

A complete FFT-based decomposition formalism for the redshift-space bispectrum

Naonori S. Sugiyama,^{1★} Shun Saito,² Florian Beutler^{3,4} and Hee-Jong Seo⁵

¹Kavli Institute for the Physics and Mathematics of the Universe (WPI), Todai Institutes for Advanced Study, The University of Tokyo, Chiba 277-8582, Japan

²Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Strasse 1, D-85740 Garching bei München, Germany

³Institute of Cosmology & Gravitation, University of Portsmouth, Portsmouth PO1 3FX, UK

⁴Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA

⁵Department of Physics and Astronomy, Ohio University, Clippinger Labs, Athens, OH 45701, USA

Accepted 2018 November 20. Received 2018 October 05; in original form 2018 March 12

ABSTRACT

To fully extract cosmological information from non-linear galaxy distribution in redshift space, it is essential to include higher-order statistics beyond the two-point correlation function. In this paper, we propose a new decomposition formalism for computing the anisotropic bispectrum in redshift space and for measuring it from galaxy samples. Our formalism uses tri-polar spherical harmonic decomposition with zero total angular momentum to compress the 3D modes distribution in the redshift-space bispectrum. This approach preserves three fundamental properties of the Universe – statistical homogeneity, isotropy, and parity-symmetry – allowing us to efficiently separate the anisotropic signal induced by redshift-space distortions (RSDs) and the Alcock–Paczynski (AP) effect from the isotropic bispectrum. The relevant expansion coefficients in terms of the anisotropic signal are reduced to one multipole index L , and the $L > 0$ modes are induced only by the RSD or AP effects. Our formalism has two advantages: (1) we can make use of Fast Fourier Transforms to measure the bispectrum and (2) it gives a simple expression to correct for the survey geometry, i.e. the survey window function. As a demonstration, we measure the decomposed bispectrum from the Baryon Oscillation Spectroscopic Survey (BOSS) Data Release 12, and, for the first time, present a 14σ detection of the anisotropic bispectrum in the $L = 2$ mode.

Key words: dark matter – large-scale structure of Universe – cosmology: observations – cosmology: theory.

1 INTRODUCTION

In the current picture of structure formation, the standard theory of inflation (Starobinsky 1980; Guth 1981; Sato 1981; Albrecht & Steinhardt 1982; Linde 1982) predicts primordial density perturbations that are nearly Gaussian. If the density fluctuation of galaxies kept being a purely Gaussian random field during its evolution with cosmic time, the two-point function or its Fourier transform, the power spectrum, would completely characterize the spatial distribution of galaxies. However, the non-linear growth of structure produces strong non-Gaussian fluctuations through the mode-coupling of different scales, resulting in higher order statistics (e.g. Peebles 1980; Bernardeau et al. 2002). Consequently, cosmological information on the two-point statistics leaks to these non-Gaussian (non-linear) fluctuations. The three-point function or its Fourier space counterpart, the bispectrum, is thus a powerful

tool for extracting a variety of important cosmological information on a given galaxy distribution, which complements that from the two-point statistics.

With this motivation, there have been various efforts to measure the three-point statistics. The first measurements of the galaxy three-point function and the bispectrum were carried out in angular catalogues by Peebles & Groth (1975), Groth & Peebles (1977), and Fry & Seldner (1982), respectively. Subsequently, many analyses of the three-point statistics have been performed in spectroscopic galaxy surveys by Kayo et al. (2004), Jing & Boerner (2004), Wang et al. (2004), Gaztanaga et al. (2005), Nichol et al. (2006), Kulkarni et al. (2007), Gaztanaga et al. (2009), McBride et al. (2011a,b), Marin (2011), Marin et al. (2013), Guo et al. (2014, 2015), and Slepian et al. (2017a) in configuration space and by Scoccimarro et al. (2001), Feldman et al. (2001), Verde et al. (2002), and Gil-Marín et al. (2015a,b) in Fourier space. Most of these studies were limited to certain choices of triangular configurations of the three-point statistics, because such measurements for all possible configurations have been computationally challenging. Over the

* E-mail: nao.s.sugiyama@gmail.com

past few years, with the progress of algorithm to measure the three-point function (Slepian & Eisenstein 2015, 2016) and the bispectrum (Scoccimarro 2015), one has started to make use of the information of full triangular configurations on the three-point statistics to constrain cosmological parameters in configuration space (Slepian et al. 2017b) and Fourier space (Gil-Marín et al. 2017; Pearson & Samushia 2018).

There remains one important issue in measuring the three-point statistics in the galaxy redshift surveys: all of the previous analyses measure *only one component*, i.e. the monopole, after spherically averaging over the line-of-sight (LOS) direction. The observed clustering pattern of galaxies is anisotropically distorted by the peculiar velocities of galaxies along the LOS, known as redshift-space distortions (RSDs; see Hamilton 1997 for a review). An additional anisotropic signal arises due to the conversion from the observed redshifts into radial distances with incorrect cosmological parameters, which is known as the Alcock–Paczyński (AP) effect (Alcock & Paczyński 1979). These two effects, the RSD and AP effects, each leave their distinctive mark on the observed galaxy clustering.

In the case of the power spectrum that depends only on one wave vector, \mathbf{k} , the anisotropic distortion is quantified by an angle between the wave vector \mathbf{k} and the LOS unit vector \hat{n} and thus can be entirely decomposed using the Legendre polynomials $\mathcal{L}_\ell(\hat{\mathbf{k}} \cdot \hat{\mathbf{n}})$, where \mathcal{L}_ℓ denotes the Legendre polynomial at ℓ -th order. Note that the LOS vector is locally defined with respect to each galaxy pair (local plain-parallel approximation; Beutler et al. 2014). In the bispectrum case, because of its three angular dependences, $\hat{\mathbf{k}}_1$, $\hat{\mathbf{k}}_2$, and $\hat{\mathbf{n}}$, one can arbitrarily choose a coordinate system to characterize its LOS dependence. Scoccimarro, Couchman & Frieman (1999) selected $\hat{\mathbf{k}}_1$ as the z -axis and decomposed the LOS dependence into spherical harmonics: $B(\mathbf{k}_1, \mathbf{k}_2, \hat{\mathbf{n}}) = \sum_{LM} B_{LM}(\mathbf{k}_1, \mathbf{k}_2) Y_L^M(\hat{\mathbf{n}})$. Hashimoto, Rasera & Taruya (2017) defined $(\hat{\mathbf{k}}_1 \times \hat{\mathbf{k}}_2)$ as the z -axis in the same bispectrum decomposition. More recently, Slepian & Eisenstein (2018) proposed in configuration space to take the LOS as the z -axis and to expand the three-point function into the product of two spherical harmonics: $\xi(\mathbf{r}_1, \mathbf{r}_2, \hat{\mathbf{n}}) = \sum_{\ell_1 \ell_2 m} \xi_{\ell_1 \ell_2}^m(\mathbf{r}_1, \mathbf{r}_2, \hat{\mathbf{n}}) Y_{\ell_1}^m(\hat{\mathbf{r}}_1) Y_{\ell_2}^{m*}(\hat{\mathbf{r}}_2)$, where these spherical harmonic functions have the same m mode (spin). We stress here that all of these expansion coefficients include the $M \neq 0$ ($m \neq 0$) modes and hence are variable under rotations. As the Universe is thought to be statistically homogeneous, isotropic, and parity symmetric, any decomposition that preserves these fundamental properties can significantly simplify the analysis.

The main goal of this paper is therefore to propose a more efficient way to distinguish the anisotropic signal on the bispectrum (and the three-point function) induced by the RSD or AP effects from the isotropic signal. To reach the goal, we present a new decomposition of the bispectrum into the a tri-polar spherical harmonic (TripoSH; Varshalovich, Moskalev & Khersonskii 1988) basis $\{\{Y_{\ell_1}(\hat{\mathbf{k}}_1) \otimes Y_{\ell_2}(\hat{\mathbf{k}}_2)\}_{\ell_{12}} \otimes Y_L(\hat{\mathbf{n}})\}_{JM}$, which is a tensor product of spherical harmonics with three different arguments. The three-point function can also be decomposed into a similar basis: $\{\{Y_{\ell_1}(\hat{\mathbf{r}}_1) \otimes Y_{\ell_2}(\hat{\mathbf{r}}_2)\}_{\ell_{12}} \otimes Y_L(\hat{\mathbf{n}})\}_{JM}$. In this basis, the assumption of statistical isotropy allows only the $J = 0$ mode, i.e. zero total angular momentum. Furthermore, the parity symmetry assumption restricts allowed multipoles to $\ell_1 + \ell_2 + L = \text{even}$. This kind of basis was previously utilized to deal with the angular bispectrum of the cosmic microwave background (CMB) radiation (for the latest results, see Ade et al. 2014, 2016 and references therein) and the wide-angle effect in large-scale structure (e.g. Szapudi 2004b). The bispectrum is then characterized by two wavenumbers, k_1 and k_2 ,

and three multipoles, ℓ_1 , ℓ_2 , and L , where we use an uppercase index L for anisotropic distortions along the LOS. We can then single out only the anisotropic signal by computing the $L \neq 0$ modes.

The formalism has two advantages. First, we can make use of Fast Fourier Transform (FFT) to measure our bispectrum multipoles by extending the Scoccimarro (2015) estimator to ours. For the three-point function multipoles, our estimator is closely related to that presented by Slepian & Eisenstein (2016). Second, we can correct for survey geometry effects on the bispectrum/three-point function in a manner similar to the Wilson et al. (2017); Beutler et al. (2017) formalism, which has been developed for the power spectrum analysis.

As a demonstration of the efficiency of our estimators, we measure the bispectrum multipoles from a publicly available galaxy sample, the Baryon Oscillation Spectroscopic Survey (BOSS) Data Release 12 (BOSS DR12; Alam et al. 2015)¹ and investigate the statistical properties of the measured bispectrum multipoles by computing their covariance matrices and signal-to-noise ratios (S/N). The covariance matrix shows an interesting feature that the diagonal elements (i.e. $k_1 = k_2$) of the bispectrum multipoles estimated at different scales, i.e. at $k_1 (= k_2)$ and $k'_1 (= k'_2)$ for $k_1 \neq k'_1$, are nearly uncorrelated like in the case of the power spectrum. We further find that for each of the $L = 0$ and 2 modes, the lowest order of the bispectrum multipoles yields the highest S/N in the L mode. These findings can significantly simplify the bispectrum analysis for such galaxy samples. While we leave a detailed analysis of the redshift-space bispectrum for future work, we will, for the first time, present a 14σ detection of the lowest order $L = 2$ mode and a 6σ detection of the next leading order $L = 2$ mode, which are induced only by anisotropic signals on the bispectrum (i.e. the RSD or AP effects).

The outline of the paper is as follows. In Section 2 we present the new decomposition formalism of the bispectrum and three-point function to characterize the RSD effect on them. In Section 3 we provide a comparison with previous works. In Section 4 we describe the estimators to measure our bispectrum and three-point function multipoles. In Section 5 we demonstrate our formalism by measuring the bispectrum multipoles of the BOSS DR12 galaxy sample and compute their covariance matrices and S/Ns. In Section 6 we show how to correct for the survey geometry effect when theoretically modeling the observed bispectrum and three-point function multipoles. We present a summary and conclusions in Section 7. We provide six appendices for further clarification and systematics test: in Appendix A we summarize the identities used for the derivations of the equations used in this paper; in Appendix B we investigate the dependence of how to choose the LOS direction on the bispectrum measurements; in Appendix C we provide a comparison of several interpolation schemes of density fields; in Appendix D we show how the FKP weighting works in our decomposition formalism; in Appendix E we describe the Hartlap factor and the Percival factor; and in Appendix F we detail the bispectrum multipoles including survey window corrections.

2 NEW DECOMPOSITION FORMALISM

The three-point function and its Fourier transform, the bispectrum, are potential tools for extracting cosmological information that leaks to non-Gaussian fluctuations of galaxy clustering. To put these in context, we compute the number density of galaxies, n_g , in fractional

¹<http://www.sdss.org/dr12/>

units relative to the background density, $\bar{n}_g: n_g = \bar{n}_g (1 + \delta)$, where δ is the density contrast. The three-point function is then defined as $\langle \delta(\mathbf{x}_1)\delta(\mathbf{x}_2)\delta(\mathbf{x}_3) \rangle$, which is the ensemble average of the product of the three density contrasts at points $\mathbf{x}_1, \mathbf{x}_2$, and \mathbf{x}_3 .

The Universe is thought to be statistically homogeneous, and hence, the three-point function can be characterized by two relative coordinates, $\mathbf{r}_1 = \mathbf{x}_1 - \mathbf{x}_3$ and $\mathbf{r}_2 = \mathbf{x}_2 - \mathbf{x}_3$:

$$\zeta(\mathbf{r}_1, \mathbf{r}_2) = \langle \delta(\mathbf{x}_1)\delta(\mathbf{x}_2)\delta(\mathbf{x}_3) \rangle. \quad (1)$$

Using the Fourier transform of the density contrast, $\delta(\mathbf{k}) = \int d^3x e^{-i\mathbf{k}\cdot\mathbf{x}} \delta(\mathbf{x})$, we define the bispectrum as

$$\langle \delta(\mathbf{k}_1)\delta(\mathbf{k}_2)\delta(\mathbf{k}_3) \rangle = (2\pi)^3 \delta_D(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) B(\mathbf{k}_1, \mathbf{k}_2), \quad (2)$$

where δ_D represents the Dirac delta function. In the bispectrum case, its statistical homogeneity corresponds to the triangle condition, $\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3 = 0$.

The observed position of galaxies \mathbf{x} is displaced from the real-space galaxy position \mathbf{x}_r by the physical peculiar velocity of galaxies \mathbf{v} along the LOS direction:

$$\mathbf{x} = \mathbf{x}_r + \frac{\mathbf{v}(\mathbf{x}_r) \cdot \hat{\mathbf{x}}}{aH(a)} \hat{\mathbf{x}}, \quad (3)$$

where a is the scale factor, $H(a)$ is the Hubble expansion parameter, and $\hat{\mathbf{x}} = \mathbf{x}/|\mathbf{x}|$ is a unit vector pointing to the galaxy from the origin. The observed galaxy density is then distorted along the LOS direction, the so-called RSD effect. It is commonly assumed that the anisotropic distortion of the galaxy clustering is characterized by only one global LOS direction $\hat{\mathbf{n}}$, the so-called *global* plane-parallel approximation: $\zeta(\mathbf{r}_1, \mathbf{r}_2, \hat{\mathbf{n}})$ and $B(\mathbf{k}_1, \mathbf{k}_2, \hat{\mathbf{n}})$. However, this assumption does not hold for actual galaxy data, because observed galaxy positions have their own LOS directions, and the three-point function should depend on three LOS directions, $\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2$, and $\hat{\mathbf{x}}_3$, in the most general case. Throughout this paper, we apply the *local* plane-parallel approximation, $\hat{\mathbf{x}}_1 \approx \hat{\mathbf{x}}_2 \approx \hat{\mathbf{x}}_3$, and choose $\hat{\mathbf{x}}_3$ as the LOS direction for the triangular configuration formed by $\mathbf{x}_1, \mathbf{x}_2$, and \mathbf{x}_3 when measuring the bispectrum from a galaxy sample in Section 4. To validate this choice, we show in Appendix B that the difference among the bispectrum measurements choosing each of $\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2$, and $\hat{\mathbf{x}}_3$ as the LOS is negligibly small in the BOSS analysis. We will discuss the modeling of the bispectrum under the *local* plane-parallel approximation by taking into account survey geometry effects in Section 6.

The three angular dependences in the bispectrum, $\hat{\mathbf{k}}_1, \hat{\mathbf{k}}_2$, and $\hat{\mathbf{n}}$, can be generally decomposed into spherical harmonics Y_ℓ^m :

$$B(\mathbf{k}_1, \mathbf{k}_2, \hat{\mathbf{n}}) = \sum_{\ell_1 \ell_2 L} \sum_{m_1 m_2 M} B_{\ell_1 \ell_2 L}^{m_1 m_2 M}(k_1, k_2) \times y_{\ell_1}^{m_1}(\hat{\mathbf{k}}_1) y_{\ell_2}^{m_2}(\hat{\mathbf{k}}_2) y_L^M(\hat{\mathbf{n}}), \quad (4)$$

where $y_\ell^m = \sqrt{4\pi/(2\ell+1)} Y_\ell^m$ is a normalized spherical harmonic function, and the corresponding expansion coefficients are given by

$$B_{\ell_1 \ell_2 L}^{m_1 m_2 M}(k_1, k_2) = N_{\ell_1 \ell_2 L} \int \frac{d^2 \hat{\mathbf{k}}_1}{4\pi} \int \frac{d^2 \hat{\mathbf{k}}_2}{4\pi} \int \frac{d^2 \hat{\mathbf{n}}}{4\pi} \times y_{\ell_1}^{m_1*}(\hat{\mathbf{k}}_1) y_{\ell_2}^{m_2*}(\hat{\mathbf{k}}_2) y_L^{M*}(\hat{\mathbf{n}}) B(\mathbf{k}_1, \mathbf{k}_2, \hat{\mathbf{n}}), \quad (5)$$

with $N_{\ell_1 \ell_2 L} = (2\ell_1 + 1)(2\ell_2 + 1)(2L + 1)$. Throughout this paper, we use uppercase indices L, M for the expansion with respect to the angles relevant to LOS. Namely, the modes with $L > 0$ are induced by the RSD or AP effects.

Alternatively we can decompose the bispectrum using TripoSH expansion (Varshalovich et al. 1988):

$$B(\mathbf{k}_1, \mathbf{k}_2, \hat{\mathbf{n}}) = \sum_{J M_J} \sum_{\ell_1 \ell_2 L \ell_{12}} B_{\ell_1 \ell_2 L \ell_{12}}^{J M_J}(k_1, k_2) S_{\ell_1 \ell_2 L \ell_{12}}^{J M_J}(\hat{\mathbf{k}}_1, \hat{\mathbf{k}}_2, \hat{\mathbf{n}}), \quad (6)$$

where we defined a normalized TripoSH basis as

$$S_{\ell_1 \ell_2 L \ell_{12}}^{J M_J}(\hat{\mathbf{k}}_1, \hat{\mathbf{k}}_2, \hat{\mathbf{n}}) = \sum_{m_1 m_2 m_{12} M} C_{\ell_1 m_1; \ell_2 m_2}^{\ell_{12} m_{12}} C_{\ell_{12} m_{12}; L M}^{J M_J} \times y_{\ell_1}^{m_1}(\hat{\mathbf{k}}_1) y_{\ell_2}^{m_2}(\hat{\mathbf{k}}_2) y_L^M(\hat{\mathbf{n}}), \quad (7)$$

and the TripoSH coefficients are then given by

$$B_{\ell_1 \ell_2 L \ell_{12}}^{J M_J}(k_1, k_2) = \sum_{m_1 m_2 m_{12} M} C_{\ell_1 m_1; \ell_2 m_2}^{\ell_{12} m_{12}} C_{\ell_{12} m_{12}; L M}^{J M_J} \times B_{\ell_1 \ell_2 L}^{m_1 m_2 M}(k_1, k_2), \quad (8)$$

with the Clebsch–Gordan coefficients

$$C_{\ell_1 m_1; \ell_2 m_2}^{\ell_3 m_3} = (-1)^{\ell_1 - \ell_2 + m_3} \sqrt{2\ell_3 + 1} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ m_1 & m_2 & -m_3 \end{pmatrix}. \quad (9)$$

The remarkable feature of the TripoSH formalism is to parameterize departures from statistical isotropy regarding total angular momenta, J and M_J . If the assumption of statistical isotropy breaks, the corresponding expansion coefficients yield the $J \geq 1$ modes, otherwise only the $J = 0$ is non-zero. A similar decomposition formalism in two-point statistics, the bipolar spherical harmonic expansion, has been used to search for the breaking of the statistical isotropy assumption of the Universe in the CMB (e.g. see Planck Collaboration XVI 2016a; Planck Collaboration XX 2016b and references therein) and large-scale structure (Pullen & Hirata 2010; Shiraishi, Sugiyama & Okumura 2017; Sugiyama, Shiraishi & Okumura 2018).

In this work, we assume that the three-point function has arisen through a physical process that is statistically isotropic and parity-symmetric as well as homogeneous. The statistical isotropy induces only the $J = 0$ mode, i.e. zero total angular momentum, in the TripoSH formalism, resulting in

$$B_{\ell_1 \ell_2 L \ell_{12}}^{J=0, M_J=0}(k_1, k_2) = \delta_{\ell_{12} L}^{(K)} (-1)^{\ell_1 - \ell_2 + L} \times \sum_{m_1 m_2 M} \begin{pmatrix} \ell_1 & \ell_2 & L \\ m_1 & m_2 & M \end{pmatrix} B_{\ell_1 \ell_2 L}^{m_1 m_2 M}(k_1, k_2), \quad (10)$$

where the statistical isotropy is satisfied through the summation weighted by the Wigner 3- j symbol over all possible m_1, m_2 , and M with $m_1 + m_2 + M = 0$, and $\delta_{\ell_{12} L}^{(K)}$ is the Kronecker delta defined such that $\delta_{\ell_{12} L}^{(K)} = 1$ if $\ell_{12} = L$, otherwise zero. Since the parity symmetry restricts allowed multipoles to $\ell_1 + \ell_2 + L = \text{even}$, it is useful to define the following bispectrum multipoles to simplify the final expressions:

$$B_{\ell_1 \ell_2 L}(k_1, k_2) = H_{\ell_1 \ell_2 L} \sum_{m_1 m_2 M} \begin{pmatrix} \ell_1 & \ell_2 & L \\ m_1 & m_2 & M \end{pmatrix} B_{\ell_1 \ell_2 L}^{m_1 m_2 M}(k_1, k_2), \quad (11)$$

where $H_{\ell_1 \ell_2 L} = \begin{pmatrix} \ell_1 & \ell_2 & L \\ 0 & 0 & 0 \end{pmatrix}$ filters even $\ell_1 + \ell_2 + L$ components. The bispectrum multipoles defined in the above correspond to the expansion coefficients of the following bispectrum

decomposition:

$$B(\mathbf{k}_1, \mathbf{k}_2, \hat{n}) = \sum_{\ell_1 + \ell_2 + L = \text{even}} B_{\ell_1 \ell_2 L}(k_1, k_2) S_{\ell_1 \ell_2 L}(\hat{k}_1, \hat{k}_2, \hat{n}), \quad (12)$$

with the basis function

$$S_{\ell_1 \ell_2 L}(\hat{k}_1, \hat{k}_2, \hat{n}) = \frac{1}{H_{\ell_1 \ell_2 L}} \sum_{m_1 m_2 M} \begin{pmatrix} \ell_1 & \ell_2 & L \\ m_1 & m_2 & M \end{pmatrix} \times y_{\ell_1}^{m_1}(\hat{k}_1) y_{\ell_2}^{m_2}(\hat{k}_2) y_L^M(\hat{n}). \quad (13)$$

The bispectrum multipoles contain all physical information under the three statistical assumptions: homogeneity, isotropy, and parity-symmetry of the Universe.

In the same manner, as the bispectrum, we can expand the three-point function in spherical harmonics,

$$\zeta(\mathbf{r}_1, \mathbf{r}_2, \hat{n}) = \sum_{\ell_1 \ell_2 L} \sum_{m_1 m_2 M} \zeta_{\ell_1 \ell_2 L}^{m_1 m_2 M}(r_1, r_2) \times y_{\ell_1}^{m_1}(\hat{r}_1) y_{\ell_2}^{m_2}(\hat{r}_2) y_L^M(\hat{n}). \quad (14)$$

Under the assumptions of statistical isotropy and parity-symmetry, we have

$$\zeta(\mathbf{r}_1, \mathbf{r}_2, \hat{n}) = \sum_{\ell_1 + \ell_2 + L = \text{even}} \zeta_{\ell_1 \ell_2 L}(r_1, r_2) S_{\ell_1 \ell_2 L}(\hat{r}_1, \hat{r}_2, \hat{n}), \quad (15)$$

and the three-point function multipoles are given by

$$\zeta_{\ell_1 \ell_2 L}(r_1, r_2) = H_{\ell_1 \ell_2 L} \sum_{m_1 m_2 M} \begin{pmatrix} \ell_1 & \ell_2 & L \\ m_1 & m_2 & M \end{pmatrix} \zeta_{\ell_1 \ell_2 L}^{m_1 m_2 M}(r_1, r_2). \quad (16)$$

The bispectrum and three-point function multipoles are related to each other according to Hankel transformations:

$$B_{\ell_1 \ell_2 L}(k_1, k_2) = (-i)^{\ell_1 + \ell_2} (4\pi)^2 \int dr_1 r_1^2 \int dr_2 r_2^2 \times j_{\ell_1}(k_1 r_1) j_{\ell_2}(k_2 r_2) \zeta_{\ell_1 \ell_2 L}(r_1, r_2) \\ \zeta_{\ell_1 \ell_2 L}(r_1, r_2) = i^{\ell_1 + \ell_2} \int \frac{dk_1 k_1^2}{2\pi^2} \int \frac{dk_2 k_2^2}{2\pi^2} \times j_{\ell_1}(r_1 k_1) j_{\ell_2}(r_2 k_2) B_{\ell_1 \ell_2 L}(k_1, k_2), \quad (17)$$

where j_ℓ are the spherical Bessel functions of order ℓ . While the three-point function multipoles satisfy the reality condition: $\zeta_{\ell_1 \ell_2 L}^* = \zeta_{\ell_1 \ell_2 L}$, the bispectrum multipoles $B_{\ell_1 \ell_2 L}$ may be imaginary if $\ell_1 + \ell_2 = \text{odd}$.

In the context of RSDs, we can show in perturbation theories that the main contribution to the L mode of the bispectrum comes from the terms proportional to $(\hat{n} \cdot \mathbf{v})^{L/2}$ (Sugiyama et al., in preparation). Therefore, we expect that the L mode has similar information on the velocity field to the ℓ mode of the power spectrum multipoles, expanded in Legendre polynomials. To validate this expectation, we will show in Section 5.5 that there is a strong correlation between the bispectrum multipoles and the power spectrum multipoles if $L = \ell$, by measuring the cross-covariance matrix between them from mock catalogues described in Section 5.1.

We end this section by discussing allowed combinations of the multipoles, ℓ_1 , ℓ_2 , and L . First, since $B(\mathbf{k}_1, \mathbf{k}_2, \hat{n}) = B(\mathbf{k}_2, \mathbf{k}_1, \hat{n})$, we obtain $B_{\ell_1 \ell_2 L}(k_1, k_2) = B_{\ell_2 \ell_1 L}(k_2, k_1)$. Therefore, we can focus only on $\ell_1 \geq \ell_2$ without loss of generality. Second, anisotropies due to RSDs are axially symmetric around the LOS direction in the plane-parallel approximation. The axial symmetry then restricts us to $L = \text{even}$. Third, the Wigner 3- j symbol $\begin{pmatrix} \ell_1 & \ell_2 & L \\ m_1 & m_2 & M \end{pmatrix}$ in equation (11), which appears as a result of the rotational symmetry

Table 1. Allowed combinations of the multipoles (ℓ_1, ℓ_2, L) in $\zeta_{\ell_1 \ell_2 L}$ and $B_{\ell_1 \ell_2 L}$, which are determined by the four conditions: (1) $\ell_1 \geq \ell_2$; (2) $L = \text{even}$; (3) $|\ell_1 - \ell_2| \leq L \leq \ell_1 + \ell_2$; and (4) $\ell_1 + \ell_2 + L = \text{even}$. As a demonstration, we show here the first four combinations for each of $L = 0, 2$, and 4.

L	(ℓ_1, ℓ_2)				
$L = 0$	(0,0)	(1,1)	(2,2)	(3,3)	...
$L = 2$	(2,0)	(1,1)	(3,1)	(2,2)	...
$L = 4$	(4,0)	(3,1)	(2,2)	(5,1)	...

assumption, satisfies the following selection rule: $|\ell_1 - \ell_2| \leq L \leq \ell_1 + \ell_2$. Finally, with the parity symmetry condition $\ell_1 + \ell_2 + L = \text{even}$, we find $\ell_1 + \ell_2 = \text{even}$ because of $L = \text{even}$, resulting in that the bispectrum multipoles should be real: $B_{\ell_1 \ell_2 L} = B_{\ell_1 \ell_2 L}^*$. Throughout this paper, we focus especially on the $L = 0, 2$, and 4 modes to clarify the relation to the power spectrum monopole, quadrupole, and hexadecapole. As a demonstration, we write down the first four combinations of (ℓ_1, ℓ_2, L) for these modes in Table 1. Specifically, in what follows we analyze six bispectrum multipoles $B_{\ell_1 \ell_2 L}$ for $(\ell_1, \ell_2, L) = (0, 0, 0), (1, 1, 0), (2, 2, 0), (2, 0, 2), (1, 1, 2)$, and $(4, 0, 4)$.

3 COMPARISON WITH PREVIOUS WORKS

In the previous section, we have presented a new decomposition formalism that is fairly general and independent of the choice of the coordinate system. In this section, we will show how our formalism is related to other coordinate choices used in the previous work.

3.1 The $L = 0$ mode, i.e. monopole

In the absence of both the RSD and AP effects, the bispectrum (the three-point function) is a function of \mathbf{k}_1 and \mathbf{k}_2 (r_1 and r_2), and hence, it is common to expand it in Legendre polynomials \mathcal{L}_ℓ (Szapudi 2004a; Pan & Szapudi 2005; Slepian & Eisenstein 2015, 2016):

$$B(\mathbf{k}_1, \mathbf{k}_2) = \sum_{\ell} B_{\ell}(k_1, k_2) \mathcal{L}_{\ell}(\hat{k}_1 \cdot \hat{k}_2) \\ \zeta(\mathbf{r}_1, \mathbf{r}_2) = \sum_{\ell} \zeta_{\ell}(r_1, r_2) \mathcal{L}_{\ell}(\hat{r}_1 \cdot \hat{r}_2). \quad (18)$$

This decomposition is possible even in redshift space after averaging over the LOS: $B(\mathbf{k}_1, \mathbf{k}_2) = \int \frac{d^2 \hat{n}}{4\pi} B(\mathbf{k}_1, \mathbf{k}_2, \hat{n})$. The corresponding expansion coefficients B_{ℓ} (hereafter, the Legendre coefficients) are then given by

$$B_{\ell}(k_1, k_2) = (2\ell + 1) \int \frac{d^2 \hat{k}_1}{4\pi} \int \frac{d^2 \hat{k}_2}{4\pi} \int \frac{d^2 \hat{n}}{4\pi} \times \mathcal{L}_{\ell}(\hat{k}_1 \cdot \hat{k}_2) B(\mathbf{k}_1, \mathbf{k}_2, \hat{n}), \quad (19)$$

and $\zeta_{\ell}(r_1, r_2)$ can be computed using a Hankel transform (equation 3 in Szapudi 2004a):

$$\zeta_{\ell}(r_1, r_2) = (-1)^{\ell} \int \frac{dk_1 k_1^2}{2\pi^2} \int \frac{dk_2 k_2^2}{2\pi^2} \times j_{\ell}(r_1 k_1) j_{\ell}(r_2 k_2) B_{\ell}(k_1, k_2). \quad (20)$$

The $L = 0$ mode of $\zeta_{\ell_1 \ell_2 L}$ and $B_{\ell_1 \ell_2 L}$ reproduces the Legendre coefficients ζ_{ℓ} and B_{ℓ} . Using the relation between spherical harmonics and Legendre polynomials (equation A2), equation (11) for

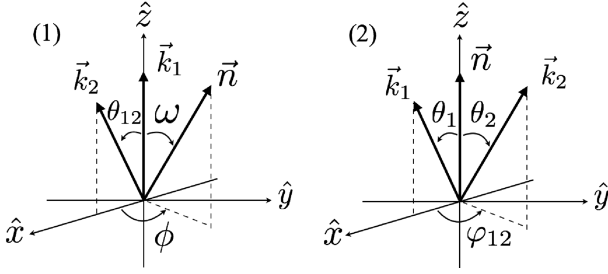


Figure 1. The configuration of the variables used to compute the bispectrum multipoles. The left- and right-hand coordinate systems choose \hat{k}_1 and \hat{n} as the z -axis, respectively.

$L = 0$ leads to

$$\begin{aligned} \zeta_{\ell_1 \ell_2 L=0}(r_1, r_2) &= \delta_{\ell_1 \ell_2}^{(K)} \zeta_{\ell_1}(r_1, r_2) \\ B_{\ell_1 \ell_2 L=0}(k_1, k_2) &= \delta_{\ell_1 \ell_2}^{(K)} B_{\ell_1}(k_1, k_2). \end{aligned} \quad (21)$$

Equation (17) for $L = 0$ reduces to equation (20). Thus, our formalism can be regarded as a generalized one of the Legendre coefficients, and the $L \geq 2$ modes will provide additional information in terms of anisotropic signals induced by the RSD or AP effects.

3.2 Coordinate systems

While the bispectrum and three-point function multipoles shown in Section 2 are independent of the choice of coordinate systems, it is convenient to choose specific coordinate axes when theoretically modeling them. In this section, we present two choices of coordinates: (1) \hat{k}_1 is chosen as the \hat{z} -axis, which is frequently used for the bispectrum in the literature (e.g. Scoccimarro et al. 1999); and (2) the LOS direction \hat{n} is taken as the \hat{z} -axis, which is recently proposed by Slepian & Eisenstein (2018) to characterize RSDs in the three-point function. The configuration of the variables in these coordinate systems is shown in Fig. 1. We present how to calculate the bispectrum multipoles (equation 11) in these specific coordinates; the three-point function multipoles can then be computed through the Hankel transform in equation (17). In Section 6, we adopt Scoccimarro's coordinate system to compute the theoretical bispectrum multipoles using perturbation theory.

3.2.1 Case 1: $\hat{k}_1 = \hat{z}$

Following Scoccimarro et al. (1999), we choose \hat{k}_1 as the z -axis without loss of generality and adopt the following coordinate system:

$$\begin{aligned} \mathbf{k}_1 &= \{0, 0, k_1\} \\ \mathbf{k}_2 &= \{k_2 \sin \theta_{12}, 0, k_2 \cos \theta_{12}\} \\ \hat{n} &= \{\sin \omega \cos \phi, \sin \omega \sin \phi, \cos \omega\}. \end{aligned} \quad (22)$$

The bispectrum in redshift space is then characterized by the five parameters: $B = B(k_1, k_2, \theta_{12}, \omega, \phi)$. The three parameters k_1 , k_2 , and θ_{12} define the shape of the triangle; the remaining two angles ω and ϕ characterize the orientation of the triangle with respect to the LOS direction. The spherical harmonic functions with \hat{k}_1 , \hat{k}_2 , and \hat{n} then become $y_{\ell_1}^{m_1*}(\hat{z}) = \delta_{0m_1}^{(K)}$, $y_{\ell_2}^{m_2*}(\hat{k}_2) = y_{\ell_2}^{m_2*}(\theta_{12}, 0)$, and $y_L^{M*}(\hat{n}) = y_L^{M*}(\omega, \phi)$, respectively. Substituting these into equa-

tions (5) and (11) leads to

$$\begin{aligned} B_{\ell_1 \ell_2 L}(k_1, k_2) &= N_{\ell_1 \ell_2 L} H_{\ell_1 \ell_2 L} \int \frac{d \cos \omega d \phi}{4\pi} \int \frac{d \cos \theta_{12}}{2} \\ &\times \left[\sum_M \begin{pmatrix} \ell_1 & \ell_2 & L \\ 0 & -M & M \end{pmatrix} y_{\ell_2}^{-M*}(\theta_{12}, 0) y_L^{M*}(\omega, \phi) \right] \\ &\times B(k_1, k_2, \theta_{12}, \omega, \phi). \end{aligned} \quad (23)$$

To more clearly show the relation to the previous work, we decompose $B(k_1, k_2, \theta_{12}, \omega, \phi)$ into spherical harmonics (equation 19 in Scoccimarro et al. 1999):

$$B(k_1, k_2, \theta_{12}, \omega, \phi) = \sum_{LM} B_{LM}(k_1, k_2, \theta_{12}) Y_L^M(\omega, \phi). \quad (24)$$

Then, the above bispectrum multipoles B_{LM} are related to ours $B_{\ell_1 \ell_2 L}$ as follows:

$$\begin{aligned} B_{\ell_1 \ell_2 L}(k_1, k_2) &= \frac{N_{\ell_1 \ell_2 L} H_{\ell_1 \ell_2 L}}{\sqrt{(4\pi)(2L+1)}} \int \frac{d \cos \theta_{12}}{2} \\ &\times \left[\sum_M \begin{pmatrix} \ell_1 & \ell_2 & L \\ 0 & -M & M \end{pmatrix} y_{\ell_2}^{-M*}(\cos \theta_{12}, 0) \right] \\ &\times B_{LM}(k_1, k_2, \theta_{12}). \end{aligned} \quad (25)$$

3.2.2 Case 2: $\hat{n} = \hat{z}$

We choose \hat{n} as the z -axis and adopt the following coordinate system:

$$\begin{aligned} \mathbf{k}_1 &= \{k_1 \sin \theta_1, 0, k_1 \cos \theta_1\} \\ \mathbf{k}_2 &= \{k_2 \sin \theta_2 \cos \varphi_{12}, k_2 \sin \theta_2 \sin \varphi_{12}, k_2 \cos \theta_2\} \\ \hat{n} &= \{0, 0, 1\}. \end{aligned} \quad (26)$$

In the same manner as Section 3.2.1, we obtain

$$\begin{aligned} B_{\ell_1 \ell_2 L}(k_1, k_2) &= N_{\ell_1 \ell_2 L} H_{\ell_1 \ell_2 L} \int \frac{d \cos \theta_1}{2} \int \frac{d \cos \theta_2 d \varphi_{12}}{4\pi} \\ &\times \left[\sum_m \begin{pmatrix} \ell_1 & \ell_2 & L \\ m & -m & 0 \end{pmatrix} y_{\ell_1}^{m*}(\theta_1, 0) y_{\ell_2}^{-m*}(\theta_2, \varphi_{12}) \right] \\ &\times B(k_1, k_2, \theta_1, \theta_2, \varphi_{12}). \end{aligned} \quad (27)$$

Slepian & Eisenstein (2018) proposed to decompose the three-point function into a product of two spherical harmonics: $\zeta(\mathbf{r}_1, \mathbf{r}_2, \hat{n} = \hat{z}) = \sum_{\ell_1 \ell_2 m} \zeta_{\ell_1 \ell_2}^m(r_1, r_2) y_{\ell_1}^m(\hat{r}_1) y_{\ell_2}^{m*}(\hat{r}_2)$. The Fourier transform of the expansion coefficients $\zeta_{\ell_1 \ell_2}^m$ is given by

$$\begin{aligned} B_{\ell_1 \ell_2}^m(k_1, k_2) &= (2\ell_1 + 1)(2\ell_2 + 1) \int \frac{d \cos \theta_1}{2} \int \frac{d \cos \theta_2 d \varphi_{12}}{4\pi} \\ &\times y_{\ell_1}^{m*}(\theta_1, 0) y_{\ell_2}^m(\theta_2, \varphi_{12}) \\ &\times B(k_1, k_2, \theta_1, \theta_2, \varphi_{12}). \end{aligned} \quad (28)$$

Equation (27) can be then represented as

$$\begin{aligned} B_{\ell_1 \ell_2 L}(k_1, k_2) &= (2L + 1) H_{\ell_1 \ell_2 L} \sum_m \begin{pmatrix} \ell_1 & \ell_2 & L \\ m & -m & 0 \end{pmatrix} (-1)^m B_{\ell_1 \ell_2}^m(k_1, k_2). \end{aligned} \quad (29)$$

The three-point function multipoles presented in this paper are related to $\zeta_{\ell_1 \ell_2}^m$ as follows:

$$\begin{aligned} \zeta_{\ell_1 \ell_2 L}(r_1, r_2) &= (2L + 1) H_{\ell_1 \ell_2 L} \sum_m \begin{pmatrix} \ell_1 & \ell_2 & L \\ m & -m & 0 \end{pmatrix} (-1)^m \zeta_{\ell_1 \ell_2}^m(r_1, r_2). \end{aligned} \quad (30)$$

4 ESTIMATORS

In this section, we present the estimators to measure the bispectrum and three-point function multipoles, $B_{\ell_1 \ell_2 L}$ and $\zeta_{\ell_1 \ell_2 L}$, using the FFT algorithm. The algorithm to measure the three-point statistics using FFT was first proposed by Scoccimarro (2015) for the bispectrum and by Slepian & Eisenstein (2016) for the three-point function. We extend their estimators to our decomposition formalism. Section 4.1 describes how to measure the number density of galaxies weighted by spherical harmonics with the LOS direction $y_L^{M*}(\hat{x})$. Using the y_L^{M*} -weighted density field, we build the estimators of $B_{\ell_1 \ell_2 L}$ and $\zeta_{\ell_1 \ell_2 L}$ in Sections 4.2 and 4.3. While these estimators can be applied to an observed galaxy catalogue, in Section 4.4 we show that our estimator can be straightforwardly applied to N -body simulations in which the global-plane parallel approximation is often adopted.

Measuring the bispectrum and three-point function from galaxy data, we use the Cartesian coordinates such that $\mathbf{x} = \chi(z)\{\cos \delta_g \cos \alpha_g, \cos \delta_g \sin \alpha_g, \sin \delta_g\}$, where χ is the comoving radial distance to galaxies, α_g and δ_g are, respectively, the right ascension and declination angles of galaxies. To be consistent with such data coordinate system, we choose the north pole as our z -axis for our decomposition. Again, the resulting bispectrum multipoles in our formalism do not depend on this axis choice.

4.1 Number density of galaxies

The number density of galaxies is given by

$$n(\mathbf{x}) = \sum_i^{N_{\text{gal}}} w(\mathbf{x}_i) \delta_{\text{D}}(\mathbf{x} - \mathbf{x}_i), \quad (31)$$

where \mathbf{x}_i represents the observed position of galaxy i , N_{gal} denotes the total number of observed galaxies, and the weight function $w(\mathbf{x})$ may include systematic weights (e.g. Reid et al. 2016) and the FKP weight (Feldman, Kaiser & Peacock 1994) (for details, see Section 5.2). We estimate the mean number density $\bar{n}(\mathbf{x})$ from a synthetic random catalogue,

$$\bar{n}(\mathbf{x}) = \alpha \sum_i^{N_{\text{ran}}} w(\mathbf{x}_i) \delta_{\text{D}}(\mathbf{x} - \mathbf{x}_i), \quad (32)$$

where N_{ran} is the total number of objects in the random catalogue, and $\alpha = \sum_i^{N_{\text{gal}}} w_i / \sum_i^{N_{\text{ran}}} w_i$. By subtracting $\bar{n}(\mathbf{x})$ from $n(\mathbf{x})$, we obtain the density fluctuation:

$$\delta n(\mathbf{x}) = n(\mathbf{x}) - \bar{n}(\mathbf{x}). \quad (33)$$

We then define the y_L^{M*} -weighted density fluctuation as (Hand et al. 2017; Sugiyama et al. 2018)

$$\begin{aligned} \delta n_L^M(\mathbf{x}) &\equiv y_L^{M*}(\hat{x}) \delta n(\mathbf{x}) \\ &= \left(\sum_i^{N_{\text{gal}}} -\alpha \sum_i^{N_{\text{ran}}} \right) y_L^{M*}(\hat{x}_i) w(\mathbf{x}_i) \delta_{\text{D}}(\mathbf{x} - \mathbf{x}_i). \end{aligned} \quad (34)$$

We note here that the leading order of δn_L^M is the same as the normal density fluctuation: $\delta n(\mathbf{x}) = \delta n_0^0(\mathbf{x})$.

4.2 Bispectrum

Since the LOS direction is not globally but locally determined by galaxy positions, we apply the local plane-parallel approximation $\hat{x}_1 \approx \hat{x}_2 \approx \hat{x}_3$ and choose \hat{x}_3 as the LOS direction in the main text (see Section 2 and Appendix B). Under this approximation, we

present the estimator of the bispectrum multipoles as follows:

$$\begin{aligned} \widehat{B}_{\ell_1 \ell_2 L}(k_1, k_2) &= H_{\ell_1 \ell_2 L} \sum_{m_1 m_2 M} \begin{pmatrix} \ell_1 & \ell_2 & L \\ m_1 & m_2 & M \end{pmatrix} \\ &\times \frac{N_{\ell_1 \ell_2 L}}{I} \int \frac{d^2 \hat{k}_1}{4\pi} y_{\ell_1}^{m_1*}(\hat{k}_1) \int \frac{d^2 \hat{k}_2}{4\pi} y_{\ell_2}^{m_2*}(\hat{k}_2) \\ &\times \int \frac{d^3 k_3}{(2\pi)^3} (2\pi)^3 \delta_{\text{D}}(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) \\ &\times \delta n(\mathbf{k}_1) \delta n(\mathbf{k}_2) \delta n_L^M(\mathbf{k}_3), \end{aligned} \quad (35)$$

where the triangle condition $\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3 = 0$ is satisfied through the delta function, and $\delta n_L^M(\mathbf{k})$ is the Fourier transform of the y_L^{M*} -weighted density fluctuation (equation 34): $\delta n_L^M(\mathbf{k}) = \int d^3 x e^{-i\mathbf{k}\cdot\mathbf{x}} \delta n_L^M(\mathbf{x})$ and $\delta n(\mathbf{k}) = \delta n_0^0(\mathbf{k})$. As mentioned in Section 2, the $L=0$ mode reduces to the estimator of the Legendre coefficients of the bispectrum:

$$\begin{aligned} \widehat{B}_\ell(k_1, k_2) &= \widehat{B}_{\ell \ell L=0}(k_1, k_2) \\ &= \frac{2\ell+1}{I} \int \frac{d^2 \hat{k}_1}{4\pi} \int \frac{d^2 \hat{k}_2}{4\pi} \mathcal{L}_\ell(\hat{k}_1 \cdot \hat{k}_2) \\ &\times \int \frac{d^3 k_3}{(2\pi)^3} (2\pi)^3 \delta_{\text{D}}(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) \\ &\times \delta n(\mathbf{k}_1) \delta n(\mathbf{k}_2) \delta n(\mathbf{k}_3). \end{aligned} \quad (36)$$

The normalization I in equation (35) is given by

$$I = \int d^3 x \bar{n}^3(\mathbf{x}). \quad (37)$$

This normalization depends on a grid-cell resolution to compute the mean density $\bar{n}(\mathbf{x})$, and hence, it is difficult to make I numerically converge to a certain value in a large survey volume. However, the value of I does not affect the final results, because three-point window functions, which are defined in Section 6 to estimate survey geometry effects, have the same normalization, and the theoretical model of the bispectrum is represented as a convolution of the theory and the window function. Namely, the normalization I cancels out when comparing the theory including the survey window effect with the measurements.

The angular integrations in equation (35) represent integration over a spherical shell in Fourier space centered at each bin $k = |\mathbf{k}|$,

$$\begin{aligned} \int \frac{d^2 \hat{k}}{4\pi} &= \frac{(2\pi)^3}{4\pi k^2} \int \frac{d^3 k'}{(2\pi)^3} \delta_{\text{D}}(|\mathbf{k}| - |\mathbf{k}'|) \\ &= \frac{1}{N_{\text{mode}}(k)} \sum_{k-\Delta k/2 < k' < k+\Delta k/2} \end{aligned} \quad (38)$$

where $N_{\text{mode}}(k)$ is the number of Fourier modes in each k -bin and Δk is the bin width.

The FFT algorithm to compute $\delta n_L^M(\mathbf{k})$ requires the interpolation of functions on a regular grid in position space. The Fourier transform of the y_L^{M*} -weighted density fluctuation measured by FFTs, $\delta n_L^M(\mathbf{k})|_{\text{FFT}}$, includes the effect of the mass assignment function $W_{\text{mass}}(\mathbf{k})$ (Jing 2005). Although there are some efforts to reduce the aliasing effect (e.g. Sefusatti et al. 2016), we adopt in this paper the simplest method to correct such effects: we divide $\delta n_L^M(\mathbf{k})|_{\text{FFT}}$ by $W_{\text{mass}}(\mathbf{k})$, namely $\delta n_L^M(\mathbf{k}) = \delta n_L^M(\mathbf{k})|_{\text{FFT}} / W_{\text{mass}}(\mathbf{k})$. The mass assignment window function is given by

$$W_{\text{mass}}(\mathbf{k}) = \prod_{i=x,y,z} \left[\text{sinc} \left(\frac{\pi k_i}{2k_N} \right) \right]^p, \quad (39)$$

where $\text{sinc}(x) = \sin(x)/x$, and $k_N = \pi/H_G$ is the Nyquist frequency with the grid spacing H_G on an axis. The indexes $p = 1, p = 2$,

and $p = 3$ correspond to the nearest grid point (NGP), cloud-in-cell (CIC), and triangular-shaped cloud (TSC) assignment functions, respectively (Hockney & Eastwood 1981). Appendix C investigates the differences between the three mass assignment functions.

The integrals in equation (35) do not decouple into a product of Fourier transforms due to the delta function $\delta_D(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3)$. When discretizing the integrals, they become double sums in the number of grid-cells, leading to an N_{grid}^2 process, where N_{grid} is the number of FFT grid-cells. To compute these integrals using FFTs, we insert the relation

$$(2\pi)^3 \delta_D(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) = \int d^3x e^{i\mathbf{x} \cdot (\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3)} \quad (40)$$

in equation (35), and then derive

$$\begin{aligned} \widehat{B}_{\ell_1 \ell_2 L}(k_1, k_2) &= H_{\ell_1 \ell_2 L} \sum_{m_1 m_2 M} \begin{pmatrix} \ell_1 & \ell_2 & L \\ m_1 & m_2 & M \end{pmatrix} \\ &\times \frac{N_{\ell_1 \ell_2 L}}{I} \int d^3x F_{\ell_1}^{m_1}(\mathbf{x}; k_1) F_{\ell_2}^{m_2}(\mathbf{x}; k_2) G_L^M(\mathbf{x}), \end{aligned} \quad (41)$$

where

$$\begin{aligned} F_{\ell}^m(\mathbf{x}; k) &= \int \frac{d^2\hat{k}}{4\pi} e^{i\hat{k} \cdot \mathbf{x}} y_{\ell}^{m*}(\hat{k}) \frac{\delta n|_{\text{FFT}}(\mathbf{k})}{W_{\text{mass}}(\mathbf{k})} \\ G_L^M(\mathbf{x}) &= \int \frac{d^3k}{(2\pi)^3} e^{i\mathbf{k} \cdot \mathbf{x}} \frac{\delta n_L^M|_{\text{FFT}}(\mathbf{k})}{W_{\text{mass}}(\mathbf{k})}. \end{aligned} \quad (42)$$

Since the angular integration in F_{ℓ}^m can be represented as a three-dimensional integral with a delta function through equation (38), it can be computed using inverse FFTs. Thus, this estimator requires only FFT, i.e. $\mathcal{O}(N_{\text{grid}} \ln N_{\text{grid}})$, processes (see also Scoccimarro 2015).

Once we have measured the bispectrum multipoles from galaxies using its estimator (equation 35), it is necessary to subtract from that the shot noise terms, which arise from a sampling of a catalogue with a finite number of galaxies. To estimate the shot noise in the bispectrum, we substitute equation (34) into equation (35) and write down the bispectrum estimator in the particle description:

$$\begin{aligned} \widehat{B}_{\ell_1 \ell_2 L}(k_1, k_2) &= H_{\ell_1 \ell_2 L} \sum_{m_1 m_2 M} \begin{pmatrix} \ell_1 & \ell_2 & L \\ m_1 & m_2 & M \end{pmatrix} \\ &\times \frac{N_{\ell_1 \ell_2 L}}{I} \int \frac{d^2\hat{k}_1}{4\pi} y_{\ell_1}^{m_1*}(\hat{k}_1) \int \frac{d^2\hat{k}_2}{4\pi} y_{\ell_2}^{m_2*}(\hat{k}_2) \\ &\times \int \frac{d^3k_3}{(2\pi)^3} (2\pi)^3 \delta_D(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) \\ &\times \left(\sum_i^{N_{\text{gal}}} -\alpha \sum_i^{N_{\text{ran}}} \right) \left(\sum_j^{N_{\text{gal}}} -\alpha \sum_j^{N_{\text{ran}}} \right) \left(\sum_k^{N_{\text{gal}}} -\alpha \sum_k^{N_{\text{ran}}} \right) \\ &\times w(\mathbf{x}_i) w(\mathbf{x}_j) w(\mathbf{x}_k) y_L^{M*}(\hat{x}_k) e^{-i\mathbf{k}_1 \cdot \mathbf{x}_i} e^{-i\mathbf{k}_2 \cdot \mathbf{x}_j} e^{-i\mathbf{k}_3 \cdot \mathbf{x}_k}, \end{aligned} \quad (43)$$

where \mathbf{x}_i , \mathbf{x}_j , and \mathbf{x}_k denote the positions of galaxies i , j , and k , respectively. Then, there are four cases to contribute to the shot noise: (1) $i = j = k$; (2) $i = k \neq j$; (3) $j = k \neq i$; and (4) $i = j \neq k$. The total shot noise is thus

$$\begin{aligned} S_{\ell_1 \ell_2 L}(k_1, k_2) &= S_{\ell_1 \ell_2 L}|_{i=j=k} + S_{\ell_1 \ell_2 L}|_{i=j \neq k} \\ &+ S_{\ell_1 \ell_2 L}|_{i=k \neq j} + S_{\ell_1 \ell_2 L}|_{j=k \neq i}. \end{aligned} \quad (44)$$

Each term of the right-hand side in the above expression is given by

$$\begin{aligned} S_{\ell_1 \ell_2 L}|_{i=j=k} &= \delta_{\ell_1 0}^{(K)} \delta_{\ell_2 0}^{(K)} \delta_{L 0}^{(K)} (1/I) \bar{S}_{L=0}^M \\ S_{\ell_1 \ell_2 L}|_{i \neq j=k}(k_1) &= \delta_{\ell_1 L}^{(K)} \delta_{\ell_2 0}^{(K)} \frac{2L+1}{I} \int \frac{d^2\hat{k}_1}{4\pi} y_L^M(\hat{k}_1) \\ &\times \left[\delta n(\mathbf{k}_1) N_L^{M*}(\mathbf{k}_1) - \frac{\bar{S}_L^M C_{\text{shot}}(\mathbf{k}_1)}{W_{\text{mass}}^2(\mathbf{k}_1)} \right] \\ S_{\ell_1 \ell_2 L}|_{i=k \neq j}(k_2) &= \delta_{\ell_2 L}^{(K)} \delta_{\ell_1 0}^{(K)} \frac{2L+1}{I} \int \frac{d^2\hat{k}_2}{4\pi} y_L^M(\hat{k}_2) \\ &\times \left[\delta n(\mathbf{k}_2) N_L^{M*}(\mathbf{k}_2) - \frac{\bar{S}_L^M C_{\text{shot}}(\mathbf{k}_2)}{W_{\text{mass}}^2(\mathbf{k}_2)} \right] \\ S_{\ell_1 \ell_2 L}|_{i=j \neq k}(k_1, k_2) &= H_{\ell_1 \ell_2 L} \sum_{m_1 m_2 M} \begin{pmatrix} \ell_1 & \ell_2 & L \\ m_1 & m_2 & M \end{pmatrix} \\ &\times \frac{N_{\ell_1 \ell_2 L}}{I} i^{\ell_1 + \ell_2} \int d^3x j_{\ell_1}(k_1 x) j_{\ell_2}(k_2 x) \\ &\times y_{\ell_1}^{m_1*}(\hat{x}) y_{\ell_2}^{m_2*}(\hat{x}) \int \frac{d^3k_3}{(2\pi)^3} e^{i\mathbf{k}_3 \cdot \mathbf{x}} \\ &\times \left[\delta n_L^M(\mathbf{k}_3) N_{0*}^M(\mathbf{k}_3) - \frac{\bar{S}_L^M C_{\text{shot}}(\mathbf{k}_3)}{W_{\text{mass}}^2(\mathbf{k}_3)} \right] \end{aligned} \quad (45)$$

where

$$\begin{aligned} \bar{S}_L^M &= \left(\sum_i^{N_{\text{gal}}} -\alpha^3 \sum_i^{N_{\text{ran}}} \right) [w(\mathbf{x}_i)]^3 y_L^{M*}(\hat{x}_i) \\ N_L^M(\mathbf{k}) &= \left(\sum_i^{N_{\text{gal}}} +\alpha^2 \sum_i^{N_{\text{ran}}} \right) [w(\mathbf{x}_i)]^2 y_L^M(\hat{x}_i) e^{-i\mathbf{k} \cdot \mathbf{x}_i}. \end{aligned} \quad (46)$$

The function $C_{\text{shot}}(\mathbf{k})$ is an analytical scale-dependent function to correct for the mass alignment effect on the shot noise (see equation 20 in Jing 2005). We expect that $S_{\ell_1 \ell_2 L}|_{i=j=k}$ should also depend on scales when taking into account the mass alignment effect and that we can derive its analytical expression similar to the power spectrum case. However, we leave more carefully investigations of these corrections for future work.

4.3 Three-point function

Now we move onto the derivation of the FFT-based estimator of the three-point function multipoles. Since the three-point function multipoles are related to the bispectrum multipoles through the Hankel transform, substituting equation (35) into equation (17) leads to

$$\begin{aligned} \widehat{\zeta}_{\ell_1 \ell_2 L}(r_1, r_2) &= H_{\ell_1 \ell_2 L} \sum_{m_1 m_2 M} \begin{pmatrix} \ell_1 & \ell_2 & L \\ m_1 & m_2 & M \end{pmatrix} \\ &\times \frac{N_{\ell_1 \ell_2 L}}{I} \int \frac{d^2\hat{r}_1}{4\pi} y_{\ell_1}^{m_1*}(\hat{r}_1) \int \frac{d^2\hat{r}_2}{4\pi} y_{\ell_2}^{m_2*}(\hat{r}_2) \\ &\times \int d^3x_1 \int d^3x_2 \int d^3x_3 \\ &\times \delta_D(\mathbf{r}_1 - \mathbf{x}_{13}) \delta_D(\mathbf{r}_2 - \mathbf{x}_{23}) \\ &\times y_L^{M*}(\hat{x}_3) \delta n(\mathbf{x}_1) \delta n(\mathbf{x}_2) \delta n(\mathbf{x}_3), \end{aligned} \quad (47)$$

where $\mathbf{x}_{13} = \mathbf{x}_1 - \mathbf{x}_3$ and $\mathbf{x}_{23} = \mathbf{x}_2 - \mathbf{x}_3$. The above expression can be rewritten into an FFT-based form

$$\begin{aligned} \widehat{\zeta}_{\ell_1 \ell_2 L}(r_1, r_2) &= H_{\ell_1 \ell_2 L} \sum_{m_1 m_2 M} \begin{pmatrix} \ell_1 & \ell_2 & L \\ m_1 & m_2 & M \end{pmatrix} \\ &\times \frac{N_{\ell_1 \ell_2 L}}{I} \int d^3 x F_{\ell_1}^{m_1}(\mathbf{x}; r_1) F_{\ell_2}^{m_2}(\mathbf{x}; r_2) G_L^M(\mathbf{x}), \end{aligned} \quad (48)$$

where

$$\begin{aligned} F_{\ell}^m(\mathbf{x}; r) &= i^{\ell} \int \frac{dk k^2}{2\pi^2} j_{\ell}(rk) F_{\ell}^m(\mathbf{x}; k) \\ &= i^{\ell} \int \frac{d^3 k}{(2\pi)^3} e^{i\mathbf{k}\cdot\mathbf{x}} j_{\ell}(rk) y_{\ell}^{m*}(\hat{\mathbf{k}}) \frac{\delta n|_{\text{FFT}}(\mathbf{k})}{W_{\text{mass}}(\mathbf{k})}. \end{aligned} \quad (49)$$

The $L = 0$ mode of equation (48) reproduces the estimator of the Legendre coefficients:

$$\widehat{\zeta}_{\ell}(r_1, r_2) = \widehat{\zeta}_{\ell \ell L=0}(r_1, r_2). \quad (50)$$

This $L = 0$ mode estimator is the same as the FFT-based one presented in Slepian & Eisenstein (2016).

When we compute the three-point function using FFTs, we need to subtract the shot-noise term from the measured three-point function, just like in the case of the bispectrum. While there are the four shot-noise terms in the bispectrum, three of them, $S|_{i \neq j = k}$, $S|_{i = k \neq j}$, and $S|_{i = j \neq k}$, contribute to the three-point function at $r_1 = 0$, $r_2 = 0$, and $r_1 = 0$ and $r_2 = 0$, respectively. Therefore, we need to consider only one term coming from $S|_{i = j \neq k}$:

$$\begin{aligned} S_{\ell_1 \ell_2 L}|_{i=j \neq k}(r_1, r_2) &= H_{\ell_1 \ell_2 L} \sum_{m_1 m_2 M} \begin{pmatrix} \ell_1 & \ell_2 & L \\ m_1 & m_2 & M \end{pmatrix} \\ &\times \frac{\delta_{r_1 r_2}^{(K)}}{N_{\text{mode}}(r)} \frac{N_{\text{grid}}}{V_{\text{FFT}}} (-1)^{\ell_1 + \ell_2} \frac{N_{\ell_1 \ell_2 L}}{I} \\ &\times \int \frac{d^2 \hat{r}_1}{4\pi} y_{\ell_1}^{m_1*}(\hat{r}_1) y_{\ell_2}^{m_2*}(\hat{r}_1) \int \frac{d^3 k_3}{(2\pi)^3} e^{i\mathbf{k}_3 \cdot \mathbf{r}_1} \\ &\times \left[\delta n_L^M(\mathbf{k}_3) N_0^{0*}(\mathbf{k}_3) - \frac{\bar{S}_L^M C_{\text{shot}}(\mathbf{k}_3)}{W_{\text{mass}}^2(\mathbf{k}_3)} \right], \end{aligned} \quad (51)$$

where V_{FFT} is the volume of the Cartesian box in which the galaxies are placed before the FFT is performed, and $N_{\text{mode}}(r)$ is the number of r -modes in each r -bin. Thus, the shot noise of the FFT-based estimator of the three-point function multipoles only contributes to the $r_1 = r_2$ case.

4.4 Estimators under the global plane-parallel approximation

In a situation like a cubic box simulation with periodic boundaries, we can define the global LOS direction and can choose the direction as the z -axis without loss of generality: the spherical harmonics with the LOS direction become $y_L^{M*}(\hat{n}) = y_L^{M*}(\hat{z}) = \delta_{M0}^{(K)}$. Under this situation, we no longer need to compute the y_L^{M*} -weighted density fluctuations but the normal density fluctuation $\delta n = \delta n_{L=0}^M$, and the estimators of the bispectrum and three-point function multipoles

are simplified to

$$\begin{aligned} \widehat{B}_{\ell_1 \ell_2 L}(k_1, k_2) &= H_{\ell_1 \ell_2 L} \sum_m \begin{pmatrix} \ell_1 & \ell_2 & L \\ m & -m & 0 \end{pmatrix} \\ &\times \frac{N_{\ell_1 \ell_2 L}}{I} \int d^3 x F_{\ell_1}^m(\mathbf{x}; k_1) F_{\ell_2}^{-m}(\mathbf{x}; k_2) G_0^0(\mathbf{x}) \\ \widehat{\zeta}_{\ell_1 \ell_2 L}(r_1, r_2) &= H_{\ell_1 \ell_2 L} \sum_m \begin{pmatrix} \ell_1 & \ell_2 & L \\ m & -m & 0 \end{pmatrix} \\ &\times \frac{N_{\ell_1 \ell_2 L}}{I} \int d^3 x F_{\ell_1}^m(\mathbf{x}; r_1) F_{\ell_2}^{-m}(\mathbf{x}; r_2) G_0^0(\mathbf{x}), \end{aligned} \quad (52)$$

where the normalization factor I is not represented by equation (37) but $I = (N_p^3/V_{\text{box}}^2)$, with V_{box} and N_p being the volume of a simulation box and the number of particles in the box, respectively.

Before closing this section, let us comment on an example of our computational time with our estimator. We carried out numerical computations on XC50 at Center for Computational Astrophysics, National Astronomical Observatory of Japan, which consists of a suite of computational nodes with the CPU of Intel Xeon Gold 6148 (20 cores with 2.4–3.7GHz).

The FFT-based algorithm given by equation (41) requires some FFT operations for each combination of k -bins, (k_1, k_2) and the summation of all possible m_1 , m_2 , and M modes. Consequently, the computational complexity of our estimator is $\mathcal{O}(N_{\text{multipole}} \times N_b^2 \times N_{\text{grid}} \ln N_{\text{grid}})$, where N_b is the number of k -bins and $N_{\text{multipole}}$ depends on the bispectrum multipoles that we compute.

For example, let us consider the lowest order bispectrum multipole, $B_{000}(k_1, k_2)$, with $N_b = 11$ and $N_{\text{grid}} = 512^3$, where the number of k -bin pairs of $B_{000}(k_1, k_2)$ is $N_b(N_b + 1)/2 = 66$ because of the symmetry between k_1 and k_2 . We are then able to compute B_{000} from a single mock (a single simulation box) in 1.22 (0.92) CPU hours using our serial code. As will be shown in Section 5.6, the diagonal part of B_{000} , i.e. $B_{000}(k_1 = k_2, k_2)$, dominates the signal to noise of the bispectrum monopole measurements. When we only compute this diagonal part that the number of k -bins is 11, measuring $B_{000}(k_1 = k_2, k_2)$ takes 0.23 (0.2) CPU hours for a mock (a box simulation), thus reducing enormously the computing time needed for the bispectrum measurements. For the higher multipoles, their computational time can be estimated by multiplying that of B_{000} by $N_{\text{multipole}}$, where $N_{\text{multipole}} = 1$ (1), 3 (3), 5 (5), 5 (1), 9 (3), and 9 (1) for B_{000} , B_{110} , B_{220} , B_{202} , B_{112} , and B_{404} for the mock (the box simulation).

Now here is the estimate of computational time to complete the bispectrum measurements for the BOSS survey that spans in two distinct sky regions (North and South Galactic Caps) in three redshift bins ($0.3 < z < 0.5$, $0.4 < z < 0.6$, and $0.5 < z < 0.75$), where each region has 2048 MultiDark-Patchy mock catalogues (Kitaura et al. 2016; Klypin et al. 2016). Sticking to four multipoles, B_{000} , B_{110} , B_{220} , and B_{202} , we can complete the bispectrum measurements in $(1 + 3 + 5 + 5) \times 6 \times 2048 \times 1.22/24 = 8745$ CPU days for the ‘full’ bispectrum (i.e. $k_1 \neq k_2$ is allowed) and 1650 CPU days for the ‘diagonal’ bispectrum (i.e. $k_1 = k_2$) only.

5 MEASUREMENTS

Having shown the new decomposition formalism of the bispectrum (Section 2) and its estimator (Section 4), we now turn to our primary goal to understand how the bispectrum multipoles provide

cosmological information. In Section 5.1, we briefly introduce the galaxy data set, the random catalogues used for corrections of survey geometry effects, and the mock catalogues used to estimate the covariance matrix. After setting the properties of our 3D Cartesian grid in Section 5.2, we will measure the bispectrum multipoles of the CMASS North Galactic Cap (NGC) sample and the corresponding mock catalogues in Sections 5.3 and 5.4. We then estimate the covariance matrix of the bispectrum multipoles in Section 5.5. Finally, we compute cumulative S/N ratios estimated from the measurements of the bispectrum multipoles to see the impact of the bispectrum measurements on the information content of galaxy clustering in Section 5.6.

5.1 DATA

We use the CMASS NGC sample with 586 003 galaxies in the redshift range $0.43 < z < 0.75$ (White et al. 2011; Parejko et al. 2013; Bundy et al. 2015; Leauthaud et al. 2016; Saito et al. 2016). This sample is drawn from the Data Release 12 (DR12; Alam et al. 2015) of the BOSS (Bolton et al. 2012; Dawson et al. 2013), which is part of the Sloan Digital Sky Survey III (SDSS-III; Eisenstein et al. 2011), and is selected from multicolour SDSS imaging (Fukugita et al. 1996; Gunn et al. 1998; Smith et al. 2002; Gunn et al. 2006; Doi et al. 2010). We also use the associated random catalogues that quantify the survey geometry of BOSS.

To estimate the errors on the bispectrum, namely the covariance matrix, we utilize the MultiDark-Patchy mock catalogues (MD-Patchy mocks; Kitaura et al. 2016; Klypin et al. 2016), which are designed to reproduce the BOSS CMASS data set. These mocks have been calibrated to an N -body-based reference sample using approximate gravity solvers and analytical–statistical biasing models and incorporate observational effects including the survey geometry, veto mask, and fiber collisions. We decided to use all 2048 mocks for the lowest order bispectrum multipole, B_{000} , and 600 mocks for the other multipoles. The fiducial cosmology for these mocks assumes a Λ CDM cosmology with $(\Omega_\Lambda, \Omega_m, \Omega_b, \sigma_8, h) = (0.692885, 0.307115, 0.048, 0.8288, 0.6777)$.

5.2 Prescription for measurements

We use the TSC assignment function to bin the CMASS NGC data set in a cubic box with a total volume of $(3.5 h^{-1} \text{Gpc})^3$ split into 512^3 grid cells. This corresponds to a grid-cell resolution of $\sim 6.8 h^{-1} \text{Mpc}$ and a Nyquist frequency of $k_N = 0.46 h \text{Mpc}^{-1}$. We then use the Fast Fourier Transform in the West (FFTW) library² to perform the FFTs. Appendix C provides a comparison of several assignment schemes, NGP, CIC, and TSC, with a different number of grid-cells (512^3 and 1024^3), showing that the TSC scheme with 512^3 grid cells achieves sub-per cent accuracy for $k \lesssim 0.2 h \text{Mpc}^{-1}$.

To correct for several observational artefacts in the catalogues, we use a completeness weight for each galaxy (Ross et al. 2012; Anderson et al. 2014; Reid et al. 2016),

$$w_c(\mathbf{x}) = w_{\text{sys}}(\mathbf{x}) (w_{\text{cp}}(\mathbf{x}) + w_{\text{noz}}(\mathbf{x}) - 1), \quad (53)$$

where \mathbf{x} is the observed galaxy position and w_{cp} , w_{noz} , and w_{sys} denote the redshift failure weight, the collision weight, and the angular systematics weight, respectively. The details about the observational systematic weights are described in Reid et al. (2016). Additionally, we use a signal-to-noise weight, the so-called FKP weight,

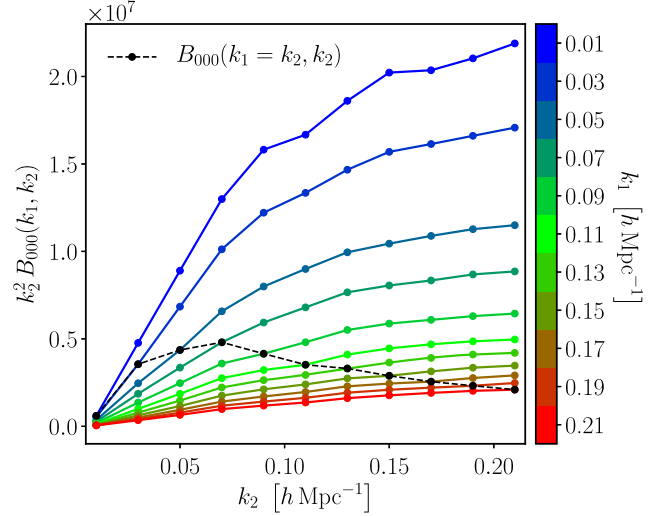


Figure 2. Means of the measurements of the bispectrum monopole, $B_{000}(k_1, k_2)$ (equation 41), from the MD-Patchy mocks for CMASS NGC, with the shot noise terms (equation 45) subtracted. For plotting purposes, we fix k_1 to the value at each k_1 -bin and plot $B_{000}(k_1, k_2)$ as a function of k_2 . Since the number of k_1 -bins is 11, we plot 11 coloured lines: from top to bottom, they denote $B_{000}(k_1 = 0.01, k_2)$, $B_{000}(k_1 = 0.03, k_2)$, ..., $B_{000}(k_1 = 0.21, k_2)$. We also plot the diagonal bispectrum monopole, $B_{000}(k_1 = k_2, k_2)$ (black dashed line), which intersects with $B_{000}(k_1, k_2)$ at each k_2 -bin.

proposed by Feldman et al. 1994, $w_{\text{FKP}}(\mathbf{x}) = 1/[1 + \bar{n}(\mathbf{x})P_0]$, with $P_0 = 10^4 (h^{-1} \text{Mpc})^3$. For Gaussian errors, this weight function works even for higher order statistics (Scoccimarro 2000). We investigate this weighting scheme in more detail in Appendix D. By multiplying the completeness weight by the FKP weight, we finally define the local weight function that we use in our analysis:

$$w(\mathbf{x}) = w_c(\mathbf{x}) w_{\text{FKP}}(\mathbf{x}). \quad (54)$$

We adopt a k -range of $0.01 h \text{Mpc}^{-1} \leq k \leq 0.21 h \text{Mpc}^{-1}$ in bins of $\Delta k = 0.02 h \text{Mpc}^{-1}$. The number of k -bins is 11 and the number of k -bin pairs for $B_{\ell_1 \ell_2 L}(k_1, k_2)$ is $11 \times (11 + 1)/2 = 66$ if $\ell_1 = \ell_2$, otherwise $11 \times 11 = 121$.

5.3 Lowest order of the bispectrum multipoles

We begin with measuring the lowest order of the bispectrum multipoles $B_{000}(k_1, k_2)$, which is hereafter referred to ‘bispectrum monopole’, from the MD-Patchy mocks. In Fig. 2, we show $B_{000}(k_1, k_2)$ measured using equation (41) after subtracting the shot noise terms (equation 45). For displaying purposes, we fix k_1 to the value at each k_1 -bin: $k_1 = 0.01, 0.03, \dots, 0.21$, and plot $B_{000}(k_1, k_2)$ as a function of k_2 . Each line in the figure is the mean bispectra in the mocks. From top (blue) to bottom (red) of the coloured lines, the value of the fixed k_1 increases. For comparison, we also plot the mean of the diagonal elements of $B_{000}(k_1, k_2)$, i.e. $B_{000}(k_1 = k_2, k_2)$ (black dashed), which intersects with $B_{000}(k_1, k_2)$ (coloured lines) at each k_2 -bin. In what follows, when we focus only on the diagonal elements, i.e. $B_{\ell_1 \ell_2 L}(k_1 = k_2, k_2)$, we refer to them as ‘diagonal bispectrum multipoles’.

5.4 Higher multipoles

We next turn to the higher order bispectrum multipoles, especially the $L \geq 2$ modes that characterize the anisotropic signal, which is induced by the RSD or AP effects. Fig. 3 presents the measurements

²<http://fftw.org>

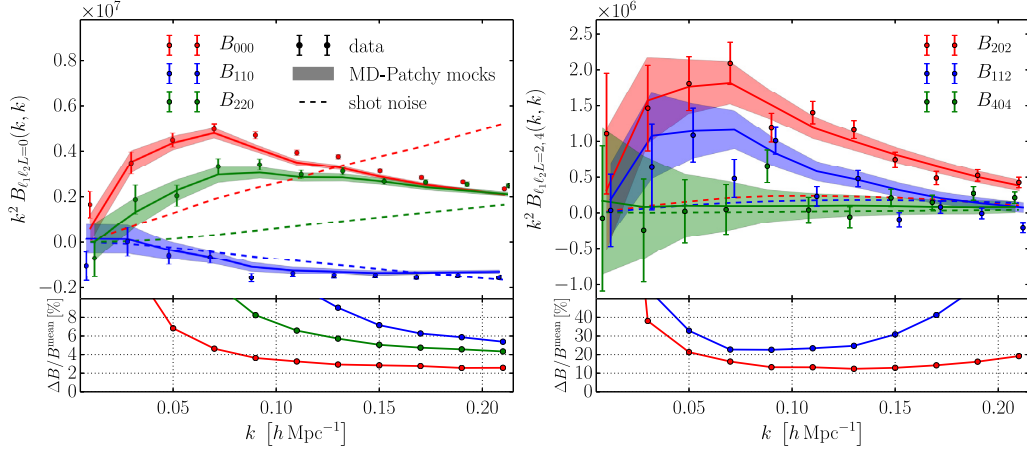


Figure 3. Diagonal bispectrum multipoles $B_{\ell_1 \ell_2 L}(k, k)$ for the $L = 0$ mode (left-hand panel) and the $L \geq 2$ modes (right-hand panel) measured from CMASS NGC, with the shot noise terms subtracted. The error bars on the data points are the 1σ errors, $\Delta B_{\ell_1 \ell_2 L}$, estimated from the MD-Patchy mocks. The shaded regions are the measurements from the MD-Patchy mocks with the 1σ errors. For comparison, we plot the means of the shot noises measured from the MD-Patchy mocks for each bispectrum multipole (dashed lines). This figure demonstrates reasonable agreement between the MD-Patchy mocks and the observed galaxy sample at large scales $k \lesssim 0.1 h \text{ Mpc}^{-1}$. The bottom parts in both the panels show the ratios of $\Delta B_{\ell_1 \ell_2 L}$ to the mean values of the measured $B_{\ell_1 \ell_2 L}$ in the mocks.

of the diagonal bispectrum multipoles from the CMASS NGC sample (data points). The error bars on the data points are the 1σ errors estimated from the MD-Patchy mocks (for details, see Section 5.5). The solid lines denote the mean measured bispectrum multipoles in the MD-Patchy mocks, and the shaded regions are the 1σ errors on the mean values. As will be shown in Section 5.6, the diagonal elements of the bispectrum monopole mainly contribute to the S/N of the full bispectrum monopole. We expect similar results for higher bispectrum multipoles, and therefore, we focus on the diagonal bispectrum multipoles in this subsection.

The left-hand panel in Fig. 3 displays the first three bispectrum multipoles of the $L = 0$ mode: B_{000} , B_{110} , and B_{220} , where they correspond to the Legendre coefficients of the bispectrum, $B_{\ell} = 0, 1, 2$ (equation 19). While we find agreement between the results from the galaxy sample and the mocks at large scales ($k \lesssim 0.1 h \text{ Mpc}^{-1}$), they start to significantly depart from each other beyond the 1σ errors at small scales ($k \gtrsim 0.1 h \text{ Mpc}^{-1}$). We will ignore this tension in the subsequent analysis, and leave a more careful calibration of the mock catalogues to the measure bispectra for future work.

In the right-hand panel of Fig. 3, we plot three bispectrum multipoles of the $L \geq 2$ modes: two in the $L = 2$ mode and one in the $L = 4$ mode. As mentioned in Section 2, the $L = 2$ and 4 modes can be interpreted as the analogues of the power spectrum quadrupole and hexadecapole, respectively, because they mainly yield from the terms proportional to $(\hat{n} \cdot \mathbf{v})^{L/2}$. As will be discussed in Section 5.6 in a quantitative context, we can measure and detect the anisotropic signal in the bispectrum $L = 2$ multipole with a high significance. On the other hand, the $L = 4$ mode seems consistent with the null hypothesis of no signal. These results are similar to those in the power spectrum case in the BOSS analysis (e.g. see Beutler et al. 2017): while the power spectrum quadrupole is measured with high statistical significance, the power spectrum hexadecapole is not detected.

Fig. 3 also includes the means of the shot noise terms measured from the MD-Patchy mocks for each bispectrum multipole (dashed lines). For the $L = 0$ mode, the shot noise terms are comparable to or larger than the corresponding bispectrum multipoles at

small scales ($k \sim 0.2 h \text{ Mpc}^{-1}$). This fact indicates that the shot noise dominates the errors on the $L = 0$ mode measurements at $k \sim 0.2 h \text{ Mpc}^{-1}$ in the BOSS analysis. Even for the $L > 0$ modes, the shot noise terms are not zero and significantly impact our measurements for all multipoles because of the anisotropic signals in the shot noise, i.e. $S_{\ell_1 \ell_2 L | i \neq j = k}$, $S_{\ell_1 \ell_2 L | i = k \neq j}$, and $S_{\ell_1 \ell_2 L | i = j \neq k}$ in equation (45).

The bottom parts in both panels of Fig. 3 show the ratios of the 1σ errors on the measured diagonal bispectrum multipole, $\Delta B_{\ell_1 \ell_2 L}$, to the means of $B_{\ell_1 \ell_2 L}$ measured in the MD-Patchy mocks. The values of these ratios depend on the bin width Δk , and smaller bin widths, e.g. $\Delta k = 0.01 h \text{ Mpc}^{-1}$, will induce higher values because the number of k -modes at each bin decreases. Nevertheless, these plots provide a rough estimate of accuracy required in the modeling of the bispectrum multipoles. To use smaller scales than $k = 0.1 h \text{ Mpc}^{-1}$ for the bispectrum analysis, we should achieve ~ 3 per cent accuracy for the theoretical model of B_{000} and ~ 10 per cent accuracy for B_{202} in the BOSS analysis.

5.5 Covariance matrices

Now we move onto investigating the properties of the covariance matrix between the bispectrum multipoles, which describe statistical uncertainties of the bispectrum measurements. Constraining cosmological parameters, we commonly analyze the bispectrum in combination with the power spectrum, and hence, we also study the cross-covariance matrix between the power spectrum multipoles and the bispectrum multipoles to see how the bispectrum is correlated with the power spectrum.

Let \mathbf{X} be a data vector of measured quantities. The covariance matrix for N_{mock} mocks is then estimated by

$$\mathbf{C} = \frac{1}{N_{\text{mock}} - 1} \sum_n^{N_{\text{mock}}} (\mathbf{X}^{(n)} - \bar{\mathbf{X}})^T (\mathbf{X}^{(n)} - \bar{\mathbf{X}}), \quad (55)$$

where $\mathbf{X}^{(n)}$ is the data vector obtained from the n -th mock, and the mean expectation value over the mocks is given by $\bar{\mathbf{X}} = (1/N_{\text{mock}}) \sum_n^{N_{\text{mock}}} \mathbf{X}^{(n)}$. The correlation matrix \mathbf{r} is defined as $r_{ij} = C_{ij} / (C_{ii} C_{jj})^{1/2}$, where r_{ij} and C_{ij} are the ij -element of the correlation

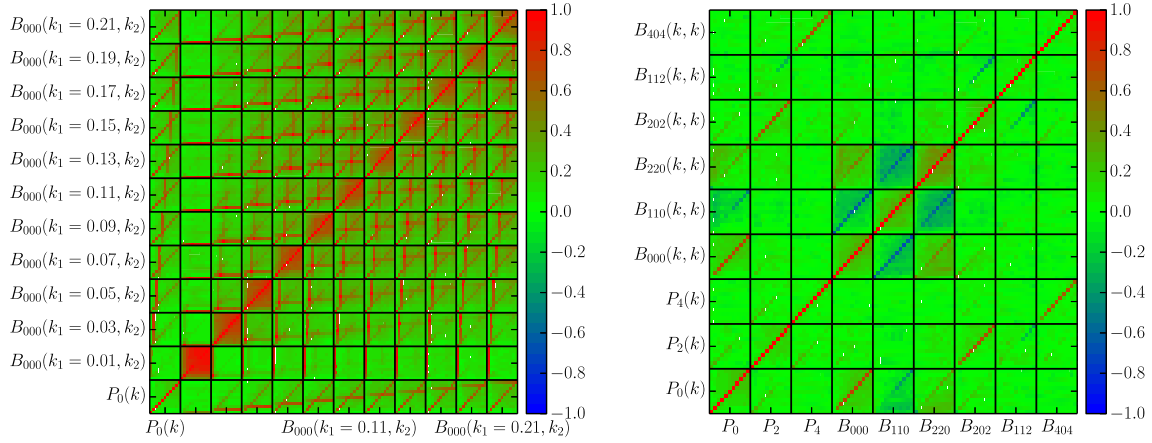


Figure 4. Correlation matrices of the power spectrum and bispectrum multipoles, estimated from the MD-Patchy mocks for CMASS NGC. The left-hand panel focuses on the power spectrum and bispectrum monopoles, $P_0(k)$ and $B_{000}(k_1, k_2)$; the right-hand panel shows the correlations among higher multipole terms of the power spectrum and the diagonal bispectrum. Each block of both the panels separated by horizontal and vertical division lines contains 11 k -bins in the k -range $k = 0.01 - 0.21 h \text{ Mpc}^{-1}$. The colour indicates the level of correlation, where red and green represent high and low correlations, respectively. Remarkably, the diagonal bispectrum multipoles $B_{\ell_1 \ell_2 L}(k, k)$ estimated at different scales, i.e. at k and k' for $k \neq k'$, are nearly uncorrelated like the power spectrum case, implying that the diagonal ones would provide dominant information on bispectrum measurements.

matrix \mathbf{r} and the covariance matrix \mathbf{C} , respectively, and the indexes i and j run over the number of bins. The case $i = j$ of the covariance, C_{ii} , gives the variance. The error bars shown in Fig. 3 are the square root of the variance, $C_{ii}^{1/2}$, of the data vector that consists of the diagonal bispectrum multipoles, $X_i = B_{\ell_1 \ell_2 L}(k_i, k_i)$.

First we focus on the power spectrum and bispectrum monopoles, namely P_0 and B_{000} . To represent $B_{000}(k_1, k_2)$ depending on k_1 and k_2 , we fix k_1 to each k_1 -bin value and describe it as a function of k_2 (see Section 5.3). We then consider the data set that consists of

$$X_1 = \{P_0(k), B_{000}(k_1 = 0.01, k_2), B_{000}(k_1 = 0.03, k_2), \dots, B_{000}(k_1 = 0.19, k_2), B_{000}(k_1 = 0.21, k_2)\}, \quad (56)$$

where each component of the data set, $P_0(k)$ or $B_{000}(k_1, k_2)$, contains 11 k -bins in the k -range from $k = 0.01 h \text{ Mpc}^{-1}$ to $k = 0.21 h \text{ Mpc}^{-1}$. The left-hand panel of Fig. 4 displays the correlation matrix of the above data set, which is divided into 12×12 blocks by 11 horizontal and vertical division lines. Each block contains 11×11 elements because of the 11 k -bins. The covariance matrix of the bispectrum monopole depends on four variables: k_1, k_2, k'_1 , and k'_2 . We find from the left-hand panel that there are strong correlations among the bispectrum monopoles estimated at given scales if $k_1 = k'_1$ or $k_2 = k'_2$, i.e. between $B_{000}(k_1, k_2)$ and $B_{000}(k_1, k'_2)$ and between $B_{000}(k_1, k_2)$ and $B_{000}(k'_1, k_2)$. Similarly, the bispectrum monopole $B_{000}(k_1, k_2)$ and the power spectrum monopole $P_0(k)$ are correlated if $k_1 = k$ or $k_2 = k$, i.e. between $B_{000}(k_1, k_2)$ and $P(k_1)$ and between $B_{000}(k_1, k_2)$ and $P(k_2)$. Strong correlation is observed whenever one of the two bispectrum scales k_1 and k_2 is matched with one of the scales of the other data vector. These results demonstrate that using only a few terms of $B_{000}(k_1, k_2)$ with different values of k_1 is a poor choice, and the full correlations considering all k_1 values should be used for any quantitative analysis.

Secondly, we consider the correlation matrix including higher multipoles of the power spectrum and bispectrum. Here, we only use the diagonal bispectrum multipoles $B_{\ell_1 \ell_2 L}(k, k)$. The data set then consists of

$$X_2 = \{P_0(k), P_2(k), P_4(k), B_{000}(k, k), B_{110}(k, k), B_{220}(k, k), B_{202}(k, k), B_{112}(k, k), B_{404}(k, k)\}. \quad (57)$$

The right-hand panel of Fig. 4 shows the correlation matrix of X_2 , which is divided into 9×9 blocks. A remarkable feature in the panel is that the values of off-diagonal elements in each block are very small, implying that the estimates of the diagonal bispectrum multipoles are nearly uncorrelated unless $k = k'$. We can reproduce this characteristic feature using analytical calculations in perturbation theory (Sugiyama et al., in preparation). We expect from this result that the diagonal bispectrum multipoles dominate the S/N on bispectrum measurements (see Section 5.6). As expected in Section 2, we find strong correlations between P_ℓ and $B_{\ell_1 \ell_2 L}$ if $\ell = L$, otherwise weak correlations.

5.6 Signal-to-noise ratios

A useful way to quantify the information context of galaxy clustering is to estimate cumulative S/Ns associated with the galaxy clustering measurements. The feature of the cumulative S/N has been well studied using simulations for the power spectrum alone (Rimes & Hamilton 2005, 2006; Takahashi et al. 2009; Blot et al. 2015), and for a combination of the power spectrum and the bispectrum (Byun et al. 2017; Chan & Blot 2017). In this subsection, we compute the ‘observed’ S/Ns by fitting the measured bispectrum multipoles from CMASS NGC to those from the MD-Patchy mocks.

5.6.1 Joint analysis of the power spectrum and bispectrum monopoles

To illustrate how the bispectrum measurements provide cosmological information on galaxy clustering in addition to the power spectrum measurements, we estimate the cumulative S/N of the amplitude of density fluctuations in the joint analysis of the power spectrum and bispectrum monopoles. Namely, we assume that the observed density contrast δ_{obs} is expressed by its theoretical prediction δ_{theory} times a free parameter α : $\delta_{\text{obs}} = \alpha \delta_{\text{theory}}$, and define the S/N as the statistical significance of α . We then fit the measurements of P_0 and B_{000} from CMASS NGC to $\alpha^2 P_0^{\text{mean}}$ and $\alpha^3 B_{000}^{\text{mean}}$, where P_ℓ^{mean} and $B_{\ell_1 \ell_2 L}^{\text{mean}}$ are the mean values in the MD-Patchy mocks. We compute the covariance matrices of P_0 and B_{000} from the MD-Patchy mocks with the fiducial value of α being

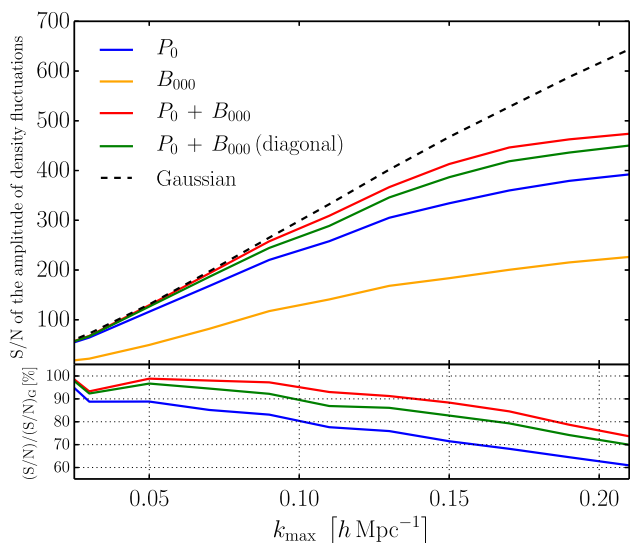


Figure 5. Cumulative signal-to-noise ratios of the amplitude of the measured density fluctuations from CMASS NGC as a function of k_{\max} . The blue line shows the S/N derived from the power spectrum monopole alone. The orange line shows the bispectrum monopole alone. Combining the power spectrum and bispectrum monopoles leads to the red line. The black dashed line shows the theoretical prediction using the Gaussian approximation, i.e. the linear power spectrum alone. For comparison, we also plot the result from the joint analysis using the diagonal bispectrum monopole, $B_{000}(k, k)$ (green solid line). The bottom part shows the ratios of the estimated S/Ns to the Gaussian prediction. This figure shows that while the S/N of the power spectrum alone is degraded to ~ 80 per cent at $k_{\max} = 0.1 h \text{ Mpc}^{-1}$ and ~ 60 per cent at $k_{\max} = 0.2 h \text{ Mpc}^{-1}$ compared to the Gaussian prediction, the additional information of the bispectrum recovers the S/N up to ~ 95 per cent at $k_{\max} = 0.1 h \text{ Mpc}^{-1}$ and ~ 75 per cent at $k_{\max} = 0.2 h \text{ Mpc}^{-1}$. Furthermore, we find that the joint analysis using only the diagonal elements of the bispectrum monopole (green line) already gets very close to the red line, indicating that the diagonal bispectrum provides dominant information on the bispectrum measurements.

unity. We perform the standard likelihood analysis through the likelihood function $L \propto \exp(-\chi^2/2)$ with the χ^2 statistics and estimate the mean of α , which is represented as $\langle \alpha \rangle$, and its 1σ error, $\Delta\alpha$, with the flat prior $0 \leq \alpha \leq 10$. In this analysis, we account for a Hartlap factor (Hartlap, Simon & Schneider 2007) and a Percival factor (Percival et al. 2014) (for details, see Appendix E). Finally, we define the S/N as $S/N \equiv \langle \alpha \rangle / \Delta\alpha$. We show the S/N as a function of the maximum wavenumber k_{\max} with fixing the minimum wavenumber to $k_{\min} = 0.01 h \text{ Mpc}^{-1}$. For comparison, we also conduct the same analysis using only P_0 or B_{000} .

Fig. 5 compares the S/Ns computed by several analyses. The blue solid line shows the analysis considering only the power spectrum monopole, the orange solid line represents the analysis using only the bispectrum monopole, and the red solid line denotes the joint analysis of the power spectrum and bispectrum monopoles, where the data set used in the joint analysis is given by equation (56). For comparison, we also plot a theoretical prediction using the power spectrum monopole including the shot-noise term in the Gaussian limit (black dashed). The bottom part of this figure shows the ratios of the estimated S/Ns to the Gaussian prediction. If the galaxy distribution is entirely Gaussian, the power spectrum encodes a complete description of galaxy clustering. In this sense, the black dashed line shows the upper limit of the S/N that we can obtain from the CMASS NGC sample. In fact, we observe the well-known non-linear degradation of the S/N of the power spectrum. The degra-

dition is ~ 80 per cent at $k_{\max} = 0.1 h \text{ Mpc}^{-1}$ and ~ 60 per cent at $k_{\max} = 0.2 h \text{ Mpc}^{-1}$ compared to the Gaussian prediction. We find that the additional information of the bispectrum monopole recovers the S/N up to ~ 95 per cent at $k_{\max} = 0.1 h \text{ Mpc}^{-1}$ and ~ 75 per cent at $k_{\max} = 0.2 h \text{ Mpc}^{-1}$.

We also consider another joint analysis, using the diagonal bispectrum monopole (the green solid line shown in Fig. 5), where the data set we use consists of $X = \{P_0(k), B_{000}(k, k)\}$. We stress that the number of bins of the diagonal bispectrum monopole, which is 11, is smaller than that of the full bispectrum monopole depending on k_1 and k_2 by a factor of 6. Nevertheless, this analysis well recovers the S/N up to ~ 90 per cent at $k_{\max} = 0.1 h \text{ Mpc}^{-1}$ and ~ 70 per cent at $k_{\max} = 0.2 h \text{ Mpc}^{-1}$. This is because the diagonal bispectrum allows us to extract nearly independent information at different scales (Section 5.5). We conclude from this result that the diagonal bispectrum monopole dominates the signal to noise of the bispectrum monopole measurements, and we expect similar results for higher bispectrum multipoles.

5.6.2 Bispectrum multipoles

Also of interest is the detectability of the bispectrum multipoles themselves, especially for the $L \neq 0$ modes arising from the RSD or AP effects. To show this, we apply the same analysis as done in Section 5.6.1 to the diagonal bispectrum multipoles, where we adopt the theoretical templates as $\beta B_{\ell_1 \ell_2 L}^{\text{mean}}$ with β being a varying parameter, and the S/N of the bispectrum multipoles is then defined as $S/N = \langle \beta \rangle / \Delta\beta$.

The left-hand panel of Fig. 6 shows the S/Ns estimated from the first three multipoles of the $L = 0$ mode: B_{000} (red), B_{110} (blue), and B_{220} (green), whose measurements are displayed by the same coloured data points with the 1σ error bars in the left-hand panel of Fig. 3. As expected, the largest S/N arises from the bispectrum monopole. The S/Ns of both B_{110} and B_{220} are about half of the S/N of B_{000} . Interestingly, we find the S/N of B_{110} to be smaller than that of B_{220} . This is because the amplitude of B_{110} is about 1.5 times smaller than B_{220} at e.g. $k = 0.2 h \text{ Mpc}^{-1}$ (Fig. 3), while the 1σ error of B_{110} is about 1.25 times smaller than B_{220} at that scale.

The right-hand panel of Fig. 6 shows the S/Ns of three diagonal bispectrum multipoles with $L \geq 2$: B_{202} (red), B_{112} (blue), and B_{404} (green), whose measurements are displayed in the right-hand panel of Fig. 3. We find that the S/Ns of these $L > 0$ modes are all smaller than those of the first three multipoles of the $L = 0$ mode. For the $L = 2$ mode, the S/N of B_{202} is about two times as large as that of B_{112} , showing that B_{202} dominates the signal-to-noise of the $L = 2$ mode measurements. The lowest order of the $L = 4$ mode, B_{404} , is consistent with the null hypothesis of no signal: $S/N < 2$ up to $k_{\max} = 0.15 h \text{ Mpc}^{-1}$ and $S/N < 3.5$ at $k_{\max} = 0.21 h \text{ Mpc}^{-1}$.

The most striking feature of the right-hand panel is that the S/N of B_{202} is 14 at $k_{\max} > 0.15 h \text{ Mpc}^{-1}$, corresponding to a 14σ detection of the anisotropic bispectrum signal in the lowest order $L = 2$ mode. We have also detected the next leading order $L = 2$ mode, B_{112} , at the 6σ level. Since we use the mean bispectra measured from the MD-Patchy mocks as the theoretical template to fit the measurements, we ignore the AP effect in this analysis. Nevertheless, we can well fit the measurements with $\chi^2/\text{d.o.f.} = 9.3/7$ for B_{202} at $k_{\max} = 0.15 h \text{ Mpc}^{-1}$ and $\chi^2/\text{d.o.f.} = 11.5/6$ for B_{112} at $k_{\max} = 0.13 h \text{ Mpc}^{-1}$, indicating that the anisotropic signals detected here are well explained by our models of the RSD effect alone without the AP effect, i.e. by the mean bispectra measured from the MD-Patchy mocks.

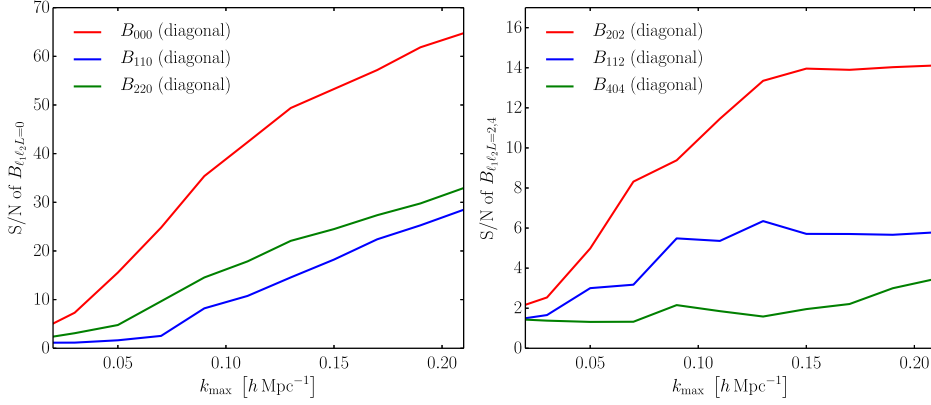


Figure 6. Cumulative S/Ns of the diagonal bispectrum multipoles. The left-hand panel shows the S/Ns for the $L = 0$ mode of the bispectrum multipoles; the right-hand panel plots those for the $L \geq 2$ modes. For each $L = 0$ and $L = 2$ mode, the lowest order of the bispectrum multipoles, B_{000} ($L = 0$) or B_{202} ($L = 2$), provides the highest S/N in the L mode. This figure shows a 14σ detection of B_{202} , which is induced only by anisotropic signals on the bispectrum (the RSD or AP effects).

6 SURVEY WINDOW FUNCTIONS

In this section, we discuss the effects of survey geometry on the bispectrum. There have been a lot of efforts to include the survey geometry effect in Fourier-space data analysis because the survey geometry distorts the observed density fluctuation and can introduce spurious anisotropic signals. To quantify the survey geometry distortion, Beutler et al. (2014), for the first time, proposed to expand the survey window functions in Legendre polynomials and defined the window function multipoles. This treatment was further developed by Wilson et al. (2017), based on configuration-space calculations. Sugiyama et al. (2018) extended the Wilson et al. (2017) formalism to the Bipolar spherical harmonic expansion formalism to explore the rotational invariance breaking of the survey geometry. We apply these formalisms, which have been developed in the power spectrum analysis, to the bispectrum analysis and present a framework how one can correct for the survey geometry effect when theoretically modeling the bispectrum.

The outline of this section is as follows. In Section 6.1, we define a three-point window function in configuration space and decompose it into the TripoSH basis with zero total angular momentum like the three-point function multipoles given in equation (16). In Section 6.2, we show the derivation of the bispectrum multipoles including the survey window corrections. In Section 6.3, we compute the masked bispectrum multipoles in the second-order perturbation theory and compare them with the measurements from simulations with and without the survey geometry effect.

6.1 Window function multipoles

We define the multipole components of the three-point window function in configuration space by replacing the density fluctuation δn (equation 47) by the mean number density measured from a random catalogue \bar{n} (equation 32):

$$\begin{aligned}
 Q_{\ell_1 \ell_2 L}(r_1, r_2) &= H_{\ell_1 \ell_2 L} \sum_{m_1 m_2 M} \begin{pmatrix} \ell_1 & \ell_2 & L \\ m_1 & m_2 & M \end{pmatrix} \\
 &\times \frac{N_{\ell_1 \ell_2 L}}{I} \int \frac{d^2 \hat{r}_1}{4\pi} y_{\ell_1}^{m_1*}(\hat{r}_1) \int \frac{d^2 \hat{r}_2}{4\pi} y_{\ell_2}^{m_2*}(\hat{r}_2) \\
 &\times \int d^3 x_1 \int d^3 x_2 \int d^3 x_3 \\
 &\times \delta_{\text{D}}(\mathbf{r}_1 - \mathbf{x}_{13}) \delta_{\text{D}}(\mathbf{r}_2 - \mathbf{x}_{23}) \\
 &\times y_L^{M*}(\hat{x}_3) \bar{n}(\mathbf{x}_1) \bar{n}(\mathbf{x}_2) \bar{n}(\mathbf{x}_3). \quad (58)
 \end{aligned}$$

These three-point window function multipoles can be computed by the same algorithm used to measure the three-point function multipoles, allowing us to use FFTs.

In this work, we measure four components of $Q_{\ell_1 \ell_2 L}$ from the random catalogue associated with CMASS NGC and plot them in Fig. 7: Q_{000} (solid) and Q_{110} (dashed) of the $L = 0$ mode in the left-hand panel; Q_{202} (solid) and Q_{112} (dashed) of the $L = 2$ mode in the right-hand panel. To display the window function multipoles depending on two comoving distances r_1 and r_2 , we fix r_1 to $r_1 = 0$ (red), 200 (green), and $550 h^{-1}$ Mpc (blue), and plot $Q_{\ell_1 \ell_2 L}(r_1, r_2)$ as a function of r_2 . On large scales going beyond the survey regime, all of the window function multipoles should become zero, because we cannot find pairs (triplets) of galaxies at these scales, which is the so-called survey edge effect. On the other hand, on small scales where the survey edge effect no longer matters, the window function multipoles that depend only on r_2 after fixing r_1 will be constant if $\ell_2 = 0$ or zero for $\ell_2 \geq 1$. In particular, $Q_{000}(r_1 = 0, r_2)$, by definition, approaches unity in the limit $r_2 \rightarrow 0$. Since the window function multipoles are symmetric under $r_1 \leftrightarrow r_2$ and $\ell_1 \leftrightarrow \ell_2$, i.e. $Q_{\ell_1 \ell_2 L}(r_1, r_2) = Q_{\ell_2 \ell_1 L}(r_2, r_1)$, for example, $Q_{000}(r_1 = 550, r_2)$ does not become unity at $r_2 = 0$ but becomes the same as $Q_{000}(r_1 = 0, r_2 = 550)$. We expect similar results for all of the other multipoles and any survey regions such as CMASS SGC, LOWZ NGC, and LOWZ SGC, even though their exact shapes will depend on each survey geometry.

6.2 Masked bispectrum multipoles

Now we are ready to derive the bispectrum multipoles including survey window corrections. To do so, we first compute the ensemble average of the estimator of the three-point function multipoles (equation 47):

$$\begin{aligned}
 \langle \hat{\xi}_{\ell_1 \ell_2 L}(r_1, r_2) \rangle &= H_{\ell_1 \ell_2 L} \sum_{m_1 m_2 M} \begin{pmatrix} \ell_1 & \ell_2 & L \\ m_1 & m_2 & M \end{pmatrix} \\
 &\times \frac{N_{\ell_1 \ell_2 L}}{I} \int \frac{d^2 \hat{r}_1}{4\pi} y_{\ell_1}^{m_1*}(\hat{r}_1) \int \frac{d^2 \hat{r}_2}{4\pi} y_{\ell_2}^{m_2*}(\hat{r}_2) \\
 &\times \int d^3 x_1 \int d^3 x_2 \int d^3 x_3 y_L^{M*}(\hat{x}_3) \\
 &\times \delta_{\text{D}}(\mathbf{r}_1 - \mathbf{x}_{13}) \delta_{\text{D}}(\mathbf{r}_2 - \mathbf{x}_{23}) \\
 &\times \langle \delta n(\mathbf{x}_1) \delta n(\mathbf{x}_2) \delta n(\mathbf{x}_3) \rangle. \quad (59)
 \end{aligned}$$

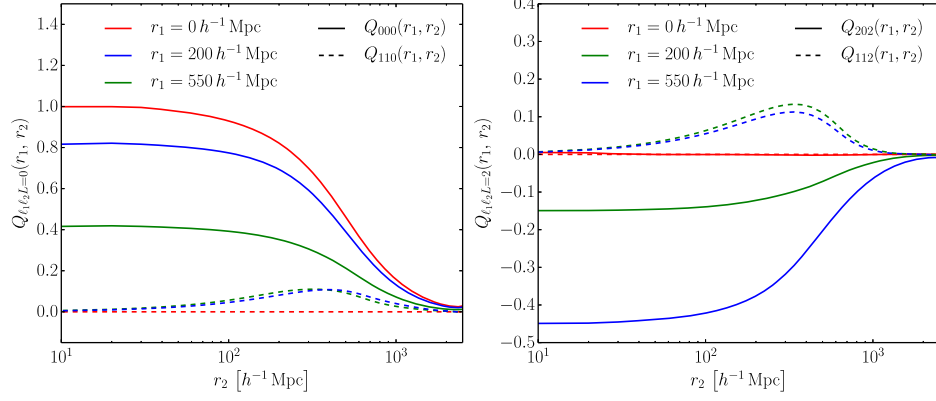


Figure 7. Multipole components of the three-point survey window function, $Q_{\ell_1 \ell_2 L}(r_1, r_2)$, as given in equation (58), which are estimated for survey geometry of CMASS NGC. For display purposes, we fix r_1 to 0 (red), 200 (green), and $550 h^{-1}$ Mpc (blue), and plot $Q_{\ell_1 \ell_2 L}(r_1, r_2)$ as a function r_2 . We show four components: Q_{000} (solid) and Q_{110} (dashed) in the left-hand panel; Q_{202} (solid) and Q_{112} (dashed) in the right-hand panel. At large scales, all of the window function multipoles become zero, while at small scales, they become constant if $\ell_2 = 0$, otherwise zero.

The theoretical expression of the three-point function of the observed density fluctuation δn (equation 33) is described as

$$\langle \delta n(\mathbf{x}_1) \delta n(\mathbf{x}_2) \delta n(\mathbf{x}_3) \rangle = \bar{n}(\mathbf{x}_1) \bar{n}(\mathbf{x}_2) \bar{n}(\mathbf{x}_3) \left[\zeta(\mathbf{x}_{13}, \mathbf{x}_{23}, \hat{x}_3) - \bar{\zeta} \right], \quad (60)$$

where we applied the local plane-parallel approximation and chose \hat{x}_3 as the LOS direction. $\bar{\zeta}$, which is the so-called integral constraint (Peacock & Nicholson 1991), which comes from the difference between the measured mean density from a finite survey volume and the true value. We determine the integral constraint to satisfy the condition:

$$\int dr_1 r_1^2 \int dr_2 r_2^2 \langle \widehat{\zeta}_{000}(r_1, r_2) \rangle = 0. \quad (61)$$

The three-point function depending on two relative coordinates, \mathbf{x}_{13} and \mathbf{x}_{23} , and one LOS direction, \hat{x}_3 , can be decomposed into the TripoSH basis with zero total angular momentum (equation 15):

$$\zeta(\mathbf{x}_{13}, \mathbf{x}_{23}, \hat{x}_3) = \sum_{\ell_1 + \ell_2 + L = \text{even}} \zeta_{\ell_1 \ell_2 L}(|\mathbf{x}_{13}|, |\mathbf{x}_{23}|) S_{\ell_1 \ell_2 L}(\hat{x}_{13}, \hat{x}_{23}, \hat{x}_3). \quad (62)$$

Inserting equations (60) and (62) in equation (59), and using equations (A4) and (A5), we obtain the masked three-point function multipoles:

$$\begin{aligned} \langle \widehat{\zeta}_{\ell_1 \ell_2 L}(r_1, r_2) \rangle_{\text{model}} &= N_{\ell_1 \ell_2 L} \sum_{\ell'_1 + \ell'_2 + L' = \text{even}} \sum_{\ell''_1 + \ell''_2 + L'' = \text{even}} \\ &\times \begin{Bmatrix} \ell'_1 & \ell'_2 & L' \\ \ell_1 & \ell_2 & L \end{Bmatrix} \left[\frac{H_{\ell_1 \ell_2 L} H_{\ell'_1 \ell'_2 L'} H_{\ell''_1 \ell''_2 L''}}{H_{\ell'_1 \ell'_2 L'} H_{\ell''_1 \ell''_2 L''}} \right] \\ &\times Q_{\ell'_1 \ell'_2 L'}(r_1, r_2) \zeta_{\ell'_1 \ell'_2 L'}(r_1, r_2) \\ &- Q_{\ell_1 \ell_2 L}(r_1, r_2) \bar{\zeta}, \end{aligned} \quad (63)$$

where 'model' means that this masked model will be compared with the measured estimator. The bracket with 9 multipole indices, $\{ \dots \}$, denotes the Wigner-9j symbol whose definition is given in equation (A5). The integral constraint in the above expression is

given by

$$\bar{\zeta} = \frac{1}{\int dr_1 r_1^2 \int dr_2 r_2^2 Q_{000}(r_1, r_2)} \sum_{\ell_1 + \ell_2 + L = \text{even}} \int dr_1 r_1^2 \int dr_2 r_2^2 \frac{Q_{\ell_1 \ell_2 L}(r_1, r_2) \zeta_{\ell_1 \ell_2 L}(r_1, r_2)}{H_{\ell_1 \ell_2 L}^2 N_{\ell_1 \ell_2 L}}. \quad (64)$$

We can schematically represent the masked three-point function multipoles as follows:

$$\langle \widehat{\zeta}_{\ell_1 \ell_2 L}(r_1, r_2) \rangle_{\text{model}} = Q_{000}(r_1, r_2) \zeta_{\ell_1 \ell_2 L}(r_1, r_2) - Q_{\ell_1 \ell_2 L}(r_1, r_2) \bar{\zeta} + [\text{the other multipole components}]. \quad (65)$$

On the right-hand side of the above expression, the first term, $Q_{000} \zeta_{\ell_1 \ell_2 L}$, is directly related to what we measure via the estimator $\widehat{\zeta}_{\ell_1 \ell_2 L}$. This term only includes the survey edge effect (Fig. 7). The second term, $Q_{\ell_1 \ell_2 L} \bar{\zeta}$, is the integral constraint correction. The survey window function does not only distort the individual multipoles, but also correlate different multipoles (the second line in the above expression). For an example, the masked three-point function monopole is given by (see also Appendix F)

$$\begin{aligned} \langle \widehat{\zeta}_{000}(r_1, r_2) \rangle_{\text{model}} &= Q_{000}(r_1, r_2) [\zeta_{000}(r_1, r_2) - \bar{\zeta}] \\ &+ \frac{1}{3} Q_{110}(r_1, r_2) \zeta_{110}(r_1, r_2) \\ &+ \frac{1}{5} Q_{220}(r_1, r_2) \zeta_{220}(r_1, r_2) + \dots \end{aligned} \quad (66)$$

In the configuration-space analysis, it is not common to normalize the three-point function using the factor I (equation 37) but the window function monopole $Q_{000}(r_1, r_2)$. equation (65) then becomes

$$\frac{\langle \widehat{\zeta}_{\ell_1 \ell_2 L}(r_1, r_2) \rangle_{\text{model}}}{Q_{000}(r_1, r_2)} = \zeta_{\ell_1 \ell_2 L}(r_1, r_2) + [\text{the other correction terms}]. \quad (67)$$

Thus, we can remove the window effect from the first term, but the other correction terms remain. As far as we measure $\zeta_{\ell_1 \ell_2 L}$ via the FFT-based estimator (equation 48) and fit it to a theoretical model, the above expression is necessary.³

³By solving the survey edge correction equation, we can correct for the survey geometry effect in configuration space only using the measured

Finally, we obtain the masked bispectrum multipoles through the Hankel transform of $\langle \widehat{\zeta}_{\ell_1 \ell_2 L} \rangle_{\text{model}}$:

$$\begin{aligned} & \left\langle \widehat{B}_{\ell_1 \ell_2 L}(k_1, k_2) \right\rangle_{\text{model}} \\ &= (-i)^{\ell_1 + \ell_2} (4\pi)^2 \int dr_1 r_1^2 \int dr_2 r_2^2 \\ & \quad \times j_{\ell_1}(k_1 r_1) j_{\ell_2}(k_2 r_2) \left\langle \widehat{\zeta}_{\ell_1 \ell_2 L}(r_1, r_2) \right\rangle_{\text{model}}. \end{aligned} \quad (68)$$

This masked bispectrum model is used to compare the theory with the measured bispectrum estimator during the likelihood fitting. There are two points to note about the above expression. The first one is that the masked bispectrum multipoles are proportional to $1/I$ (equation 37) through the window function multipoles. Therefore, comparing the theory with the measurements, the value of I no longer matters as mentioned in Section 4. The second one is that the masked bispectrum multipoles are derived using the local plane-parallel approximation in equation (60). This means that we can take into account the local plane-parallel approximation through the masked bispectrum multipoles as mentioned in Section 2, even though we can use the global plane-parallel formalism to compute the three-point functions that are necessary to calculate the masked three-point function in equation (63).

6.3 Comparison with simulations

A direct way to estimate the survey geometry effect on the bispectrum measurements is to compare two types of simulations. The first one is the MD-Patchy mocks that reproduce the survey geometry of the BOSS galaxy sample. The second one is cubic box simulations that have the same galaxy description as the MD-Patchy mock catalogues with a volume of $(2.5 h^{-1} \text{Mpc})^3$. These box simulations provide the bispectrum measurement under the periodic boundary condition, without an artefact due to non-uniform survey footprint, and therefore, can be straightforwardly compared with the analytic bispectrum derived from perturbation theories. The MD-Patchy mocks include the light-cone evolution of galaxy clustering that are produced with different redshift snapshots of the box simulation. As an approximation, we ignore the redshift dependence and compare one snapshot at $z = 0.576$ with the MD-Patchy mocks. To validate this approximation, we also measured the bispectra at three redshifts, $z = 0.466$, $z = 0.505$, and $z = 0.533$; we find that there is no significant difference among these snapshots, and hence, the evolution of galaxy clustering does not produce a large change of the shape of the bispectrum multiples in the redshift range $0.466 < z < 0.576$. We expect from this result that the difference between the box simulations and the MD-Patchy mocks mainly arises from the survey geometry effect. We note that we do not take account of the discreteness effect due to binning which may introduce systematic bias in the measured bispectrum at low- k at a few percent level. We defer a more detail analysis investigating this effect for future works.

In the box simulation, we apply the global-plane parallel approximation and choose the z -axis as the LOS direction. Then, we measure the bispectrum multipoles via the estimator given by equation (52).

quantities themselves. This prescription has been developed by Slepian & Eisenstein (2015) and Slepian & Eisenstein (2018). If we focus only on the $L = 0$ mode, equation (63) divided by Q_{000} reduces to equation (32) in Slepian & Eisenstein (2015).

Fig. 8 shows four diagonal bispectrum multipoles: B_{000} (upper left), B_{110} (bottom left), B_{202} (upper right), and B_{112} (bottom right). The black dashed lines denote the means of the measured bispectra in 200 realizations of the box simulation at redshift $z = 0.576$. The black solid lines with error bars in Fig. 8 are the results from the MD-Patchy mocks, which are the same as shown in Fig. 3. We find clear differences between the box simulations and the MD-Patchy mocks at scales larger than $k = 0.01 h \text{Mpc}^{-1}$ for B_{000} , $k = 0.1 h \text{Mpc}^{-1}$ for B_{110} , $k = 0.02 h \text{Mpc}^{-1}$ for B_{202} , and $k = 0.03 h \text{Mpc}^{-1}$ for B_{112} . For the $L = 0$ mode (B_{000} and B_{110}), the MD-Patchy mocks including the survey window effect tend to be in agreement with the box mocks that provide the true measurements on small scales. On the other hand, there is a small offset on small scales ($k \sim 0.2 h \text{Mpc}^{-1}$) in the $L = 2$ mode (B_{202} and B_{112}) between the two mocks, which could be due to the velocity calibrations made to the MD-Patchy mocks to match their quadrupole power spectrum to the measured quadrupole of BOSS DR12 (Kitaura et al. 2016).

The main purpose of this subsection is to show that our window function treatment with equation (68) can indeed model the difference between the box simulations and the MD-Patchy mocks. To construct the bispectrum multipole models with the window function effect, we model the true bispectrum multipoles using second-order perturbation theory, the so-called tree-level solution (for details, see Appendix F). We compute the tree-level bispectra with the fiducial cosmological parameters mentioned in Section 5.1 at $z = 0.576$, where this parameter set corresponds to the box simulation shown in Fig. 8. In this model, only two nuisance parameters are necessary to describe the clustering of a sample of galaxies: b_1 and b_2 , the first- and second-order Eulerian bias parameters. We determine these bias parameters by performing a standard minimum χ^2 analysis by comparing the four bispectrum multipoles, B_{000} , B_{110} , B_{202} , and B_{112} , predicted by the theory with those measured from the box simulation resulting in the best-fitting values: $b_1 = 1.69$ and $b_2 = 1.89$. Notice that, since we adopt this model just for the purpose of reproducing the box simulation results at large scales, we do not interpret values of b_1 and b_2 as the physical galaxy bias parameters. We compute the three-point function multipoles $\zeta_{\ell_1 \ell_2 L}(r_1, r_2)$ through the Hankel transform given in equation (17).

We adopt the simplest expression of the masked three-point function that only considers the first term in equation (65):

$$\left\langle \widehat{\zeta}_{\ell_1 \ell_2 L}(r_1, r_2) \right\rangle_{\text{model}} = Q_{000}(r_1, r_2) \zeta_{\ell_1 \ell_2 L}(r_1, r_2). \quad (69)$$

This approximation is adequate in the BOSS analysis because we have checked in Appendix F that the other additional corrections (the second line of equation 65) produces a negligibly small change in the predictions of the bispectrum multipoles, $\lesssim 30$ per cent of the 1σ errors estimated from the MD-Patchy mocks.

The red-solid and red-dashed lines in Fig. 8 show the tree-level solutions with and without the window corrections, respectively. The tree-level solution including our survey window corrections can model well the observed survey geometry effect. In particular, the perturbation theory approach explains a characteristic feature that the survey window effect on B_{110} appears even up to $k = 0.1 h \text{Mpc}^{-1}$, which is consistent with the result from the two mocks.

While the survey geometry effect on the bispectrum becomes smaller on small scales as shown in Fig. 8, it is important to estimate the impact of the effect on clustering analysis more quantitatively. To do so, we compute the ratios of the differences between the bispectrum multipoles with and without the window effect, $B_{\ell_1 \ell_2 L} - B_{\ell_1 \ell_2 L}^{\text{eff}}$, to the 1σ error estimated from the MD-

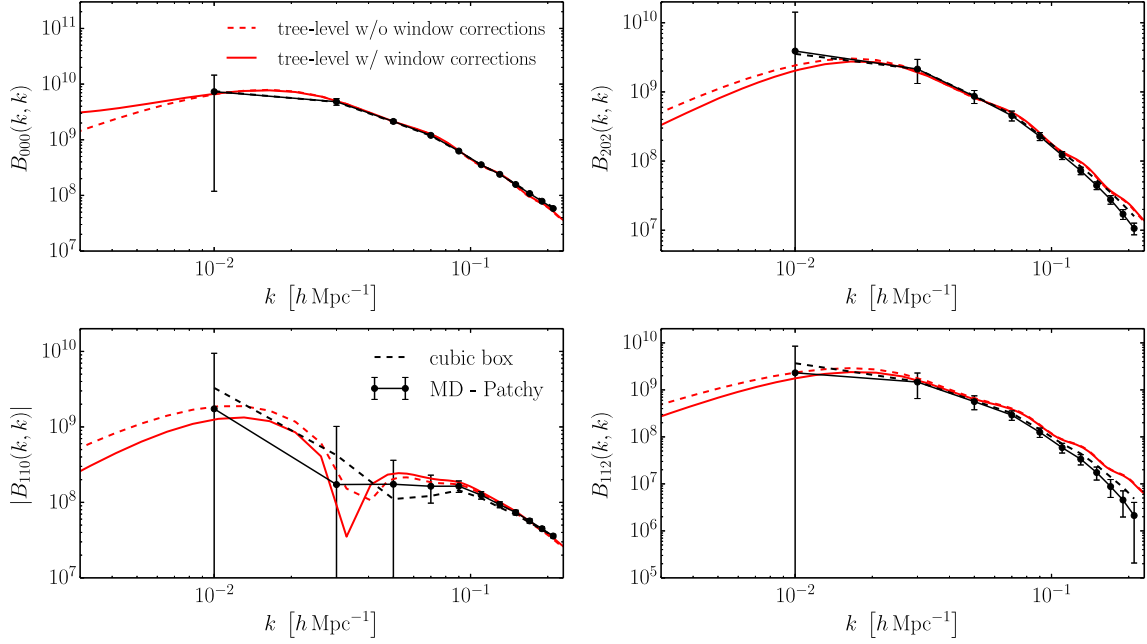


Figure 8. Diagonal bispectrum multipoles with (solid lines) and without (dashed lines) the survey geometry effect, for B_{000} (upper left), B_{110} (bottom left), B_{202} (upper right), and B_{112} (bottom right). In the context of simulations, we compare the measurements from the MD-Patchy mocks (black solid lines with error bars) with the mean values in the 200 cubic box simulations at $z = 0.576$ that has the same galaxy distribution as the MD-Patchy mocks (black dashed lines). In the framework of perturbation theories, we take into account the survey geometry effect using equation (68), and plot the tree-level solutions with (red solid) and without (red dashed) the survey geometry effect. This figure shows that the survey geometry effect is noticeable mainly on scales larger than $k = 0.01 h \text{ Mpc}^{-1}$ for B_{000} , $k = 0.1 h \text{ Mpc}^{-1}$ for B_{110} , $k = 0.02 h \text{ Mpc}^{-1}$ for B_{202} , and $k = 0.03 h \text{ Mpc}^{-1}$ for B_{112} , and will be smaller on small scales, even though Fig. 9 will show that the effective difference compared to the 1σ error can be still substantial even on small scales.

Patchy mocks, $\Delta B_{\ell_1 \ell_2 L}$, in Fig. 9. Here, the reference bispectrum is the one without the window effect, and we consider four diagonal bispectrum multipoles: B_{000} (blue), B_{110} (green), B_{202} (red), and B_{112} (cyan). The ratios shown in this figure increase at smaller scales, because the 1σ errors decrease with increasing the number of modes, i.e. with decreasing scales, faster than the window effect is reduced. Remarkably, the ratios for B_{000} and B_{110} reach ~ 100 per cent and ~ 60 per cent at $k = 0.2 h \text{ Mpc}^{-1}$, respectively. This fact indicates that we cannot ignore the window effect in clustering analysis even on small scales.

Fig. 8 also clarifies the scales where the tree-level approximation breaks down. For B_{000} and B_{202} , the tree-level solutions work well up to $k \sim 0.1 h \text{ Mpc}^{-1}$ at $z \sim 0.5$ (compare the red dashed lines with the black dashed lines), while for B_{110} and B_{112} , they cannot explain the simulation results at even larger scales. These results indicate that to model higher bispectrum multipoles or to use small-scale information on the bispectrum, we need to go beyond the tree-level approximation. We leave improvements on the modeling of the bispectrum for future work (e.g. Hashimoto et al. 2017; Yamamoto, Nan & Hikage 2017; Nan, Yamamoto & Hikage 2018).

7 CONCLUSIONS

This work aims to efficiently extract information of the RSD and AP effects from the three-point statistics to improve cosmological parameters constraints.

We have presented a new type of decomposition of the three-point statistics into the TripoSH basis (equation 6) with zero total angular momentum, which makes use of the fact that density fluctuations in the Universe are expect to satisfy homogeneity, isotropy, and parity

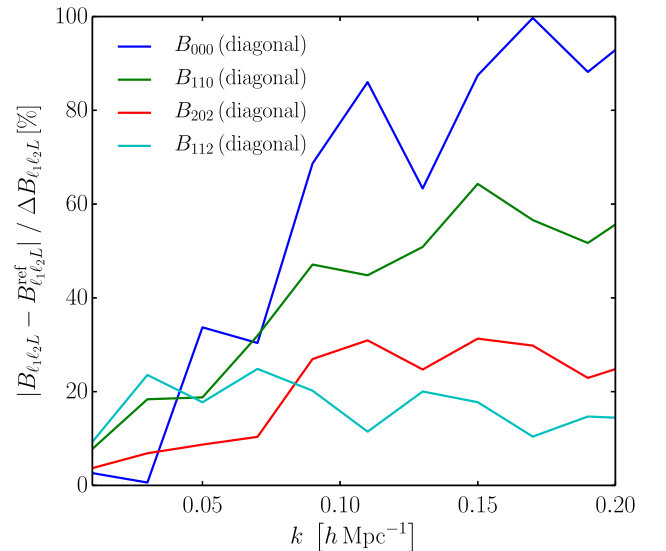


Figure 9. Ratios of the difference between the bispectrum multipoles with and without the survey window corrections, $B_{\ell_1 \ell_2 L} - B_{\ell_1 \ell_2 L}^{\text{ref}}$ to the 1σ errors estimated from the MD-Patchy mocks, $\Delta B_{\ell_1 \ell_2 L}$. In this figure, we compute the bispectrum multipoles using the tree-level solution, and the reference bispectrum multipoles, $B_{\ell_1 \ell_2 L}^{\text{ref}}$, are the true values without including the survey window effect. This figure shows that although the absolute difference shown in Fig. 8 is small on small scales, the effective difference compared to the 1σ error reaches 50 – 100 per cent for B_{000} and B_{110} , indicating that the window corrections cannot be ignored in clustering analysis even on small scales.

symmetry. In the relevant expansion coefficients for the bispectrum $B_{\ell_1\ell_2L}(k_1, k_2)$ (equation 11), the L mode is an analogue of the ℓ mode of the power spectrum multipoles, P_ℓ , expanded in Legendre polynomials with respect to the LOS, and hence, the $L \neq 0$ modes can never be generated unless in the presence of the RSD or AP effects. The corresponding coefficients of the three-point function $\zeta_{\ell_1\ell_2L}(r_1, r_2)$ can then be computed through a two-dimensional Hankel transform (equation 17).

We have presented the bispectrum and three-point function estimators within our new decomposition formalism both for the global plane-parallel approximation (i.e. for idealized periodic-box simulations, equation 52) and for the local plane-parallel approximation (i.e. for observed galaxy data, equations 41 and 48). The resulting estimators can be computed using FFTs, which allows a complexity of $\mathcal{O}(N \log N)$, where N represents the number of grid cells used in the FFT.

Using our estimators, we have measured the bispectrum multipoles from the CMASS NGC sample drawn from BOSS DR12 and the associated MD-Patchy mocks. We find an excellent agreement between the results from the mocks and the galaxy sample on large scales ($k \gtrsim 0.1 h \text{ Mpc}^{-1}$). On smaller scales ($k \lesssim 0.1 h \text{ Mpc}^{-1}$), they start to depart from each other beyond the 1σ error.

We have investigated the covariance matrices of the bispectrum multipoles (Fig. 4), calculated from the MD-Patchy mocks. Since the bispectrum multipoles are characterized by two scales, k_1 and k_2 , their covariance matrices depend on four variables: k_1, k_2, k'_1 , and k'_2 . As far as we see the $k_1 = k'_1$ (or $k_1 = k'_2$) case, there are strong correlations between the bispectrum multipoles, even if $k_2 \neq k'_2$ (or $k_2 \neq k'_1$). Remarkably, there are weak correlations between the diagonal elements of the bispectrum multipoles estimated at two different scales, i.e. between $B_{\ell_1\ell_2L}(k, k)$ and $B_{\ell_1\ell_2L}(k', k')$, indicating that they include nearly independent information at different scales.

We have estimated the cumulative S/Ns of the amplitude of density fluctuations from the CMASS NGC sample (Fig. 5). In the analysis only considering the power spectrum monopole, there is the well-known non-linear degradation of the S/N by ~ 80 per cent at $k_{\text{max}} = 0.1 h \text{ Mpc}^{-1}$ and ~ 60 per cent at $k_{\text{max}} = 0.2 h \text{ Mpc}^{-1}$ compared to the Gaussian prediction due to the leakage of the information to higher order statistics. The joint analysis of the power spectrum and bispectrum multipoles recovers the S/N up to 95 per cent at $k = 0.1 h \text{ Mpc}^{-1}$ and 75 per cent at $k = 0.2 h \text{ Mpc}^{-1}$. We find that the joint analysis using the diagonal bispectrum monopole well recovers the S/N up to 90 per cent at $k = 0.1 h \text{ Mpc}^{-1}$ and 70 per cent at $k = 0.2 h \text{ Mpc}^{-1}$ because of their nearly independent information at different scales.

We have also computed the S/Ns of the diagonal bispectrum multipoles themselves (Fig. 6). For each of the $L = 0$ and 2 modes, the lowest order of the bispectrum multipoles, B_{000} ($L = 0$) or B_{202} ($L = 2$), yields the highest S/N. One of the most remarkable results of this paper is that, we, for the first time, detected the $L = 2$ mode, which is an analogue of the power spectrum quadrupole: we measure the lowest order $L = 2$ mode, B_{202} , with a significance of 14σ and also the next leading order, B_{112} , with a significance of 6σ .

The finding that most of the bispectrum information is contained in the diagonal terms could significantly simplify the bispectrum analysis. The measurements of the diagonal elements of a few bispectrum multipoles in both $L = 0$ and $L = 2$ modes would allow extracting most of the higher order information contained in the bispectrum.

We present a formalism for how to include the survey window function in the bispectrum model, before comparing it with the

measured bispectrum. This represents the first self-consistent treatment of the survey window function in a bispectrum analysis. The survey geometry effect appears mainly on large scales and gets close to zero at small scales (Fig. 8) as expected. Nevertheless, the survey geometry effect can significantly affect the bispectrum measurements even at small scales, because these survey geometry effects are comparable to the statistical uncertainties, which is also reduced towards small scales (Fig. 9). This fact emphasizes the need to include the survey geometry effect in any statistical analysis.

This paper has demonstrated that the TripoSH formalism can be used to measure the bispectrum in large-scale structure data sets, which complements the now well-established power spectrum analysis. The formalism can be trivially extended to higher-order statistics, like the four-point function and the trispectrum, by use of poly-polar spherical harmonic (PolypoSH) decomposition. We hope that the formalism presented in this paper will become a standard method for analyzing higher-order statistics.

ACKNOWLEDGEMENTS

NSS acknowledges financial support from Grant-in-Aid for JSPS Fellows (No. 28-1890). Numerical computations were carried out on Cray XC50 at Center for Computational Astrophysics, National Astronomical Observatory of Japan. This work was supported in part by JSPS KAKENHI Grant Number JP15H05896. FB is a Royal Society University Research Fellow. H-JS is supported by the U.S. Department of Energy, Office of Science, Office of High Energy Physics under Award Number DE-SC0014329. NSS would like to thank Eiichiro Komatsu and Max Planck Institute for Astrophysics for their warm hospitality during the middle stage of this work. NSS and SS also thank the Yukawa Institute for Theoretical Physics at Kyoto University. Discussions during the YITP workshop YITP-T-17-03 on ‘CosKASI-ICG-NAOC-YITP’ were used to complete this work.

REFERENCES

- Ade P. A. R. et al., 2014, *A&A*, 571, A24
 Ade P. A. R. et al., 2016, *A&A*, 594, A17
 Alam S. et al., 2015, *Astrophys. J. Suppl.*, 219, 12
 Albrecht A., Steinhardt P. J., 1982, *Phys. Rev. Lett.*, 48, 1220
 Alcock C., Paczyński B., 1979, *Nature*, 281, 358
 Anderson L. et al., 2014, *MNRAS*, 441, 24
 Bernardeau F., Colombi S., Gaztanaga E., Scoccimarro R., 2002, *Phys. Rept.*, 367, 1
 Beutler F. et al., 2014, *MNRAS*, 443, 1065
 Beutler F. et al., 2017, *MNRAS*, 466, 2242
 Blot L., Corasaniti P. S., Alimi J.-M., Reverdy V., Rasera Y., 2015, *MNRAS*, 446, 1756
 Bolton A. S. et al., 2012, *Astron. J.*, 144, 144
 Bundy K. et al., 2015, *ApJS*, 221, 15
 Byun J., Eggemeier A., Regan D., Seery D., Smith R. E., 2017, *MNRAS*, 471, 1581
 Chan K. C., Blot L., 2017, *Phys. Rev.*, D96, 023528
 Dawson K. S. et al., 2013, *Astron. J.*, 145, 10
 Doi M. et al., 2010, *Astron. J.*, 139, 1628
 Eisenstein D. J. et al., 2011, *Astron. J.*, 142, 72
 Feldman H. A., Kaiser N., Peacock J. A., 1994, *ApJ*, 426, 23
 Feldman H. A., Frieman J. A., Fry J. N., Scoccimarro R., 2001, *Phys. Rev. Lett.*, 86, 1434
 Fry J. N., Seldner M., 1982, *ApJ*, 259, 474
 Fukugita M., Ichikawa T., Gunn J. E., Doi M., Shimasaku K., Schneider D. P., 1996, *Astron. J.*, 111, 1748

Gaztanaga E., Norberg P., Baugh C. M., Croton D. J., 2005, *MNRAS*, 364, 620

Gaztanaga E., Cabre A., Castander F., Crocce M., Fosalba P., 2009, *MNRAS*, 399, 801

Gil-Marln H. et al., 2015b, *MNRAS*, 452, 1914

Gil-Marln H., Norena J., Verde L., Percival W. J., Wagner C., Manera M., Schneider D. P., 2015a, *MNRAS*, 451, 539

Gil-Marln H., Percival W. J., Verde L., Brownstein J. R., Chuang C.-H., Kitaura F.-S., Rodriguez-Torres S. A., Olmstead M. D., 2017, *MNRAS*, 465, 1757

Groth E. J., Peebles P. J. E., 1977, *ApJ*, 217, 385

Gunn J. E. et al., 1998, *Astron. J.*, 116, 3040

Gunn J. E. et al., 2006, *Astron. J.*, 131, 2332

Guo H. et al., 2015, *MNRAS*, 446, 578

Guo H., Li C., Jing Y. P., Borner G., 2014, *ApJ*, 780, 139

Guth A. H., 1981, *Phys. Rev.*, D23, 347

Hamilton A. J. S., 1997, *Astrophysics and Space Science Library*, Vol. 231, The Evolving Universe. Selected Topics on Large-Scale Structure and on the Properties of Galaxies. Kluwer Academic Publishers, Dordrecht, p. 185

Hand N., Feng Y., Beutler F., Li Y., Modi C., Seljak U., Slepian Z., 2018, *AJ*, 156, 160

Hand N., Li Y., Slepian Z., Seljak U., 2017, *JCAP*, 1707, 002

Hartlap J., Simon P., Schneider P., 2007, *A&A*, 464, 399

Hashimoto I., Rasera Y., Taruya A., 2017, *Phys. Rev.*, D96, 043526

Hockney R. W., Eastwood J. W., 1981, *Computer Simulation Using Particles*. Taylor & Francis, Inc., Bristol, PA

Jing Y. P., 2005, *ApJ*, 620, 559

Jing Y. P., Boerner G., 2004, *ApJ*, 607, 140

Kayo I. et al., 2004, *PASJ*, 56, 415

Kitaura F.-S. et al., 2016, *MNRAS*, 456, 4156

Klypin A., Yepes G., Gottlober S., Prada F., Hess S., 2016, *MNRAS*, 457, 4340

Kulkarni G. V., Nichol R. C., Sheth R. K., Seo H.-J., Eisenstein D. J., Gray A., 2007, *MNRAS*, 378, 1196

Leauthaud A. et al., 2016, *MNRAS*, 457, 4021

Lesgourgues J., 2011, preprint ([arXiv:1104.2932](https://arxiv.org/abs/1104.2932))

Linde A. D., 1982, *Phys. Lett.*, B108, 389

Marin F., 2011, *ApJ*, 737, 97

Marin F. A. et al., 2013, *MNRAS*, 432, 2654

McBride C. K., Connolly A. J., Gardner J. P., Scranton R., Newman J. A., Scoccimarro R., Zehavi I., Schneider D. P., 2011a, *ApJ*, 726, 13

McBride C. K., Connolly A. J., Gardner J. P., Scranton R., Scoccimarro R., Berlind A. A., Marin F., Schneider D. P., 2011b, *ApJ*, 739, 85

Nan Y., Yamamoto K., Hikage C., 2018, *JCAP*, 1807, 038

Nichol R. C. et al., 2006, *MNRAS*, 368, 1507

Pan J., Szapudi I., 2005, *MNRAS*, 362, 1363

Parejko J. K. et al., 2013, *MNRAS*, 429, 98

Peacock J. A., Nicholson D., 1991, *MNRAS*, 253, 307

Pearson D. W., Samushia L., 2018, *MNRAS*, 478, 4500

Peebles P. J. E., 1980, *The Large-Scale Structure of the Universe*. Princeton Univ. Press, Princeton, USA

Peebles P. J. E., Groth E. J., 1975, *ApJ*, 196, 1

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

Planck Collaboration et al., 2016a, *A&A*, 594, A16

Planck Collaboration et al., 2016b, *A&A*, 594, A20

Pullen A. R., Hirata C. M., 2010, *J. Cosmol. Astropart. Phys.*, 5, 027

Reid B. et al., 2016, *MNRAS*, 455, 1553

Rimes C. D., Hamilton A. J. S., 2005, *MNRAS*, 360, L82

Rimes C. D., Hamilton A. J. S., 2006, *MNRAS*, 371, 1205

Ross A. J. et al., 2012, *MNRAS*, 424, 564

Saito S. et al., 2016, *MNRAS*, 460, 1457

Sato K., 1981, *MNRAS*, 195, 467

Scoccimarro R., 2000, *ApJ*, 544, 597

Scoccimarro R., 2015, *Phys. Rev.*, D92, 083532

Scoccimarro R., Couchman H. M. P., Frieman J. A., 1999, *ApJ*, 517, 531

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

Sefusatti E., Crocce M., Scoccimarro R., Couchman H., 2016, *MNRAS*, 460, 3624

Shiraishi M., Sugiyama N. S., Okumura T., 2017, *Phys. Rev.*, D95, 063508

Slepian Z. et al., 2017a, *MNRAS*, 468, 1070

Slepian Z. et al., 2017b, *MNRAS*, 469, 1738

Slepian Z., Eisenstein D. J., 2015, *MNRAS*, 454, 4142

Slepian Z., Eisenstein D. J., 2016, *MNRAS*, 455, L31

Slepian Z., Eisenstein D. J., 2018, *MNRAS*, 478, 1468

Smith J. A. et al., 2002, *Astron. J.*, 123, 2121

Starobinsky A. A., 1980, *Phys. Lett.*, B91, 99

Sugiyama N. S., Shiraishi M., Okumura T., 2018, *MNRAS*, 473, 2737

Szapudi I., 2004a, *ApJ*, 605, L89

Szapudi I., 2004b, *ApJ*, 614, 51

Takahashi R. et al., 2009, *ApJ*, 700, 479

Varshalovich D. A., Moskalev A. N., Khersonskii V. K., 1988, *Quantum Theory of Angular Momentum*. World Scientific Publishing Co, NJ, USA

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

Wang Y., Yang X.-H., Mo H. J., van den Bosch F. C., Chu Y.-Q., 2004, *MNRAS*, 353, 287

White M. et al., 2011, *ApJ*, 728, 126

Wilson M. J., Peacock J. A., Taylor A. N., de la Torre S., 2017, *MNRAS*, 464, 3121

Yamamoto K., Nan Y., Hikage C., 2017, *Phys. Rev.*, D95, 043528

APPENDIX A: USEFUL IDENTITIES

Here we summarize the identities used for the derivations in this paper.

Throughout this paper, we use a normalized spherical harmonics, y_ℓ^m , defined as

$$y_\ell^m(\hat{x}) \equiv \sqrt{\frac{4\pi}{2\ell+1}} Y_\ell^m(\hat{x}), \quad (\text{A1})$$

where its leading order becomes unity: $y_0^0 = 1$.

The relation between Legendre polynomials and spherical harmonics is given by

$$\mathcal{L}_\ell(\hat{x} \cdot \hat{y}) = \sum_m y_\ell^m(\hat{x}) y_\ell^{m*}(\hat{y}). \quad (\text{A2})$$

The relation between associated Legendre polynomials and spherical harmonics is

$$y_\ell^m(\theta, \varphi) = (-1)^{(m-|m|)/2} \sqrt{\frac{(\ell-|m|)!}{(\ell+|m|)!}} \mathcal{L}_\ell^{|m|}(\cos\theta) e^{im\varphi}. \quad (\text{A3})$$

The addition rule of spherical harmonics:

$$y_{\ell_1}^{m_1}(\hat{x}) y_{\ell_2}^{m_2}(\hat{x}) = \sum_{\ell_3 m_3} (2\ell_3 + 1) \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ m_1 & m_2 & m_3 \end{pmatrix} y_{\ell_3}^{m_3*}(\hat{x}). \quad (\text{A4})$$

Five Wigner 3- j symbols can generate the Wigner 9- j symbol as follows:

$$\begin{aligned} & \sum_{m_4 m_5 m_6 m_7 m_8 m_9} (-1)^{\sum_{i=4}^9 (\ell_i - m_i)} \\ & \times \begin{pmatrix} \ell_4 & \ell_1 & \ell_7 \\ m_4 & -m_1 & m_7 \end{pmatrix} \begin{pmatrix} \ell_5 & \ell_4 & \ell_6 \\ -m_5 & -m_4 & -m_6 \end{pmatrix} \\ & \times \begin{pmatrix} \ell_6 & \ell_3 & \ell_9 \\ m_6 & -m_3 & m_9 \end{pmatrix} \begin{pmatrix} \ell_7 & \ell_9 & \ell_8 \\ -m_7 & -m_9 & -m_8 \end{pmatrix} \begin{pmatrix} \ell_8 & \ell_2 & \ell_5 \\ m_8 & -m_2 & m_5 \end{pmatrix} \\ & = (-1)^{\ell_2 + \ell_5 + \ell_8} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \left\{ \begin{matrix} \ell_1 & \ell_2 & \ell_3 \\ \ell_4 & \ell_5 & \ell_6 \\ \ell_7 & \ell_8 & \ell_9 \end{matrix} \right\}. \quad (\text{A5}) \end{aligned}$$

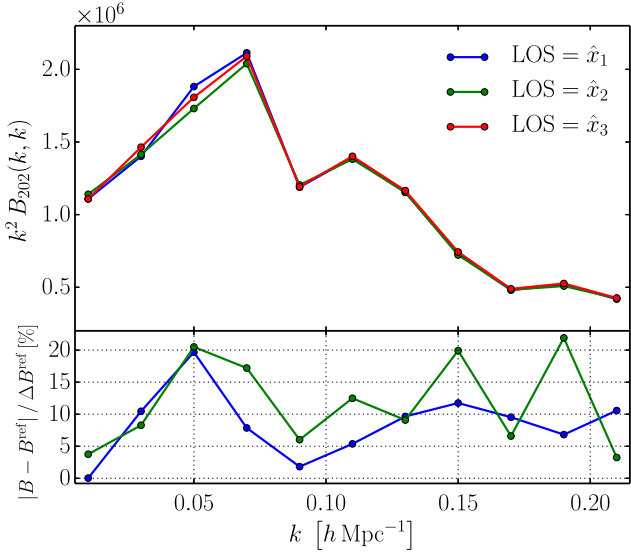


Figure B1. A comparison of the bispectrum measurements using three different LOS directions from CMASS NGC. In the top part, we measure the lowest order of the $L = 2$ mode of the bispectrum multipoles, $B_{202}(k, k)$, by choosing \hat{x}_1 , \hat{x}_2 , and \hat{x}_3 as the LOS direction. In the bottom part, we show the ratios of the differences between the measured bispectra and a reference bispectrum to the 1σ error, ΔB^{ref} , on the reference one. This figure shows that the dependence of the choice of the LOS directions on the bispectrum measurement is negligibly small in the BOSS analysis, $\lesssim 20$ per cent of ΔB^{ref} .

APPENDIX B: CHOICE OF LINE-OF-SIGHT DIRECTIONS

In the main text, when measuring the bispectrum multipoles from observed galaxies, we apply the local plane-parallel approximation, $\hat{x}_1 \approx \hat{x}_2 \approx \hat{x}_3$, and choose \hat{x}_3 as the LOS direction. However, as the approximation is not exactly satisfied in the bispectrum measurement, the $L \geq 2$ modes of the measured bispectrum multipoles may depend on the choice of the LOS direction: \hat{x}_1 , \hat{x}_2 , or \hat{x}_3 , while the $L = 0$ mode does not.

This point is illustrated in Fig. B1, which shows the lowest order of the $L = 2$ mode, $B_{202}(k, k)$, measured using three different LOS directions, \hat{x}_1 (blue), \hat{x}_2 (green), and \hat{x}_3 (red), from the CMASS NGC sample. While the top part plots the bispectrum measurements, the bottom shows the ratios of the differences between the measured bispectra and a reference bispectrum B^{ref} to the 1σ error ΔB^{ref} on the reference one. Here, the reference bispectrum is computed using \hat{x}_3 as the LOS direction (Section 4), and its 1σ error is estimated from the 600 MD-Patchy mocks (Section 5.5). We find from this figure that the dependence of the choice of the LOS directions on the bispectrum measurement is less than about 20 per cent of the 1σ error, and hence that it is negligible in the BOSS analysis. We expect that the difference among the LOS directions, to some extent, can be explained by the survey window corrections (Section 6), because the survey window function should also depend on the choice of the LOS directions.

APPENDIX C: MASS ASSIGNMENT SCHEMES

Measuring a density field from a galaxy sample requires the interpolation of functions on a regular grid in position space. In this section, we provide a comparison of several interpolation schemes, the nearest grid point (NGP), cloud-in-cell (CIC), and triangular-

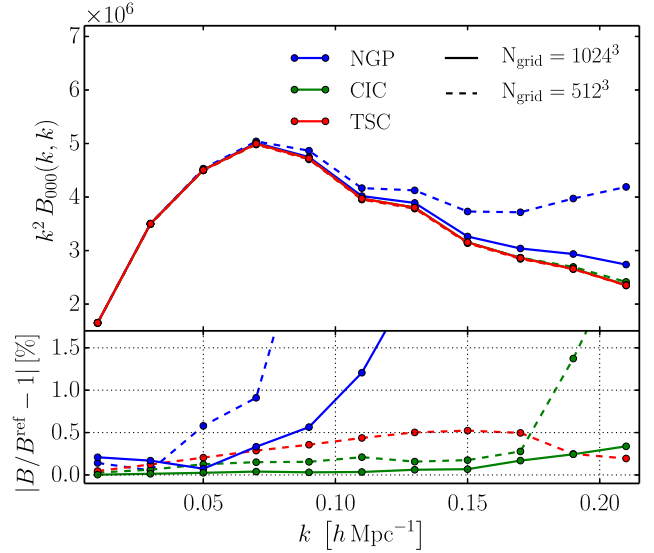


Figure C1. A comparison of the effect of the mass assignment function on the measurement of the diagonal bispectrum monopole, $B_{000}(k, k)$. In the top part, we measure $B_{000}(k, k)$ using the NGP, CIC, and TSC schemes with two different numbers of grid cells, $N_{\text{grid}} = 512^3$ and $N_{\text{grid}} = 1024^3$, from the CMASS NGC sample. The CMASS NGC galaxies are placed in a cubic box with a volume of $(3.5 h^{-1} \text{Gpc})^3$, and the Nyquist frequency is then $k_N = 0.46 h \text{Mpc}^{-1}$ for $N_{\text{grid}} = 512^3$ and $k_N = 0.92 h \text{Mpc}^{-1}$ for $N_{\text{grid}} = 1024^3$. In the bottom part, we show the ratios of the measured bispectrum monopoles to a reference one B^{ref} , computed using the TSC scheme with a 1024^3 grids. This figure shows that the TSC scheme achieves sub-per cent accuracy, $\lesssim 0.5$ per cent, up to the scale of interest, $k \lesssim 0.21 h \text{Mpc}^{-1}$.

shaped cloud (TSC) assignment schemes (Hockney & Eastwood 1981), with two different numbers of grid-cells, $N_{\text{grid}} = 512^3$ and $N_{\text{grid}} = 1024^3$. We focus especially on the diagonal bispectrum monopole, $B_{000}(k, k)$, and measure it from the CMASS NGC galaxy sample by placing the CMASS NGC galaxies in a cubic box with a volume of $(3.5 h^{-1} \text{Gpc})^3$. The Nyquist frequency then becomes $k_N = 0.46 h \text{Mpc}^{-1}$ for $N_{\text{grid}} = 512^3$ and $k_N = 0.92 h \text{Mpc}^{-1}$ for $N_{\text{grid}} = 1024^3$.

Fig. C1 illustrates the difference among the NGP (blue), CIC (green), and TSC (red) schemes with $N_{\text{grid}} = 512^3$ (dashed) and $N_{\text{grid}} = 1024^3$ (solid). The top part plots the measurements using these schemes, and the bottom shows the ratios of the measured bispectrum monopoles to a reference bispectrum monopole, where the reference one is computed using the TSC assignment function with a 1024^3 grid. The choice of assignment functions significantly affects the bispectrum measurement, compared to the power spectrum case (e.g. see fig. 2 in Hand et al. 2018). In particular, the NGP scheme has a bad convergence property when increasing N_{grid} . As expected, higher-order interpolation schemes perform better. For the TSC scheme, doubling the number of grids per side, from 512 to 1024, produces a small change in the bispectrum measurement, within 0.5 per cent, in a k -range of $0.01 h \text{Mpc}^{-1} < k < 0.21 h \text{Mpc}^{-1}$.

APPENDIX D: FKP WEIGHTING

While the FKP weight function $w_{\text{FKP}}(\mathbf{x}) = 1/[1 + \bar{n}(\mathbf{x})P_0]$ was originally obtained by minimizing the fractional variance in the power spectrum under the assumption that density fluctuations are Gaussian. Feldman et al. (1994) and Scoccimarro (2000) showed that assuming Gaussian errors, the same weight function minimizes

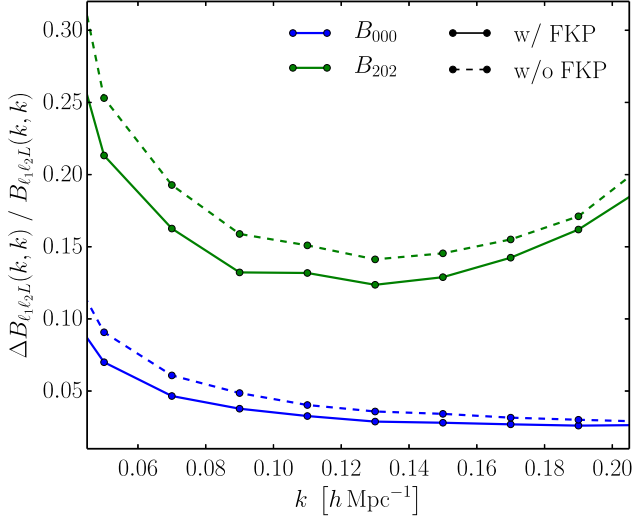


Figure D1. A comparison of the bispectrum measurements with (solid) and without (dashed) the FKP weighting from the CMASS NGC mocks, where both the lowest orders of the $L = 0$ and $L = 2$ modes, B_{000} (blue) and B_{202} (green), are computed. We plot the ratios of the 1σ error on the bispectrum measurements to the mean values of them. This figure shows that the FKP weighting decreases the errors by ~ 15 per cent in the k -range $0.01 h \text{ Mpc}^{-1} < k < 0.2 h \text{ Mpc}^{-1}$ in the BOSS analysis.

the variance of the higher-order correlation functions as well. To verify this in our decomposition formalism (Section 2), we compare two measurements of the bispectrum multipoles with and without the FKP weighting, with $P_0 = 10^4 h^{-3} \text{ Mpc}^3$. We use the 600 MD-Patchy mocks for CMASS NGC to estimate the mean value of the bispectrum measurements and their 1σ errors.

Fig. D1 provides a comparison of the bispectrum measurements with (solid) and without (dashed) the FKP weighting, where we measure the lowest orders of the $L = 0$ and $L = 2$ modes, respectively: B_{000} (blue) and B_{202} (green), and we plots the ratios of the 1σ error on the bispectrum measurements to the mean values of them. We find from this figure that the FKP weighting, as expected, decreases the errors by ~ 15 per cent in the k -range $0.01 h \text{ Mpc}^{-1} < k < 0.2 h \text{ Mpc}^{-1}$ in the BOSS analysis.

APPENDIX E: THE HARTLAP FACTOR AND THE PERCIVAL FACTOR

In this appendix, we describe the Hartlap factor and the Percival factor.

Measuring the standard χ^2 statistics in Section 5.6, it is necessary to compute the inverse of the covariance matrix. As the estimated covariance matrix C in equation (55) is inferred from a set of mocks, its inverse C^{-1} is biased due to the limited number of realizations. We account for this effect by rescaling the inverse covariance matrix as follows (Hartlap et al. 2007):

$$C_{\text{Hartlap}}^{-1} = \frac{N_{\text{mock}} - N_{\text{bin}} - 2}{N_{\text{mock}} - 1} C^{-1}, \quad (\text{E1})$$

where N_{bin} is the number of bins, and $(N_{\text{mock}} - N_{\text{bin}} - 2)/(N_{\text{mock}} - 1)$ is the so-called Hartlap factor. The standard χ^2 value is then given by

$$\chi^2 = (\mathbf{X}^{\text{obs}} - \mathbf{X}^{\text{theory}})^T C_{\text{Hartlap}}^{-1} (\mathbf{X}^{\text{obs}} - \mathbf{X}^{\text{theory}}), \quad (\text{E2})$$

where \mathbf{X}^{obs} is the vector of the observed quantities and $\mathbf{X}^{\text{theory}}$ is the vector of the corresponding theoretical predictions.

In addition to the Hartlap factor, we propagate the error in the covariance matrix to the error on parameters by scaling their variances by (equation 18 in Percival et al. 2014):

$$M = \frac{1 + B(N_{\text{bin}} - N_p)}{1 + A + B(N_p + 1)} \quad (\text{E3})$$

with

$$A = \frac{2}{\sqrt{(N_{\text{mock}} - N_{\text{bin}} - 1)(N_{\text{mock}} - N_{\text{bin}} - 4)}} \quad (\text{E4})$$

$$B = \frac{N_{\text{mock}} - N_{\text{bin}} - 2}{\sqrt{(N_{\text{mock}} - N_{\text{bin}} - 1)(N_{\text{mock}} - N_{\text{bin}} - 4)}},$$

where N_p is the number of parameters.

APPENDIX F: WINDOW CORRECTIONS

We here provide a complementary study of the survey geometry effect on the bispectrum multipoles in the framework of perturbation theory.

In the second-order perturbation theory, the first- and second-order kernels are (Scoccimarro et al. 1999)

$$Z_1(\mathbf{k}_1) = (b_1 + f\mu_1^2)$$

$$Z_2(\mathbf{k}_1, \mathbf{k}_2) = b_1 F_1(\mathbf{k}_1, \mathbf{k}_2) + \frac{b_2}{2} + f\mu^2 G_2(\mathbf{k}_2, \mathbf{k}_2) + \frac{f\mu k}{2} \left[\frac{\mu_1}{k_1} Z_1(\mathbf{k}_2) + \frac{\mu_2}{k_2} Z_1(\mathbf{k}_1) \right] \quad (\text{F1})$$

where $\mu = \hat{k} \cdot \hat{n}$ with $\mathbf{k} = \mathbf{k}_1 + \mathbf{k}_2$, $\mu_i = \hat{k}_i \cdot \hat{n}$ for $i = 1, 2$, and

$$F_2(\mathbf{k}_1, \mathbf{k}_2) = \frac{5}{7} + \frac{\mu_{12}}{2} \left(\frac{k_1}{k_2} + \frac{k_2}{k_1} \right) + \frac{2}{7} \mu_{12}^2$$

$$G_2(\mathbf{k}_1, \mathbf{k}_2) = \frac{3}{7} + \frac{\mu_{12}}{2} \left(\frac{k_1}{k_2} + \frac{k_2}{k_1} \right) + \frac{4}{7} \mu_{12}^2 \quad (\text{F2})$$

with $\mu_{12} = \hat{k}_1 \cdot \hat{k}_2$. The bispectrum is then given by

$$B(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \hat{n}) = 2Z_2(\mathbf{k}_1, \mathbf{k}_2)Z_1(\mathbf{k}_1)Z_1(\mathbf{k}_2)P_{\text{lin}}(k_1)P_{\text{lin}}(k_2) + \text{cyc.}, \quad (\text{F3})$$

where P_{lin} is the linear matter power spectrum, which is generated with CLASS (Lesgourgues 2011) in this paper.

Computing the masked three-point function multipoles (equation 63), in this appendix, we decide to use four three-point function multipoles and four window function multipoles: ζ_{000} , ζ_{110} , ζ_{202} , and ζ_{112} ; Q_{000} , Q_{110} , Q_{202} , and Q_{112} . We ignore the terms expressed by a product of the $L = 2$ modes of the three-point function and the window function, e.g. $\zeta_{202}Q_{202}$, because we found them to be negligible. Furthermore, we approximate the integral constraint (equation 64) as follows:

$$\bar{\zeta} \approx \frac{\int dr_1 r_1^2 \int dr_2 r_2^2 Q_{000}(r_1, r_2) \zeta_{000}(r_1, r_2)}{\int dr_1 r_1^2 \int dr_2 r_2^2 Q_{000}(r_1, r_2)}. \quad (\text{F4})$$

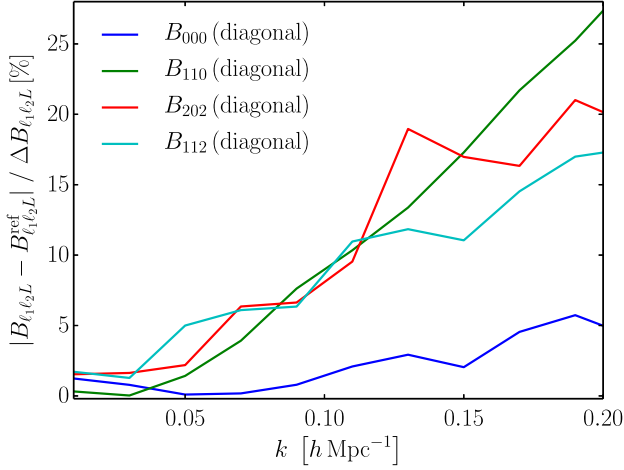


Figure F1. Ratios of the difference between the masked bispectrum multipoles computed using equations (69) and (F5) to the 1σ errors estimated from the MD-Patchy mocks, where the reference bispectrum multipoles, $B_{\ell_1 \ell_2 L}^{\text{ref}}$, are computed from equation (69). This figure shows that the difference is less than ~ 30 per cent of the 1σ error, and hence, additional correction terms given by equation (F5) to the window function effect are negligible in the BOSS analysis.

The masked three-point function multipoles are then given by

$$\begin{aligned}
 \langle \widehat{\zeta}_{000} \rangle_{\text{model}} &= Q_{000} [\zeta_{000} - \bar{\zeta}] + \frac{1}{3} Q_{110} \zeta_{110} \\
 \langle \widehat{\zeta}_{110} \rangle_{\text{model}} &= Q_{000} \zeta_{110} + Q_{110} [\zeta_{000} - \bar{\zeta}] \\
 \langle \widehat{\zeta}_{202} \rangle_{\text{model}} &= Q_{000} \zeta_{202} + \frac{1}{3} Q_{110} \zeta_{112} \\
 &\quad + \frac{1}{3} Q_{112} \zeta_{110} + Q_{202} [\zeta_{000} - \bar{\zeta}] \\
 \langle \widehat{\zeta}_{112} \rangle_{\text{model}} &= Q_{000} \zeta_{112} + \frac{2}{5} Q_{110} [\zeta_{022} + \zeta_{202}] \\
 &\quad + Q_{112} [\zeta_{000} - \bar{\zeta}] + \frac{2}{5} [Q_{022} + Q_{202}] \zeta_{110}, \quad (\text{F5})
 \end{aligned}$$

where, for brevity, we abbreviated the dependence of two comoving distances, r_1 and r_2 , on $\zeta_{\ell_1 \ell_2 L}(r_1, r_2)$ and $Q_{\ell_1 \ell_2 L}(r_1, r_2)$. By substituting these expressions into equation (17), we can obtain the masked bispectrum multipoles. The purpose of this appendix is to compare the results from the above masked three-point function with the simplest approximation, $\langle \widehat{\zeta}_{\ell_1 \ell_2 L} \rangle_{\text{model}} = Q_{000} \zeta_{\ell_1 \ell_2 L}$ (equation 69), used in the main text.

Fig. F1 shows the ratios of the difference between equations (69) and (F5), $|B_{\ell_1 \ell_2 L} - B_{\ell_1 \ell_2 L}^{\text{ref}}|$, to the 1σ errors estimated from the MD-Patchy mocks, $\Delta B_{\ell_1 \ell_2 L}$, where the reference bispectrum multipoles are computed from equation (69). Both of the numerator $|B_{\ell_1 \ell_2 L} - B_{\ell_1 \ell_2 L}^{\text{ref}}|$ and the denominator $\Delta B_{\ell_1 \ell_2 L}$ decrease toward smaller scales, but $\Delta B_{\ell_1 \ell_2 L}$ reduces faster than $|B_{\ell_1 \ell_2 L} - B_{\ell_1 \ell_2 L}^{\text{ref}}|$. Consequently, the ratio between them becomes an increasing function. We find that the difference is less than ~ 30 per cent of the 1σ error, indicating that the simplest approximation (equations 69) used in the main text is adequate for the BOSS analysis.

This paper has been typeset from a $\text{\TeX}/\text{\LaTeX}$ file prepared by the author.