B.19})$$

$$b_1^{\text{gal,eff}}(k, z|i_M) = \frac{1}{\bar{n}_{\text{gal}}(z)} \int_{M \in \text{bin}(i_M)} dM \frac{dn_h}{dM} \langle N_{\text{gal}}(M) \rangle b_1(M, z) u(k|M, z) \quad (\text{B.20})$$

being the contribution of the  $i_M$  mass bin to the number of galaxy, 1-halo power spectrum and effective galaxy bias.

Now are left only the 2PT and s2 bispectrum terms which both depend on the angle between the Fourier modes and are thus the subject of the next subsection.

## B.2 Angle dependent bispectrum

As a foreword, one could think that the bispectrum could be parametrised only by the moduli  $k_1, k_2, k_3$ . Indeed the angles between the vectors can be computed from the moduli, given that  $(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)$  form a triangle. This proves not to be a safe bet, because angles are fast varying functions of the moduli in the squeezed limit  $k_1 \ll k_2, k_3$ , so we would never be in the position to apply the Limber's approximation used previously. Thus in the following, we keep the angle-dependence of the bispectrum in the equations.

The 2PT and s2 bispectrum terms are of the general form:

$$B_{\text{halo}}(k_{123}|M_{123}, z_{123}) = b_\alpha(M_1, z_1) b_1(M_2, z_2) b_1(M_3, z_3) \mathcal{K}(\mathbf{k}_2, \mathbf{k}_3) P_{\text{DM}}(k_2|z_{12}) P_{\text{DM}}(k_3|z_{13}) + 2 \text{ perm.} \quad (\text{B.21})$$

with  $\alpha = 1, \text{s2}$  and  $\mathcal{K} = 2F^{2\text{PT}}, S_2$  for the 2PT and s2 terms respectively.

As already noted previously for the b2 term,  $P(k_2)P(k_3) \ll P(k_1)P(k_2), P(k_1)P(k_3)$  so we can neglect the corresponding term in the permutation.

The dimensionless kernel  $\mathcal{K}$  is a function of the angle between the Fourier modes and can be decomposed onto Legendre polynomials:

$$\mathcal{K}(\mathbf{k}_1, \mathbf{k}_2) = \sum_{\ell_k} \mathcal{K}_{\ell_k} P_{\ell_k}(\cos \theta_{12}) \quad (\text{B.22})$$

For example for the  $F^{2\text{PT}}$  kernel the sum runs over  $\ell_F = 0, 1, 2$  and we have

$$F_0^{2\text{PT}} = \frac{17}{21} \quad F_1^{2\text{PT}} = \frac{1}{2} \left( \frac{k_1}{k_2} + \frac{k_2}{k_1} \right) \quad F_2^{2\text{PT}} = \frac{2}{7} \quad (\text{B.23})$$

while for the  $S_2$  kernel, the sum runs only over  $\ell_S = 2$ , all other terms being zero, and we have

$$S_{2,2} = 1 \quad (\text{B.24})$$

Now we need to compute the effect of the  $\mathcal{K}_{\ell_k}$  term. Before diving into this, let us first note that the case of  $\ell_K = 0$  is in fact simple: as  $P_0 = 1$ , we are in fact back to the angle-independent case. Through section B.1 we know how to tackle this, and we get:

$$\begin{aligned} \text{Cov}_{\mathcal{K}_0} \left( N_{\text{cl}}(i_M, i_z), C_\ell^{\text{gal}}(j_z) \right) &= \frac{2}{\Delta N_{\text{gal}}(j_z)^2} \int dV_{12} \frac{\partial n_h}{\partial \delta_b}(i_M, z_1) \bar{n}_{\text{gal}}(z_2)^2 \\ &\times \mathcal{K}_0 b_\alpha^{\text{gal,eff}}(k_\ell, z_2) b_1^{\text{gal,eff}}(k_\ell, z_2) P_{\text{DM}}(k_\ell|z_2) \sigma_{\text{proj}}^2(z_1, z_2) \end{aligned} \quad (\text{B.25})$$

For the s2 term  $\mathcal{K}_0 = 0$  while for the 2PT term  $\mathcal{K}_0 = 2 \times \frac{17}{21}$ . Thus:

$$\text{Cov}_{s2,0} \left( N_{\text{cl}}(i_M, i_z), C_\ell^{\text{gal}}(j_z) \right) = 0 \quad (\text{B.26})$$

$$\begin{aligned} \text{Cov}_{2\text{PT},0} \left( N_{\text{cl}}(i_M, i_z), C_\ell^{\text{gal}}(j_z) \right) &= \frac{1}{\Delta N_{\text{gal}}(j_z)^2} \int dV_{12} \frac{\partial n_h}{\partial \delta_b}(i_M, z_1) \bar{n}_{\text{gal}}(z_2)^2 \\ &\times \frac{68}{21} b_1^{\text{gal,eff}}(k_\ell, z_2) b_1^{\text{gal,eff}}(k_\ell, z_2) P_{\text{DM}}(k_\ell|z_2) \sigma_{\text{proj}}^2(z_1, z_2) \end{aligned} \quad (\text{B.27})$$

which correspond to the results stated in the main text for the s2 and 2PT terms.

The rest of this subsection is devoted to argue that the contribution of higher  $\ell_K$  is negligible (compared to the 2PT,0 term).

To do this we need to go back to the full generality of eq. (B.4) for the angular bispectrum:

$$\begin{aligned} b_{0\ell\ell}^{\mathcal{K}_{\ell_k}}(M_1, z_{123}) &= \int \frac{d^3 \mathbf{k}_{123}}{(2\pi)^9} j_0(k_1 r_1) j_\ell(k_2 r_2) j_\ell(k_3 r_3) \frac{4\pi}{2\ell+1} \sum_m Y_{\ell m}^*(\hat{k}_2) Y_{\ell m}(\hat{k}_3) \\ &\times B_{\text{hgg}}^{\mathcal{K}_{\ell_k}}(\mathbf{k}_{123}|M_1, z_{123}) \int d^3 \mathbf{x} (4\pi)^3 \sum_{1,2,3} i^{\ell_1+\ell_2-\ell_3} j_{\ell_1}(k_1 x) j_{\ell_2}(k_2 x) j_{\ell_3}(k_3 x) \\ &\times Y_1(\hat{k}_1) Y_2(\hat{k}_2) Y_3^*(\hat{k}_3) Y_1^*(\hat{x}) Y_2^*(\hat{x}) Y_3(\hat{x}) \end{aligned} \quad (\text{B.28})$$

Now we will drop integrals over masses in the bispectrum, as the argument is based on the integrals of spherical harmonics and spherical bessel function, and the equations are already long enough. We also drop the permutation  $2 \leftrightarrow 3$  and some redshift dependences for clarity.

Basically the angular bispectrum contains:

$$\begin{aligned} b_{0\ell\ell}^{\mathcal{K}_{\ell_k}}(M_1, z_{123}) &\ni \int \frac{d^3 \mathbf{k}_{123}}{(2\pi)^9} j_0(k_1 r_1) j_\ell(k_2 r_2) j_\ell(k_3 r_3) \frac{4\pi}{2\ell+1} \sum_m Y_{\ell m}^*(\hat{k}_2) Y_{\ell m}(\hat{k}_3) \\ &\times P_{\ell_K}(\cos \theta_{12}) P(k_1) P(k_2) \int d^3 \mathbf{x} (4\pi)^3 \sum_{1,2,3} i^{\ell_1+\ell_2-\ell_3} j_{\ell_1}(k_1 x) j_{\ell_2}(k_2 x) j_{\ell_3}(k_3 x) \\ &\times Y_1(\hat{k}_1) Y_2(\hat{k}_2) Y_3^*(\hat{k}_3) Y_1^*(\hat{x}) Y_2^*(\hat{x}) Y_3(\hat{x}) \end{aligned} \quad (\text{B.29})$$

Now we can expand the Legendre polynomial onto spherical harmonics:

$$P_{\ell_K}(\cos \theta_{12}) = \frac{4\pi}{2\ell_K+1} \sum_{m_K} Y_{\ell_K, m_K}^*(\hat{k}_1) Y_{\ell_K, m_K}(\hat{k}_2) \quad (\text{B.30})$$

Then we can perform the integrals over  $\hat{k}_1, \hat{k}_2, \hat{k}_3$ . First the integral over  $\hat{k}_3$  forces  $(\ell_3, m_3) = (\ell, m)$  by orthogonality of spherical harmonics. In the same way, the integral over  $\hat{k}_1$  forces  $(\ell_1, m_1) = (\ell_K, m_K)$ . Finally the integral over  $\hat{k}_2$  is non-trivial and yields a Gaunt coefficient:

$$G(\ell, m, \ell_K, m_K, \ell_2, m_2) = \int d^2 \hat{k}_2 Y_{\ell m}^*(\hat{k}_2) Y_{\ell_K, m_K}(\hat{k}_2) Y_{\ell_2, m_2}(\hat{k}_2) \quad (\text{B.31})$$

This forces  $-m + m_K + m_2 = 0$  and  $|\ell - \ell_K| \leq \ell_2 \leq \ell + \ell_K$  with  $\ell + \ell_K + \ell_2$  even.

The integral over  $\hat{x}$  yields the same Gaunt coefficient (more precisely its complex conjugate, but Gaunt coefficients are real anyway). Thus:

$$b_{0\ell\ell}^{\mathcal{K}_{\ell_K}}(M_1, z_{123}) \ni \frac{(4\pi)^5}{(2\pi)^9} \frac{1}{(2\ell+1)(2\ell_K+1)} \sum_{2,m,m_K} \int k_{123}^2 dk_{123} x^2 dx j_0(k_1 r_1) j_\ell(k_2 r_2) j_\ell(k_3 r_3) \\ \times P(k_1) P(k_2) i^{\ell_K+\ell_2-\ell} j_{\ell_K}(k_1 x) j_{\ell_2}(k_2 x) j_\ell(k_3 x) G^2(\ell, m, \ell_K, m_K, \ell_2, m_2) \quad (\text{B.32})$$

The sum over  $m_2, m, m_K$  can be carried out analytically and yields

$$\sum_{m_2, m, m_K} G^2(\ell, m, \ell_K, m_K, \ell_2, m_2) = \frac{(2\ell_2+1)(2\ell+1)(2\ell_K+1)}{4\pi} \begin{pmatrix} \ell & \ell_2 & \ell_K \\ 0 & 0 & 0 \end{pmatrix}^2 \quad (\text{B.33})$$

The integral over  $k_3$  is the already-used identity of integral of spherical Bessel functions and yields a dirac of redshifts  $z_3$  and  $z_x$ .

Now comes the pitch. For the cases of interest (2PT and s2 bispectrum terms),  $\ell_K = 1$  or 2. Thus,  $\ell_2 = \ell - 1, \ell + 1$  or  $\ell_2 = \ell - 2, \ell, \ell + 2$  respectively. Let's compare the  $k_1$  and  $k_2$  integrals to the  $\ell_K = 0$  case. In the  $\ell_K = 0$  case we have

$$\int k_1^2 dk_1 j_0(k_1 r_1) j_0(k_1 x) P(k_1) \quad (\text{B.34})$$

$$\int k_2^2 dk_2 j_\ell(k_2 r_2) j_\ell(k_2 x) P(k_2) \quad (\text{B.35})$$

while for a general  $\ell_K$  we have

$$\int k_1^2 dk_1 j_0(k_1 r_1) j_{\ell_K}(k_1 x) P(k_1) \quad (\text{B.36})$$

$$\int k_2^2 dk_2 j_\ell(k_2 r_2) j_{\ell_2}(k_2 x) P(k_2) \quad (\text{B.37})$$

We expect both integrals to be small compared to their relative counterparts in the  $\ell_K = 0$  case. Indeed, the Bessel functions will have destructing interferences. They will not have the same frequency, unless  $z_1 = z_x$  for the first integral ( $z_2 = z_x$  for the second) where the integrals of the  $\ell_K = 0$  case are maximal. And in the  $z_1 = z_x$  (resp.  $z_2 = z_x$ ) case, the Bessel interferences will still be destructive because:

- $j_0$  and  $j_{\ell_K}$  do not peak at the same location. And the same is true for  $j_\ell$  and  $j_{\ell_2}$ .
- the asymptotic oscillations<sup>5</sup> are out of phase.

We thus conclude that these terms are negligible compared to the  $\ell_K = 0$  case. Thus  $(2\text{PT}, 2) \ll (2\text{PT}, 0)$  and  $(2\text{PT}, 1) \ll (2\text{PT}, 0)$ . Given that the  $F^{2\text{PT}}$  and  $S_2$  kernels are both of order 1 and that the corresponding biases will also be of the same order,<sup>6</sup> we also conclude that  $(\text{s2}, 2) \ll (2\text{PT}, 0)$ .

<sup>5</sup> $j_n(y) \sim \sin(y - \frac{n\pi}{2})/y$  for  $y \gg n$ .

<sup>6</sup>In fact  $b_{s2} = -\frac{2}{7}(b_1 - 1)$  following [31].

## C SSC inequalities

This appendix contains the mathematical proof of the SSC inequalities stated in the main text, as well as a discussion of their consequences.

**Definition.** A bivariate symmetric function  $\kappa(x, y)$  is said positively separable if it is of the form

$$\kappa(x, y) = \sum_i \Psi_i(x) \Psi_i(y)$$

for some family of univariate functions  $\Psi_i$ .

**Lemma.** If  $\kappa(x, y)$  is positively separable, the bivariate form

$$\langle f, g \rangle = \int dx dy f(x) g(y) \kappa(x, y)$$

is a semi-inner product over real-valued functions.

*Proofs.* The form is obviously bilinear, and symmetric because of the symmetry of  $\kappa(x, y)$ .

Let's prove positivity, i.e.  $\langle f, f \rangle \geq 0$ :

$$\begin{aligned} \langle f, f \rangle &= \int dx dy f(x) f(y) \kappa(x, y) \\ &= \sum_i \int dx dy f(x) f(y) \Psi_i(x) \Psi_i(y) \\ &= \sum_i \left( \int f \Psi_i \right)^2 \\ &\geq 0 \end{aligned} \tag{C.1}$$

**Note.**  $\langle f, f \rangle = 0 \Leftrightarrow \int f \Psi_i \forall i$ . Thus we have positive definiteness, i.e.  $\langle f, g \rangle$  is a strict inner product, iff the family  $(\Psi_i)$  is a basis of real-valued functions.

**Consequence.** We have the Cauchy-Schwarz inequality [38]

$$|\langle f, g \rangle| \leq \|f\| \cdot \|g\|$$

**Application.** Recall the equations for super-sample covariance of  $C_\ell^{\text{gal}}(i_z)$  and  $N_{\text{cl}}(i_M, i_z)$ :

$$\text{Cov}_{\text{SSC}}(N_{\text{cl}}(i_M, i_z), N_{\text{cl}}(j_M, i_z)) = \int dV_{12} \frac{\partial n_h(i_M, z_1)}{\partial \delta_b} \frac{\partial n_h(j_M, z_2)}{\partial \delta_b} \sigma_{\text{proj}}^2(z_1, z_2) \tag{C.2}$$

$$\text{Cov}_{\text{SSC}}(N_{\text{cl}}(i_M, i_z), C_\ell^{\text{gal}}(j_z)) = \int dV_{12} \frac{\bar{n}_{\text{gal}}(z_2)^2}{\Delta N_{\text{gal}}(j_z)^2} \frac{\partial n_h(i_M, z_1)}{\partial \delta_b} \frac{\partial P_{\text{gal}}(k_\ell, z_2)}{\partial \delta_b} \sigma_{\text{proj}}^2(z_1, z_2) \tag{C.3}$$

$$\begin{aligned} \text{Cov}_{\text{SSC}}(C_\ell^{\text{gal}}(i_z), C_{\ell'}^{\text{gal}}(j_z)) &= \int dV_{12} \frac{\bar{n}_{\text{gal}}(z_1)^2 \bar{n}_{\text{gal}}(z_2)^2}{\Delta N_{\text{gal}}(i_z)^2 \Delta N_{\text{gal}}(j_z)^2} \frac{\partial P_{\text{gal}}(k_\ell, z_1)}{\partial \delta_b} \frac{\partial P_{\text{gal}}(k_{\ell'}, z_2)}{\partial \delta_b} \\ &\quad \times \sigma_{\text{proj}}^2(z_1, z_2) \end{aligned} \tag{C.4}$$

where  $dV = r^2(z) \frac{dr}{dz} dz$ , the redshift integrals are implicitly over the corresponding redshift bins, and

$$\sigma_{\text{proj}}^2(z_1, z_2) = \frac{1}{2\pi^2} \int k^2 dk j_0(kr_1) j_0(kr_2) P(k|z_{12})$$

Given that  $P(k|z_{12}) = G(z_1) G(z_2) P(k, z = 0)$  with  $G(z)$  the growth function,  $\sigma_{\text{proj}}^2(z_1, z_2)$  is positively separable with

$$\sum_i \rightarrow \frac{1}{2\pi^2} \int k^2 dk \quad \text{and} \quad \Psi_i \rightarrow j_0(kr(z)) G(z) (P(k, z = 0))^{1/2}$$

We are thus in position to apply the lemma, with e.g.

$$f(z) = r^2(z) \frac{dr}{dz} \frac{\partial n_h(i_M, z)}{\partial \delta_b} \mathbb{1}_{[z \in \text{bin}(i_z)]} \quad \text{and} \quad g(z) = r^2(z) \frac{dr}{dz} \frac{\partial n_h(j_M, z)}{\partial \delta_b} \mathbb{1}_{[z \in \text{bin}(j_z)]}$$

giving

$$|\text{Cov}_{\text{SSC}}(N_{\text{cl}}(i_M, i_z), N_{\text{cl}}(j_M, j_z))| \leq [\text{Var}_{\text{SSC}}(N_{\text{cl}}(i_M, i_z)) \cdot \text{Var}_{\text{SSC}}(N_{\text{cl}}(j_M, j_z))]^{1/2} \quad (\text{C.5})$$

In particular, for a given redshift slice  $i_z = j_z$ , this inequality ensures that the correlation matrix of cluster counts has off-diagonal elements  $\leq 1$  even in the deep SSC domination regime.

The same process can be applied to the auto-covariance of the galaxy power spectrum, yielding:

$$|\text{Cov}_{\text{SSC}}(C_\ell^{\text{gal}}(i_z), C_{\ell'}^{\text{gal}}(j_z))| \leq [\text{Var}_{\text{SSC}}(C_\ell^{\text{gal}}(i_z)) \cdot \text{Var}_{\text{SSC}}(C_{\ell'}^{\text{gal}}(j_z))]^{1/2} \quad (\text{C.6})$$

Again ensuring that the auto correlation matrix of the galaxy spectrum has off-diagonal elements  $\leq 1$  in the SSC dominated regime.

Finally we can apply the Cauchy-Schwarz inequality to the cross-covariance, with

$$f(z) = r^2(z) \frac{dr}{dz} \frac{\partial n_h(i_M, z)}{\partial \delta_b} \mathbb{1}_{[z \in \text{bin}(i_z)]} \quad \text{and} \quad g(z) = r^2(z) \frac{dr}{dz} \frac{\bar{n}_{\text{gal}}(z)^2}{\Delta N_{\text{gal}}(j_z)^2} \frac{\partial P_{\text{gal}}(k_\ell, z)}{\partial \delta_b} \mathbb{1}_{[z \in \text{bin}(j_z)]}$$

yielding

$$|\text{Cov}_{\text{SSC}}(N_{\text{cl}}(i_M, i_z), C_\ell^{\text{gal}}(j_z))| \leq [\text{Var}_{\text{SSC}}(N_{\text{cl}}(i_M, i_z)) \cdot \text{Var}_{\text{SSC}}(C_\ell^{\text{gal}}(j_z))]^{1/2} \quad (\text{C.7})$$

ensuring that the normalised joint correlation matrix has off-diagonal elements  $\leq 1$ .

These inequalities are most useful when  $i_z = j_z$ . Indeed the kernel  $\sigma_{\text{proj}}^2(z_1, z_2)$  decreases quickly when  $z_1$  becomes different from  $z_2$ , thus the SSC covariance decreases quickly for different redshift bins.

**Consequence 2.** The SSC joint covariance is a positive matrix.

**Note.** This statement is in fact more general and implies in particular the Cauchy-Schwarz inequality

*Proofs.* The proof simply amounts to showing that the covariance  $C$  is a semi-inner product.

We have  $C_{ij} = \int dV_{12} \frac{\partial O_i}{\partial \delta_b} \frac{\partial O_j}{\partial \delta_b} \sigma_{\text{proj}}^2(z_1, z_2)$

with observable  $O_i$  being in our case either  $n_h(i_M, z)$  or  $\frac{\bar{n}_{\text{gal}}(z)^2}{\Delta N_{\text{gal}}(j_z)^2} P_{\text{gal}}(k_\ell, z)$ .

Thus:

$${}^T X \cdot C \cdot X = \sum_{ij} \int dV_{12} k^2 dk j_0(kr_1) j_0(kr_2) P(k|z_{12}) x_i \frac{\partial O_i}{\partial \delta_b} \frac{\partial O_j}{\partial \delta_b} x_j \quad (\text{C.8})$$

$$= \int k^2 dk \left[ \int dV j_0(kr) G(z) (P(k, z = 0))^{1/2} \sum_i x_i \frac{\partial O_i}{\partial \delta_b} \right]^2 \quad (\text{C.9})$$

$$\geq 0 \quad (\text{C.10})$$

So that  $C$  defines a semi-inner product.



**Is  $C$  definite?** To investigate definiteness, we have to examine the case of equality in eq. (C.10).

$${}^T X \cdot C \cdot X = 0 \quad \Leftrightarrow \quad \forall k \quad \int dV j_0(kr) G(z) \sum_i x_i \frac{\partial O_i}{\partial \delta_b} = 0 \quad (\text{C.11})$$

$j_0(kr) = \sin(kr)/kr$  so we have basically  $\int dr \sin(kr) f(r) = 0 \quad \forall k$  so  $f(r) = 0$  (except an unphysical dirac at  $r=0$ ). Thus

$${}^T X \cdot C \cdot X = 0 \quad \Leftrightarrow \quad \forall z \quad G(z) \sum_i x_i \frac{\partial O_i}{\partial \delta_b} = 0 \quad (\text{C.12})$$

$$\Leftrightarrow \quad \forall z \quad \frac{\partial \sum_i x_i O_i}{\partial \delta_b} = 0 \quad (\text{C.13})$$

So the covariance is not definite *iff* there is some linear combination of observables which does not react to any change of background at any redshift.

This would happen for instance if we included twice the same observable in the data vector, which we obviously do not do, or if two observables reacted exactly in the same way to a background change, in which case we would see a 100% SSC correlation for them. From the figures given in the text, this is not the case of the cluster counts; however it does appear to be the case for the galaxy power spectrum, to the limit of numerical precision. Hence  $C$  may have some zero eigenvalues, or at least eigenvalues small enough to blow up its condition number. Numerical regularization, e.g. with Principal Component Analysis (PCA), would thus be required to analyse observables deep in the SSC dominated regime.

## D Multipole binning

A binning scheme can be defined by its stakes  $\ell_b$ , such that the bin with index  $b$  contains the multipoles  $\ell_b \leq \ell < \ell_{b+1}$ , with the center of the bin noted as  $\ell_{\text{cent}}$  and its width as  $\Delta\ell$ . A weighting is usually applied to the power spectrum to make the binned elements closer to constant so as to lose less information:

$$D_\ell = w_\ell C_\ell \quad (\text{D.1})$$

$$D_b = \frac{1}{\Delta\ell} \sum_{\ell \in [\ell_b, \ell_{b+1}[} D_\ell \quad (\text{D.2})$$

where the weight  $w_\ell$  is usually  $\frac{\ell(\ell+1)}{2\pi}$  in CMB studies. In our case, the galaxy power spectrum behaves approximately as  $C_\ell \propto 1/\ell$  so that a weight  $w_\ell = \ell$  is more adapted.

The binned  $D_b$  can optionally be transformed back into a spectrum-like quantity with  $C_b = D_b/w_{\ell_{\text{cent}}}$ . The information content in  $D_b$  and  $C_b$  is the same but the latter is useful for visualization purpose when performing actual data analysis.

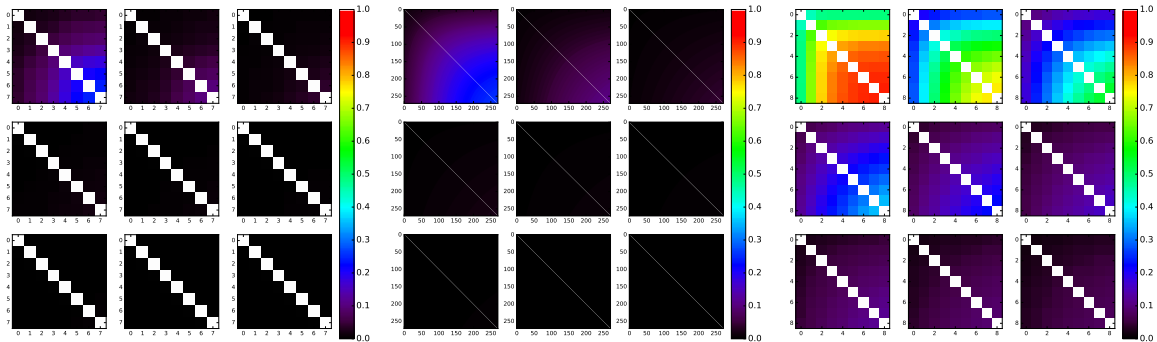
We can thus define the binning operator

$$P_{b,\ell} = \begin{cases} \frac{w_\ell}{w_{\ell_{\text{cent}}} \Delta\ell} & \text{if } \ell \in [\ell_b, \ell_{b+1}[ \\ 0 & \text{otherwise} \end{cases} \quad (\text{D.3})$$

so that

$$C_b = P_{b,\ell} C_\ell \quad (\text{D.4})$$

$$\text{Cov}(C_b, C_{b'}) = P_{b,\ell} P_{b',\ell'} C_{\ell,\ell'} \quad (\text{D.5})$$



**Figure 12.** Correlation matrix of the galaxy power spectrum. *Left:* correlation matrix of eight particular multipoles between  $\ell_{\min} = 30$  and  $\ell_{\max} = 300$ . *Center:* after interpolation, correlation matrix of all multipoles in  $30 \leq \ell \leq 300$ . *Right:* correlation matrix after binning of multipoles with a constant width  $\Delta\ell = 30$ . Each plot contains nine subplots corresponding to nine different redshift bins, going from  $z = 0.1 - 0.2$  in the upper left to  $z = 0.9 - 1$  in the bottom right corner.

To apply these equations in our case, we need predictions of the power spectrum and more difficultly of its covariance for all multipoles in the range considered. Exact analytical prediction however proves too numerically-intensive to be achieved. Nevertheless the power spectrum and its covariance prove remarkably smooth when considered in the adequate space. Thus we employ interpolation. More precisely we use spline interpolation in the space  $(\log \ell, \log C_\ell)$ , which works very accurately for the power spectrum. For the covariance, there are two small problems to overcome: the covariance can be negative, and the Gaussian term is a sharp peak on the diagonal. The sign of the covariance is fortunately constant with multipoles for a given couple of redshift bin, so that we can factor out the sign and interpolate  $\log |C_{\ell,\ell'}|$ . The diagonal of the Gaussian covariance is then interpolated separately from the other terms which have a smooth off-diagonal behaviour.

Figure 12 shows the results of the different steps: first the covariance on the original eight multipoles computed, then the interpolated covariance on the full multipole range, and finally the binned covariance matrix.

One sees that binning increases the off-diagonality of the matrix, and makes thus more evident the importance of non-Gaussianity on the information content of the galaxy power spectrum. Indeed, the Gaussian covariance being diagonal, its contribution decreases with the bin width as  $\text{Var}(C_b) \sim \text{Var}(C_{\ell_{\text{cent}}})/\Delta\ell$ , while non-Gaussian terms decrease much more slowly. For example in the case of super-sample covariance, multipoles are nearly 100% correlated together if SSC is the only covariance contributor, meaning the correlation matrix has value 1 everywhere. After binning, the bins remain obviously nearly 100% correlated together so that the binned SSC correlation matrix is as off-diagonal as the unbinned one.

We note that this result might be seen as counter-intuitive at first sight, since averaging should Gaussianize a variable by the central limit theorem. But the paradox is only at first sight: the central limit theorem ensures that the *likelihood* of  $C_b$  becomes closer to Gaussian as  $\Delta\ell$  increases, but it tells nothing about the evolution of the covariance matrix of  $C_b$ . Physically, the constant effect of mode-coupling (due to non-linearity) becomes more evident as the error bars on mode determination shrink (via binning).

## Acknowledgments

This work was supported by the São Paulo Research Foundation (FAPESP) under grants #2011/11973 and #2013/19936-6. R.R. was partially supported by a CNPq research grant. The authors acknowledge useful discussions with Michel Aguena, Scott Dodelson, Enrique Gaztañaga, Kai Hoffmann, Benjamin Joachimi and Elisabeth Krause.

## References

- [1] PLANCK collaboration, P.A.R. Ade et al., *Planck 2015 results. XIII. Cosmological parameters*, [arXiv:1502.01589](#) [INSPIRE].
- [2] J. Frieman, M. Turner and D. Huterer, *Dark Energy and the Accelerating Universe*, *Ann. Rev. Astron. Astrophys.* **46** (2008) 385 [[arXiv:0803.0982](#)] [INSPIRE].
- [3] W. Hu and A.V. Kravtsov, *Sample variance considerations for cluster surveys*, *Astrophys. J.* **584** (2003) 702 [[astro-ph/0203169](#)] [INSPIRE].
- [4] S. Majumdar and J.J. Mohr, *Self calibration in cluster studies of dark energy: Combining the cluster redshift distribution, the power spectrum and mass measurements*, *Astrophys. J.* **613** (2004) 41 [[astro-ph/0305341](#)] [INSPIRE].
- [5] M. Lima and W. Hu, *Self-calibration of cluster dark energy studies: Observable-mass distribution*, *Phys. Rev. D* **72** (2005) 043006 [[astro-ph/0503363](#)] [INSPIRE].
- [6] L. Liberato and R. Rosenfeld, *Dark energy parameterizations and their effect on dark halos*, *JCAP* **07** (2006) 009 [[astro-ph/0604071](#)] [INSPIRE].
- [7] L.R. Abramo, R.C. Batista, L. Liberato and R. Rosenfeld, *Structure formation in the presence of dark energy perturbations*, *JCAP* **11** (2007) 012 [[arXiv:0707.2882](#)] [INSPIRE].
- [8] F. de Simoni et al., *Large-scale analysis of the SDSS-III DR8 photometric luminous galaxies angular correlation function*, *Mon. Not. Roy. Astron. Soc.* **435** (2013) 3017 [[arXiv:1308.0630](#)] [INSPIRE].
- [9] A. Cooray and R.K. Sheth, *Halo models of large scale structure*, *Phys. Rept.* **372** (2002) 1 [[astro-ph/0206508](#)] [INSPIRE].
- [10] J.L. Tinker et al., *Toward a halo mass function for precision cosmology: The limits of universality*, *Astrophys. J.* **688** (2008) 709 [[arXiv:0803.2706](#)] [INSPIRE].
- [11] J.F. Navarro, C.S. Frenk and S.D.M. White, *A universal density profile from hierarchical clustering*, *Astrophys. J.* **490** (1997) 493 [[astro-ph/9611107](#)] [INSPIRE].
- [12] R. Scoccimarro, R.K. Sheth, L. Hui and B. Jain, *How many galaxies fit in a halo? Constraints on galaxy formation efficiency from spatial clustering*, *Astrophys. J.* **546** (2001) 20 [[astro-ph/0006319](#)] [INSPIRE].
- [13] U. Seljak, *Analytic model for galaxy and dark matter clustering*, *Mon. Not. Roy. Astron. Soc.* **318** (2000) 203 [[astro-ph/0001493](#)] [INSPIRE].
- [14] A. Cooray and W. Hu, *Power spectrum covariance of weak gravitational lensing*, *Astrophys. J.* **554** (2001) 56 [[astro-ph/0012087](#)] [INSPIRE].
- [15] M. Takada and W. Hu, *Power Spectrum Super-Sample Covariance*, *Phys. Rev. D* **87** (2013) 123504 [[arXiv:1302.6994](#)] [INSPIRE].
- [16] Y. Li, W. Hu and M. Takada, *Super-Sample Covariance in Simulations*, *Phys. Rev. D* **89** (2014) 083519 [[arXiv:1401.0385](#)] [INSPIRE].
- [17] M. Takada and S. Bridle, *Probing dark energy with cluster counts and cosmic shear power spectra: including the full covariance*, *New J. Phys.* **9** (2007) 446 [[arXiv:0705.0163](#)] [INSPIRE].
- [18] M. Takada and D.N. Spergel, *Joint analysis of cluster number counts and weak lensing power spectrum to correct for the super-sample covariance*, *Mon. Not. Roy. Astron. Soc.* **441** (2014) 2456 [[arXiv:1307.4399](#)] [INSPIRE].

- [19] E. Schaan, M. Takada and D.N. Spergel, *Joint likelihood function of cluster counts and  $n$ -point correlation functions: Improving their power through including halo sample variance*, *Phys. Rev. D* **90** (2014) 123523 [[arXiv:1406.3330](#)] [[INSPIRE](#)].
- [20] A. Mana, T. Giannantonio, J. Weller, B. Hoyle, G. Huetsi and B. Sartoris, *Combining clustering and abundances of galaxy clusters to test cosmology and primordial non-Gaussianity*, *Mon. Not. Roy. Astron. Soc.* **434** (2013) 684 [[arXiv:1303.0287](#)] [[INSPIRE](#)].
- [21] G. Huetsi and O. Lahav, *The cluster-galaxy cross-spectrum: an additional probe of cosmological and halo parameters*, *Astron. Astrophys.* **492** (2008) 355 [[arXiv:0810.1229](#)] [[INSPIRE](#)].
- [22] E. Krause and T. Eifler, *CosmoLike — Cosmological Likelihood Analyses for Photometric Galaxy Surveys*, [arXiv:1601.05779](#) [[INSPIRE](#)].
- [23] J.S. Bullock et al., *Profiles of dark haloes. Evolution, scatter and environment*, *Mon. Not. Roy. Astron. Soc.* **321** (2001) 559 [[astro-ph/9908159](#)] [[INSPIRE](#)].
- [24] E. Komatsu and U. Seljak, *The Sunyaev-Zel'dovich angular power spectrum as a probe of cosmological parameters*, *Mon. Not. Roy. Astron. Soc.* **336** (2002) 1256 [[astro-ph/0205468](#)] [[INSPIRE](#)].
- [25] J.L. Tinker et al., *The Large Scale Bias of Dark Matter Halos: Numerical Calibration and Model Tests*, *Astrophys. J.* **724** (2010) 878 [[arXiv:1001.3162](#)] [[INSPIRE](#)].
- [26] K. Hoffmann, J. Bel and E. Gaztanaga, *Comparing halo bias from abundance and clustering*, *Mon. Not. Roy. Astron. Soc.* **450** (2015) 1674 [[arXiv:1503.00313](#)] [[INSPIRE](#)].
- [27] J.L. Tinker and A.R. Wetzel, *What Does Clustering Tell Us About the Buildup of the Red Sequence?*, *Astrophys. J.* **719** (2010) 88 [[arXiv:0909.1325](#)] [[INSPIRE](#)].
- [28] A. Pénin, F. Lacasa and N. Aghanim, *Non-Gaussianity of the cosmic infrared background anisotropies — II. Predictions of the bispectrum and constraints forecast*, *Mon. Not. Roy. Astron. Soc.* **439** (2014) 143 [[arXiv:1312.1252](#)] [[INSPIRE](#)].
- [29] F. Lacasa, A. Pénin and N. Aghanim, *Non-Gaussianity of the cosmic infrared background anisotropies — I. Diagrammatic formalism and application to the angular bispectrum*, *Mon. Not. Roy. Astron. Soc.* **439** (2014) 123 [[arXiv:1312.1251](#)] [[INSPIRE](#)].
- [30] D.J. Eisenstein and W. Hu, *Baryonic features in the matter transfer function*, *Astrophys. J.* **496** (1998) 605 [[astro-ph/9709112](#)] [[INSPIRE](#)].
- [31] T. Baldauf, U. Seljak, V. Desjacques and P. McDonald, *Evidence for Quadratic Tidal Tensor Bias from the Halo Bispectrum*, *Phys. Rev. D* **86** (2012) 083540 [[arXiv:1201.4827](#)] [[INSPIRE](#)].
- [32] J.N. Fry, *The Galaxy correlation hierarchy in perturbation theory*, *Astrophys. J.* **279** (1984) 499 [[INSPIRE](#)].
- [33] F. Lacasa, A. Pénin and N. Aghanim, *Non-Gaussianity of the cosmic infrared background anisotropies — I. Diagrammatic formalism and application to the angular bispectrum*, *Mon. Not. Roy. Astron. Soc.* **439** (2014) 123 [[arXiv:1312.1251](#)] [[INSPIRE](#)].
- [34] F. Lacasa, *Non-Gaussianity and extragalactic foregrounds to the Cosmic Microwave Background*, [arXiv:1406.0441](#) [[INSPIRE](#)].
- [35] F. Lacasa, *Combining cosmological constraints from cluster counts and galaxy clustering*, in the proceedings of the *International Astronomical Union, IAU Symposium* **306** (2014) 216.
- [36] J. Barriga and E. Gaztanaga, *The 3-point function in the large scale structure. 1. the weakly nonlinear regime in  $n$ -body simulations*, *Mon. Not. Roy. Astron. Soc.* **333** (2002) 443 [[astro-ph/0112278](#)] [[INSPIRE](#)].
- [37] A.G. Sanchez, C.M. Baugh and R. Angulo, *What is the best way to measure baryonic acoustic oscillations?*, *Mon. Not. Roy. Astron. Soc.* **390** (2008) 1470 [[arXiv:0804.0233](#)] [[INSPIRE](#)].
- [38] H.A. Schwarz, *Über ein Flächen kleinsten Flächeninhalts betreffendes Problem der Variationsrechnung*, *Acta Societatis scientiarum Fennicae* XV, pg. 318, <http://www-stat.wharton.upenn.edu/~steele/Publications/Books/CSMC/Schwarz.pdf>.