Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **498,** 5299–5316 (2020)
Advance Access publication 2020 September 14

doi:10.1093/mnras/staa2706

Higher order spectra of weak lensing convergence maps in parametrized theories of modified gravity

D. Munshi★ and J. D. McEwen

Mullard Space Science Laboratory, University College London, Holmbury St Mary, Dorking, Surrey RH5 6NT, UK

Accepted 2020 September 1. Received 2020 September 1; in original form 2020 April 18

ABSTRACT

We compute the low-ℓ limit of the family of higher order spectra for projected (2D) weak lensing convergence maps. In this limit these spectra are computed to an arbitrary order using *tree-level* perturbative calculations. We use the flat-sky approximation and Eulerian perturbative results based on a generating function approach. We test these results for the lower order members of this family, i.e. the skew- and kurt-spectra against state-of-the-art simulated all-sky weak lensing convergence maps and find our results to be in very good agreement. We also show how these spectra can be computed in the presence of a realistic sky-mask and Gaussian noise. We generalize these results to 3D and compute the *equal-time* higher order spectra. These results will be valuable in analysing higher order statistics from future all-sky weak lensing surveys such as the *Euclid* survey at low-ℓ modes. As illustrative examples, we compute these statistics in the context of the *Horndeski* and *beyond Horndeski* theories of modified gravity. They will be especially useful in constraining theories such as the Gleyzes–Langlois–Piazza–Vernizzi (GLPV) theories and degenerate higher order scalar-tensor theories as well as the commonly used normal-branch of Dvali–Gabadadze–Porrati model, clustering quintessence models and scenarios with massive neutrinos.

Key words: gravitational lensing: weak – cosmology: theory – large-scale structure of Universe.

1 INTRODUCTION

We have a standard model of cosmology thanks to recently completed cosmic microwave background radiation (CMBR) experiments such as the *Planck* Surveyor[1] (Planck Collaboration XVI XIII I 2014, 2016a, 2020). However, many of the outstanding questions pertaining, e.g. to the nature of dark matter and dark energy or possible modification of General Relativity (GR) on cosmological scales remain unanswered (Clifton et al. 2012; Joyce et al. 2015). In addition they will also provide an estimate of the sum of the neutrino masses (Lesgourgues & Pastor 2006). Ongoing and planned future large-scale structure (LSS) surveys may resolve or will provide clues for these questions using weak lensing analyses. Observational programs of many such surveys, including *Euclid*[2] (Laureijs et al. 2011), CFHTLS,[3] PAN-STARRS,[4] Dark Energy Surveys (DES)[5] (Abott et al. 2016), WiggleZ[6] (Jurek et al. 2010), LSST[7] (Tyson et al. 2003), BOSS[8] (Eisenstein et al. 2011), KiDS (Kuijiken et al. 2015), and WFIRST (National Research Council) lists weak lensing as their main science driver.

From the early days of detection the weak lensing (Munshi et al. 2008) studies have matured to a point where weak lensing results from *Euclid* are expected to constrain the cosmological parameters to sub-per cent accuracy. However, weak lensing at smaller angular scales probes the non-linear regime of gravitational clustering, and is thus key to understanding the non-Gaussianity induced by the non-linearity and fullly exploiting in the weak lensing maps. The higher order statistics are also useful for the breaking of parameter degeneracies in studies involving the power spectrum analysis alone and they are also important in understanding the variance or error of estimation of lower order statistics. These higher order statistics including the cumulants (Bernardeau 1994b) and their correlators (Bernardeau 1996a; Calabrese et al. 2010; Munshi et al. 2011; Riquelme & Spergel 2012) are among the best-known diagnostics of the deviation from Gaussianity that characterizes the non-linear regime (Bartolo et al. 2004), with a long history analytical modelling (Bernardeau et al. 2002a). Most of these studies use extensions of perturbative results in the quasi-linear regime valid at large smoothing angular scales or variants of halo models (Cooray & Sheth 2002). Early studies concentrated on measurements of higher order correlation hierarchy in the angular space due to small survey size (Bernardeau,

★ E-mail: d.munshi@ucl.ac.uk

[1] Planck
[2] http://sci.esa.int/euclid/
[3] http://www.cfht.hawai.edu/Sciences/CFHLS
[4] http://pan-starrs.ifa.hawai.edu/
[5] https://www.darkenergysurvey.org/
[6] http://wigglez.swin.edu.au/
[7] http://www.lsst.org/llst home.shtml
[8] http://www.sdss3.org/surveys/boss.php

Mellier & van Waerbeke 2002b; Bernardeau, Waerbeke & Mellier 2003). However, the near all-sky coverage of future surveys such as *Euclid* will let us estimate higher order statistics in the harmonic domain with unprecedented accuracy (Amendola et al. 2013). While measurements of real space correlations are much simpler in the presence of complicated survey design, the measurements for different angular scales can be highly correlated (Munshi 2000; Munshi & Jain 2000). In comparison measurements in the harmonic domain are less correlated and each mode contains (nearly) independent information in the limit of all-sky coverage. The primary motivation of this study is to develop analytical predictions for one such statistic called the skew-spectrum, and test them against numerical simulations. We will borrow the concepts developed for constructing skew-spectra for the study of non-Gaussianity in the context of CMBR observations (Planck Collaboration XVII 2016b). However, we also include gravity-induced secondary non-Gaussianity. The skew-spectrum is the lowest order member in the family of higher order spectra (Munshi et al. 2011a, 2020). In a series of papers, the one-point statistics such as the skewness and kurtosis were generalized to two-point cumulant correlator, e.g. the two-to-one correlator and its higher order generalizations. These can be represented in the harmonic domain by their associated spectra such as the skew-spectrum (Munshi & Heavens 2010) and its higher order generalizations (Munshi et al. 2011a, 2020). These spectra have already been used to analyse WMAP[9] (Smidt et al. 2010) as well as *Planck* data (Planck Collaboration XVII 2016b). They are useful tools to separate individual contributions and estimate systematics. In this paper, we will concentrate on the projected skew-spectrum and kurt-spectrum in the context of weak lensing surveys (Munshi et al. 2011c).

Other similar estimators also exist, including the morphological estimators (Munshi et al. 2012), e.g. position-dependent power spectra (Munshi 2017), phase correlations (Matsubara 2007), line-correlations (Eggemeier & Smith 2017), peak-statistics (Peel et al. 2017), peak-correlations (Heavens & Gupta 2001), and extreme value statistics (Harrison & Coles 2011).

Many modified gravity theories are now severely constrained with the first detection of GW170817 (Abott et al. 2017) and its electromagnetic counterpart GRB 170817A (Goldstein et al. 2017) implying Gravity Waves travel at the speed of light with deviation smaller than few $\times\ 10^{-15}$ – see e.g. Baker et al. (2017), Sakstein & Jain (2017), Lombriser & Lima (2017), and Creminelli & Verniizzi (2017). However, some of the models we consider here are designed to evade this constraint. It is expected that the constraints on these models will be further tightened by the observations of LSS by *Euclid* and LSST. The higher order statistics we develop here can be very effectively used to test these models independently or jointly with power spectrum estimates. As a concrete example of the higher order spectra we take the modified gravity theories also known as the Horndeski's theory of gravity. These are the most general theory of gravity that has second-order equation of motion. It was proposed first in 1974 (Barthelemy 1974) and since then, it was realized that Horndeski theory contains many other theories of gravity as a special case. These include GR, Jordon–Brans–Dicke theories of gravity, Dilaton and Chameleon theories of gravity, theories involving as co-variant Galileons as well as models of Quintessence. All of these models of gravity have found use in construction of cosmological models of inflation as well as dark energy (see e.g. Deffayet et al. 2011; Gleyzes et al. 2015a, b; Langois & Noui ; Barthelemy et al. 2020; Langois & Noui 2016a, for an incomplete list of recent references). We use a recent parametrization of the gravity induced bispectrum in this model as well as models that are known as the beyond Horndeski theories to compute the skew-spectrum in the low-ℓ limit.

This paper is organized as follows. In Section 2, we review results regarding the convergence bispectrum in the context of tree-level standard perturbation theory (SPT). In Section 2.3, we introduce the skew-spectrum and relate it to the bispectrum. The corresponding results for trispectrum and kurt-spectra are derived in Section 2.4. Theoretical predictions in the context of generating functions are derived in Section 3, The generalization of higher order spectra is presented in Section 4. The higher order spectra can be derived in the presence of a mask. The corresponding results are presented in Section 5. The simulations are discussed in Section 6, the numerical results are presented in Section 7. We present the results for various modified gravity and other beyond Lambda cold dark matter (ΛCDM) scenarios in Section 8. Finally, the conclusions are drawn in Section 9.

2 MODELLING OF HIGHER ORDER WEAK LENSING SPECTRA

In this section, we will review the aspects of standard tree-level perturbative which we use to compute the bispectrum as well trispectrum as and eventually the skew-spectrum and kurt-spectrum.

2.1 Tree-level perturbative calculations

In the quasi-linear regime ($\delta \leq 1$), the evolution of density contrast δ can be described using SPT (Munshi et al. 2008). However, the treatment based on perturbation theory breaks down when density contrast at a given length-scale becomes non-linear ($\delta \geq 1$) which significantly increases the growth of clustering. We will denote the Fourier transform of the density contrast $\delta(\mathbf{r})$ by $\delta(\mathbf{k})$, where $\mathbf{r}$ is the comoving co-ordinate, and $\mathbf{k}$ denotes the comoving wavenumber. Expanding the $\delta(\mathbf{k})$, in a perturbative series, and assuming the density contrast is less than unity, for the pertubative series to be convergent, we get

$$\delta(\mathbf{k}) = \delta^{(1)}(\mathbf{k}) + \delta^{(2)}(\mathbf{k}) + \delta^{(3)}(\mathbf{k}) + \cdots . \tag{1a}$$

[9] https://map.gsfc.nasa.gov/

The nth order perturbative term denoted as $\delta^{(n)}$ is $\propto [\delta^{(1)}]^n$ where $\delta^{(1)}$ is the linear density contrast. The term $\delta^{(n)}$ is expressed using a kernel F_n using the following convolution:

$$\delta^{(n)}(\mathbf{k}) = \int \mathrm{d}\mathbf{k}_1 . \cdots \int \mathrm{d}\mathbf{k}_n \delta_{3\mathrm{D}}(\mathbf{k_1} + \cdots + \mathbf{k_n} - \mathbf{k}) F_n(\mathbf{k}_1, \cdots, \mathbf{k}_n) \delta^{(1)}(\mathbf{k}_1) \cdots \delta^{(1)}(\mathbf{k}_n); \quad \mathrm{d}\mathbf{k} = \frac{\mathrm{d}^3\mathbf{k}}{(2\pi)^{3/2}}. \tag{1b}$$

The Dirac delta function in 3D is denoted by $\delta_{3\mathrm{D}}$ and $\mathbf{k}_1$, $\mathbf{k}_2$, $\cdots$, $\mathbf{k}_n$ denotes different wavenumbers. The second-order kernel F_2 has the following expression. For the higher order kernels see (Munshi et al. 2008)

$$F_2(\mathbf{k}_1, \mathbf{k}_2) = \frac{5}{7} + \frac{1}{2}\left(\frac{k_1}{k_2} + \frac{k_2}{k_1}\right)\left(\frac{\mathbf{k}_1 \cdot \mathbf{k}_2}{k_1 k_2}\right) + \frac{2}{7}\left(\frac{\mathbf{k}_1 \cdot \mathbf{k}_2}{k_1 k_2}\right)^2, \quad k_i = |\mathbf{k}_i|. \tag{1c}$$

Throughout we will use the following convention for the 3D fourier transform (FT) and its inverse

$$\delta(\mathbf{k}) = \int \mathrm{d}\mathbf{r} \exp(-i\mathbf{k}\cdot\mathbf{r})\delta(\mathbf{r}); \quad \delta(\mathbf{r}) = \int \mathrm{d}\mathbf{k} \exp(i\,\mathbf{k}\cdot\mathbf{r})\delta(\mathbf{k}); \qquad \mathrm{d}\mathbf{r} = \frac{\mathrm{d}^3\mathbf{r}}{(2\pi)^{3/2}}. \tag{2}$$

We have suppressed the temporal dependence in equations (1a) and (1b) which will be introduced later in this section. The power spectrum P_δ and the bispectrum B_δ of δ are defined, respectively, as the two- and three-point correlation of the variable $\delta(\mathbf{k})$. The $P_{\rm lin}(k)$ denotes the linear power spectrum, i.e. $\delta_{\rm lin}(\mathbf{k}) = \delta^{(1)}(\mathbf{k})$ and $\langle \delta_{\rm lin}(\mathbf{k}_1)\delta_{\rm lin}(\mathbf{k}_2)\rangle_c = (2\pi)^3\delta_{3\mathrm{D}}(\mathbf{k}_1 + \mathbf{k}_2)P_{\rm lin}(k_1)$. Throughout angular brackets represent ensemble averaging. The subscript $_{\rm lin}$ stands for linear-order contributions.

The linearized solution for the density field is $\delta^{(1)}(\mathbf{k})$; higher order terms yield corrections to this linear solution. Using an *ideal* fluid approach known to be valid at large scales (and before shell crossing) one can write the second-order correction to the linearized density field using the kernel $F_2(\mathbf{k_1}, \mathbf{k_2})$. The cosmological structure formation is described by a set of equation which describes the Newtonian gravity coupled to the Euler and continuity equation (Munshi et al. 2008). This system of non-linear, coupled integro-differential equations are used to compute the kernels $F_2(\mathbf{k}_1, \mathbf{k}_2)$, $F_3(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)$, and their higher order counterparts. This is typically done perturbatively in an order-by-order manner.

2.2 Weak lensing statistics in projection (2D)

We will now specialize our discussion to weak lensing surveys. The weak lensing convergence κ is a line-of-sight projection of the 3D density contrast $\delta(\mathbf{r})$

$$\kappa(\widehat{\Omega}; r_{\rm s}) = \int_0^{r_{\rm s}} \mathrm{d}r\, w(r, r_{\rm s})\, \delta(r, \widehat{\Omega}); \quad w(r, r_{\rm s}) = \frac{3\Omega_{\rm M}}{2}\frac{H_0^2}{ac^2}\frac{\mathrm{d}_A(r)\mathrm{d}_A(r_{\rm s} - r)}{\mathrm{d}_A(r_{\rm s})}, \tag{3}$$

where r is the comoving distance, $\widehat{\Omega} = (\theta, \phi)$ is a unit vector that defines the position of the pixel on the surface of the sky, with θ and ϕ, respectively, representing the azimuthal and polar coordinates $\mathrm{d}\widehat{\Omega} = \sin\theta\, \mathrm{d}\theta\, \mathrm{d}\varphi$ is the measure of integration, $r_{\rm s}$ is the radial comoving distance to the source plane, c is the speed of light, a represents the scale factor, H_0 the Hubble parameter, $d_A(r)$ is the comoving angular diameter distance and the 3D density contrast δ, and $\Omega_{\rm M}$ is the cosmological density parameter. We will ignore the source distribution and assume them to be localized on a single source plane, we will also ignore photometric redshift errors. However, such complications are essential to link predictions to observational data and can readily be included in our analysis. To avoid cluttering, we will suppress the $r_{\rm s}$ dependence of $\kappa(\widehat{\Omega}, r_{\rm s})$ and $w(r, r_{\rm s})$ defined in equation (1a) in the following. The corresponding 3D power spectrum P_δ, bispectrum B_δ and trispectrum T_δ for δ are

$$\langle \delta(\mathbf{k}_1)\delta(\mathbf{k}_2)\rangle_c = (2\pi)^3\delta_{3\mathrm{D}}(\mathbf{k}_1 + \mathbf{k}_2)P_\delta(k_1); \quad k = |\mathbf{k}|; \tag{4a}$$

$$\langle \delta(\mathbf{k}_1)\delta(\mathbf{k}_2)\delta(\mathbf{k}_3)\rangle_c = (2\pi)^3\delta_{3\mathrm{D}}(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3)B_\delta(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3); \tag{4b}$$

$$\langle \delta(\mathbf{k}_1)\cdots\delta(\mathbf{k}_4)\rangle_c = (2\pi)^3\delta_{3\mathrm{D}}(\mathbf{k}_1 + \cdots + \mathbf{k}_4)T_\delta(\mathbf{k}_1, \cdots, \mathbf{k}_3). \tag{4c}$$

The subscript $_c$ denotes the fact that only connected diagrams are included in computing these statistics. The flat-sky power spectrum P^κ and bispectrum B^κ are similarly defined through (Munshi et al. 2008)

$$\langle \kappa(\mathbf{l}_1)\kappa(\mathbf{l}_2)\rangle_c = (2\pi)^2\delta_{2\mathrm{D}}(\mathbf{l}_1 + \mathbf{l}_2)P^\kappa(l_1); \tag{5a}$$

$$\langle \kappa(\mathbf{l}_1)\kappa(\mathbf{l}_2)\kappa(\mathbf{l}_3)\rangle_c = (2\pi)^2\delta_{2\mathrm{D}}(\mathbf{l}_1 + \mathbf{l}_2 + \mathbf{l}_3)B^\kappa(\mathbf{l}_1, \mathbf{l}_2, \mathbf{l}_3); \tag{5b}$$

$$\langle \kappa(\mathbf{l}_1)\cdots\kappa(\mathbf{l}_4)\rangle_c = (2\pi)^2\delta_{2\mathrm{D}}(\mathbf{l}_1 + \cdots + \mathbf{l}_3)T^\kappa(\mathbf{l}_1, \mathbf{l}_2, \mathbf{l}_3, \mathbf{l}_4). \tag{5c}$$

The wavenumbers $\mathbf{l}$, $\mathbf{l}_1$, $\cdots$ $\mathbf{l}_4$ are wavenumbers defined on the flat-patch of the sky. For a given radial distance r they are related to the projected 3D wavenumber by the relation $\mathbf{l} = \mathbf{k}_\perp/\mathrm{d}_A(r)$; where $\mathrm{d}_A(r)$ being the co-moving angular diameter distance defined before and $l = |\mathbf{l}|$. Using the *flat-sky* approximation as well as *Limber* and *pre-factor unity* approximation the projected power spectrum $P^\kappa(\mathbf{l})$ and bispectrum $B^\kappa(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)$ can be expressed, respectively, in terms of the 3D δ power spectrum $P_\delta(k)$ and bispectrum $B_\delta(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)$ (Munshi et al. 2008)

$$P^{\kappa}(l) = \int_0^{r_s} \mathrm{d}r \frac{w^2(r)}{\mathrm{d}_A^2(r)} P_\delta\left(\frac{l}{\mathrm{d}_A(r)}; r\right); \tag{6a}$$

$$B^{\kappa}(\mathbf{l}_1, \mathbf{l}_2, \mathbf{l}_3) = \int_0^{r_s} \mathrm{d}r \frac{w^3(r)}{\mathrm{d}_A^4(r)} B_\delta\left(\frac{\mathbf{l}_1}{\mathrm{d}_A(r)}, \frac{\mathbf{l}_2}{\mathrm{d}_A(r),}, \frac{\mathbf{l}_3}{\mathrm{d}_A(r)}; r\right); \tag{6b}$$

$$T^{\kappa}(\mathbf{l}_1, \mathbf{l}_2, \mathbf{l}_3, \mathbf{l}_4) = \int_0^{r_s} \mathrm{d}r \frac{w^4(r)}{\mathrm{d}_A^6(r)} T_\delta\left(\frac{\mathbf{l}_1}{\mathrm{d}_A(r)}, \frac{\mathbf{l}_2}{\mathrm{d}_A(r),}, \frac{\mathbf{l}_3}{\mathrm{d}_A(r)}, \frac{\mathbf{l}_4}{\mathrm{d}_A(r)}; r\right). \tag{6c}$$

The superscript κ correspond to the convergence field which these statistics correspond to. The function w is defined in equation (3). We will use different approximations introduced in Section 2 in equations (6a) and (6b) to compute the convergence or κ bispectrum.

2.3 Bispectrum and skew-spectrum

The spherical harmonic transform of a convergence map $\kappa(\widehat{\Omega})$, denoted as $\kappa_{\ell m}$, defined over the surface of the sky using spherical harmonics $Y_{\ell m}(\widehat{\Omega})$ can be used to define the multipoles $\kappa_{\ell m}$

$$\kappa_{\ell m} = \int \mathrm{d}\widehat{\Omega}\, Y_{\ell m}(\widehat{\Omega})\, \kappa(\widehat{\Omega}); \quad \widehat{\Omega} = (\theta, \varphi). \tag{7}$$

A Gaussian field is completely characterized by its power spectrum $\mathcal{C}_\ell^\kappa$ which is defined as $\mathcal{C}_\ell^\kappa = \langle \kappa_{\ell m} \kappa^*_{\ell m} \rangle$. Here, $\kappa^*_{\ell m}$ represents the complex conjugate of $\kappa_{\ell m}$. The flat sky power spectrum $P^\kappa(l)$ is identical to $\mathcal{C}_\ell^\kappa$ at high ℓ with the identification $l = \ell$. Bispectrum is the lowest order statistics that characterizes departure from Gaussianity that is defined as the three-point coupling of harmonic coefficients. Assuming isotropy and homogeneity the all-sky bispectrum $B^\kappa_{\ell_1\ell_2\ell_3}$ is defined as (Bartolo et al. 2004)

$$\langle \kappa_{\ell_1 m_1} \kappa_{\ell_2 m_2} \kappa_{\ell_3 m_3} \rangle_c \equiv B^\kappa_{\ell_1\ell_2\ell_3} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ m_1 & m_2 & m_3 \end{pmatrix}. \tag{8}$$

The quantity in parentheses is the well-known Wigner-3j symbol that enforces rotational invariance. It is only non-zero for the triplets (ℓ_1, ℓ_2, ℓ_3) that satisfy the triangular condition and $\ell_1 + \ell_2 + \ell_3$ is even. The reduced bispectrum $b^\kappa_{\ell_1\ell_2\ell_3}$ for convergence κ is defined through the following expression (Bartolo et al. 2004):

$$B^\kappa_{\ell_1\ell_2\ell_3} = \sqrt{\frac{(2\ell_1+1)(2\ell_2+1)(2\ell_3+1)}{4\pi}} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ 0 & 0 & 0 \end{pmatrix} b^\kappa_{\ell_1\ell_2\ell_3}. \tag{9}$$

The skew-spectrum is defined as the cross power spectrum formed by cross-correlating the squared κ^2 maps against the original map κ (Munshi & Heavens 2010)

$$\mathcal{S}^{(21)}_\ell = \frac{1}{2\ell+1} \sum_m \mathrm{Real}\{[\kappa^2]_{\ell m} [\kappa]^*_{\ell m}\} = \sum_{\ell_1\ell_2} B^\kappa_{\ell_1\ell_2\ell} J_{\ell_1\ell_2\ell}; \tag{10a}$$

$$J_{\ell_1\ell_2\ell} = \sqrt{\frac{(2\ell_1+1)(2\ell_2+1)}{(2\ell+1)}} \begin{pmatrix} \ell_1 & \ell_2 & \ell \\ 0 & 0 & 0 \end{pmatrix}. \tag{10b}$$

To avoid cluttering we will not explicitly display smoothing windows in our equations. The beam-smoothed versions of the expressions can be recovered by using the smoothed harmonics i.e. replacing $k_{\ell m}$ with $\kappa_{\ell m} b_\ell$ where b_ℓ is the smoothing beam in the harmonic domain which can be top hat or Gaussian. In case of Gaussian smoothing, the expressions are derived in an order-by-order manner (Bernardeau 1994b, 1996b) For a top-hat smoothing these expressions are derived using a generating function to an arbitrary order (Matsubara 2003). The normalized one-point skewness parameter (Bernardeau et al. 2002a) $S_3 = \langle \kappa^3 \rangle_c / \langle \kappa^2 \rangle_c^2$ can be recovered from the skew-spectrum by constructing the beam-smoothed third-order moment $\langle \kappa^3 \rangle_c$ (Munshi & Heavens 2010)

$$\mu_3 = \langle \kappa^3 \rangle_c = \sum_\ell (2\ell+1) S^{(21)}_\ell = \sum_{\ell_1\ell_2\ell} J_{\ell_1\ell_2\ell} B^\kappa_{\ell_1\ell_2\ell}. \tag{11}$$

The normalized skewness parameter S_3^κ is defined as $S_3^\kappa = \mu_3/\mu_2^2$ with $\mu_{\mathrm{N}} = \langle \kappa^{\mathrm{N}} \rangle_c$ and $S^\kappa_{\mathrm{N}} = \mu_{\mathrm{N}}/\mu_2^{\mathrm{N}-1}$. The subscript $_c$ denotes that only the connected contributions are taken into account.

The real space two-to-one correlation function can be defined in terms of the skew-spectrum as (Munshi & Heavens 2010)

$$\xi^{(21)}(\theta_{12}) = \langle \kappa^2(\widehat{\Omega}_1) \kappa(\widehat{\Omega}_2) \rangle_c = \frac{1}{4\pi} \sum_\ell (2\ell+1) S^{(21)}_\ell P_\ell(\cos\theta_{12}), \tag{12}$$

where P_ℓ represents the Legendre Polynomial, and the angular positions $\widehat{\Omega}_1$ and $\widehat{\Omega}_2$ are separated by an angle θ_{12}. Suitably normalized two-to-one correlator is the lowest order of a family of statistics also known as cumulant correlator (Bernardeau 1996a; Calabrese et al. 2010; Munshi et al. 2011; Riquelme & Spergel 2012), which has also been used in the context of weak-lensing surveys (Munshi 2000; Munshi et al. 2020).

In our notation δ_{2D} is the 2D Dirac delta function. The flat-sky bispectrum $B^{\kappa}(\mathbf{l}_1, \mathbf{l}_2, \mathbf{l}_3)$ is identical to the reduced bispectrum $b_{\ell_1\ell_2\ell_2}$ for high multipole (Bartolo et al. 2004). This can be shown by noting the following asymptotic relationship:

$$\mathcal{G}_{\ell_1 m_1, \ell_2 m_2, \ell_3 m_3} \equiv \int \mathrm{d}\widehat{\Omega} Y_{\ell_1 m_1}(\widehat{\Omega}) Y_{\ell_2 m_2}(\widehat{\Omega}) Y_{\ell_3 m_3}(\widehat{\Omega});$$
$$= \sqrt{\frac{(2\ell_1+1)(2\ell_2+1)(2\ell_3+1)}{4\pi}} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \approx (2\pi)^2 \delta_{2D}(\mathbf{l}_1 + \mathbf{l}_2 + \mathbf{l}_3). \tag{13}$$

A few comments about the skew-spectrum are in order. One-point statistics such as the skewness parameter have the advantage of having high signal-to-noise ratio. However, they lack distinguishing power as all the available information in the bispectrum is compressed into a single number. In contrast, the skew-spectrum encodes some information on the shape of the spectrum, and in principle can allow us to separate the contribution from gravity-induced non-Gaussianity or possible source of contamination from systematics. Though primordial non-Gaussianity is highly constrained in the light of *Planck* data, such contributions can also tested using the skew-spectrum.

In this paper, we consider a direct estimator for the skew-spectrum as opposed to the optimal estimator developed in Munshi & Heavens (2010) where optimality was achieved by using suitable weights to the harmonics that incorporates a match filtering as well as saturates the Cramer–Rao limit in the limit of weakly non-Gaussian limit. Indeed, a simple Fisher matrix based analysis, however, will non-longer be adequate for moderately non-Gaussian weak lensing maps. However, optimality is not of crucial importance of analysis for weak lensing maps as the secondary non-Gaussianity is expected to be detected with much higher signal-to-noise ratio. A simpler direct estimator will thus be useful for studying non-Gaussianity in weak-lensing maps.

2.4 Trispectrum and kurt-spectra

The near all-sky weak lensing maps from surveys such as *Euclid* will also allow determination of non-Gaussianity statistics beyond the lowest order, e.g. the fourth-order correlator or the trispectrum. Trispectrum can be useful not only to construct the covariance of the power spectrum estimator but also as a consistency check for the lower order estimators. In this section, we will extend the estimator present above for the bispectrum to the case of trispectrum.

The trispectrum $T^{\ell_1\ell_2}_{\ell_3\ell_4}(L)$ can be defined by the following expressions from the four-point correlation function of the spherical harmonics $\kappa_{\ell m}$ for the convergence field κ (Munshi et al. 2011a):

$$\langle \kappa_{\ell_1 m_1} \kappa_{\ell_2 m_2} \kappa_{\ell_3 m_3} \kappa_{\ell_4 m_4} \rangle_c = \sum_{LM} (-1)^M T^{\ell_1\ell_2}_{\ell_3\ell_4}(L) \begin{pmatrix} \ell_1 & \ell_2 & L \\ m_1 & m_2 & M \end{pmatrix} \begin{pmatrix} \ell_3 & \ell_4 & L \\ m_3 & m_4 & -M \end{pmatrix}; \tag{14a}$$

$$T^{\ell_1\ell_2}_{\ell_3\ell_4}(L) = (2L+1) \sum_M \sum_{m_i} \begin{pmatrix} \ell_1 & \ell_2 & L \\ m_1 & m_2 & M \end{pmatrix} \begin{pmatrix} \ell_3 & \ell_4 & L \\ m_3 & m_4 & -M \end{pmatrix} \langle \kappa_{\ell_1 m_1} \kappa_{\ell_2 m_2} \kappa_{\ell_3 m_3} \kappa_{\ell_4 m_4} \rangle_c. \tag{14b}$$

Here, M is the magnetic quantum number associated with the azimuthal quantum number L. The Wigner 3j-symbols above ensure that the triangle inequalities imposed by statistical isotropy and homogeneity of the trispectrum in the harmonic space is represented by a quadrilateral. The harmonics ℓ_1, ℓ_2, ℓ_3, and ℓ_4 represent the sides of the quadrilateral and the harmonics L represents one of the diagonal of the quadrilateral. The two kurt-spectra $\mathcal{K}^{(31)}_{\ell}$ and $\mathcal{K}^{(31)}_{\ell}$ are defined as (Munshi et al. 2011a, 2020)

$$\mathcal{K}^{(31)}_{\ell} = \frac{1}{2\ell+1} \sum_m \mathrm{Real}\{[\kappa^3]_{\ell m} [\kappa]^*_{\ell m}\} = \sum_{\ell_1\ell_2\ell_3 L} T^{\ell_3\ell}_{\ell_1\ell_2}(L) J_{\ell_1\ell_2 L} J_{L\ell_3\ell}; \tag{15a}$$

$$\mathcal{K}^{(22)}_{\ell} = \frac{1}{2\ell+1} \sum_m \{[\kappa^2]_{\ell m} [\kappa^2]^*_{\ell m}\} = \sum_{\ell_1\ell_2\ell_3\ell_4} T^{\ell_3\ell_4}_{\ell_1\ell_2}(\ell) J_{\ell_1\ell_2\ell} J_{\ell_3\ell_4\ell}. \tag{15b}$$

Thus the kurt-spectra described above are computed using either by keeping the diagonal fixed and summing over all possible configurations (the two-to-two kurt-spectra $\mathcal{K}^{(2,2)}_{\ell}$ defined in equation 15a) or by keeping one of the side fixed and summing over all possible configurations (introduced above as three-to-one kurt-spectra $\mathcal{K}^{(31)}_{\ell}$ defined in equation 15b). These *states* are linked to the collapsed and squeezed configurations. At higher order the polyspectra are characterized by a polygon. The number of polyspectra at a given order can be high since the number of diagonals and sides of such polygons can be be quite high.

Another related point is that disconnected contributions will exists even in the absence of noise. These contributions need to subtracted out when estimating from the data (Hu 2001; Okamoto & Hu 2002). The trispectrum in this case is given in equation (17) and is specified completely by the power spectrum $\mathcal{C}_{\ell}$. The corresponding spectra are given in terms of the Gaussian trispectrum $G^{l_1 l_2}_{l_3 l_4}(L)$ (Munshi et al. 2011a)

$$\mathcal{G}^{(31)}_{\ell} = \sum_{\ell_1\ell_2\ell_3 L} G^{\ell_3\ell}_{\ell_1\ell_2}(L) J_{\ell_1\ell_2 L} J_{L\ell_3\ell}; \quad \mathcal{G}^{(22)}_{\ell} = \sum_{\ell_1\ell_2\ell_3\ell_4} G^{\ell_3\ell_4}_{\ell_1\ell_2}(\ell) J_{\ell_1\ell_2\ell} J_{\ell_3\ell_4\ell}, \tag{16}$$

where the Gaussian trispectrum $G^{l_1 l_2}_{l_3 l_4}(L)$ is given by (Hu 2001; Okamoto & Hu 2002)

$$G^{l_1 l_2}_{l_3 l_4}(L) = (-1)^{l_1+l_3} \sqrt{(2l_1+1)(2l_3+1)} \mathcal{C}_{l_1} \mathcal{C}_{l_3} \delta_{L0} \delta_{l_1 l_2} \delta_{l_2 l_3}$$
$$+ (2L+1) \mathcal{C}_{l_1} \mathcal{C}_{l_2} \left[(-1)^{l_2+l_3+L} \delta_{l_1 l_3} \delta_{l_2 l_4} + \delta_{l_1 l_4} \delta_{l_2 l_3} \right]. \tag{17}$$

In (Munshi & Heavens 2010; Munshi et al. 2011a, 2020) optimal versions of skew- and kurt-spectra estimators were developed which requires weights based on target spectra. This method was used in investigating primordial spectra as the signal-to-noise ratio is rather low. However, for investigating the gravity induced secondary non-Gaussianity with surveys that have as high expected signal-to-noise ratio as *Euclid* optimization is not mandatory.

The commonly used kurtosis parameter S_4^κ (to be defined below) can be reconstructed from the kurt-spectra as follows (Munshi et al. 2011a):

$$\mu_4 = \langle\kappa^4(\widehat{\Omega})\rangle_c = \frac{1}{4\pi}\int \kappa^4(\widehat{\Omega})\mathrm{d}\widehat{\Omega} = \frac{1}{4\pi}\sum_L \sum_{\ell_1\ell_2\ell_3\ell_4} h_{\ell_1\ell_2 L}h_{\ell_3\ell_4 L}T^{\ell_1\ell_2}_{\ell_3\ell_4}(L); \tag{18a}$$

$$= \sum_\ell (2\ell+1)\mathcal{K}_\ell^{(31)} = \sum_\ell (2\ell+1)\mathcal{K}_\ell^{(2,2)}. \tag{18b}$$

We will use noise free simulations but in case of analysing noisy maps, the $\mathcal{C}_\ell$s will also include the noise contribution. The commonly used kurtosis is a normalized fourth-order one-point estimator (Bernardeau et al. 2002a) $S_4^\kappa = \frac{\langle\kappa^4\rangle_c}{\langle\kappa^2\rangle_c^3} = \left[\frac{\mu_4 - 3\mu_2^2}{\mu_2^3}\right]$. Here, $\mu_2 = 1/4\pi\sum_\ell(2\ell+1)\mathcal{C}_\ell$. The corresponding cumulant correlators for these spectra are defined in a manner similar to equation (12) (Munshi et al. 2020)

$$\xi^{31}(\theta_{12}) = \langle\kappa^3(\widehat{\Omega}_1)\kappa(\widehat{\Omega}_2)\rangle_c = \frac{1}{4\pi}\sum_\ell (2\ell+1)\mathcal{K}_\ell^{(31)}P_\ell(\cos\theta_{12}); \tag{19a}$$

$$\xi^{22}(\theta_{12}) = \langle\kappa^2(\widehat{\Omega}_1)\kappa^2(\widehat{\Omega}_2)\rangle_c = \frac{1}{4\pi}\sum_\ell (2\ell+1)\mathcal{K}_\ell^{(22)}P_\ell(\cos\theta_{12}). \tag{19b}$$

Next, we will employ tree-level perturbative calculations.

3 TREE-LEVEL PERTURBATIVE RESULTS

The unsmoothed normalized higher order cumulants or $S_N = \langle\delta^N\rangle_c/\langle\delta^2\rangle_c^{N-1}$ can be expressed in terms of the tree-level vertices denoted as ν_N using the following expressions (Bernardeau et al. 2002a):

$$S_3 = 3\nu_2; \quad S_4 = 4\nu_3 + 12\nu_2^2; \quad S_5 = 5\nu_4 + 60\nu_3\nu_2 + 60\nu_2^3. \tag{20}$$

The vertices ν_N are the angular averages of the mode-coupling kernels F_N defined in equation (1c) i.e. $\nu_N = N!\langle F_N\rangle$ introduced in Section 2.2 in the Fourier domain.

$$\nu_N = N!\langle F_N\rangle = N!\int\frac{\mathrm{d}\widehat{\Omega}_{k_1}}{4\pi}\cdots\int\frac{\mathrm{d}\widehat{\Omega}_{k_N}}{4\pi}F_N(\mathbf{k}_1.\cdots,\mathbf{k}_N); \quad \mathrm{d}\widehat{\Omega}_k = \sin\theta_k\mathrm{d}\theta_k\mathrm{d}\varphi_k. \tag{21}$$

The following generating function approach was introduced in Bernardeau (1992, 1994b). The generating functions $\mathcal{G}_\delta(\tau)$ are solved using the equations of gravitational dynamics encapsulated in Euler–Continuity–Poisson equations. Here, τ plays the role of a dummy variable. In the perturbative regime, the ν_N parameter can be computed for an arbitrary N

$$\mathcal{G}_\delta(\tau) = \sum_n \frac{\nu_N}{N!}\tau^N = -\tau + \frac{12}{14}\tau^2 - \frac{29}{42}\tau^3 + \frac{79}{147}\tau^4 - \frac{2085}{5096}\tau^5 + \cdots \tag{22}$$

Next, using equation (20), the one-point cumulants in 2D (Munshi et al. 1999b), denoted as Σ_N as opposed to S_N parameters which represent the cumulants in 3D, can be used to compute the cumulants to arbitrary order in 2D (Munshi et al. 1999b)

$$\Sigma_3 = \frac{36}{7}; \quad \Sigma_4 = \frac{2540}{49}; \quad \Sigma_5 = 793; \quad \Sigma_6 = 16370; \tag{23}$$

The generalization of the one-point cumulants i.e the S_N parameters to the two-point cumulant correlators $C_{pq} = \langle\delta_1^p\delta_2^q\rangle_c/\langle\delta^2\rangle_c^{p+q-1}\langle\delta_1\delta_2\rangle_c$ or C_{pq} parameters was introduced in Bernardeau (1996a). The lower order normalized cumulant correlators can also be expressed in terms of the tree-level vertices ν_N just as the one-point cumulants introduced in equation (20)

$$C_{21} = 2\nu_2; \quad C_{31} = 3\nu_3 + 6\nu_2^2; \quad C_{41} = 4\nu_4 + 36\nu_3\nu_2 + 24\nu_2^3. \tag{24}$$

The corresponding quantities for convergence maps can be defined in an analogous manner and are denoted with a superscript κ i.e. C^κ_{pq}. To compare with observed or simulated data smoothing of the field is necessary. The smoothed generating function $\mathcal{G}^s_\delta$ can be computed from the unsmoothed generating function $\mathcal{G}_\delta$. The generating functions $\mathcal{G}^s_\delta$ and $\mathcal{G}_\delta$ are related by the following implicit relation (Bernardeau 1995):

$$\mathcal{G}^s_\delta(\tau) = \mathcal{G}_\delta(\tau[1+\mathcal{G}^s_\delta]^{-(2+n)/4}). \tag{25}$$

A *top hat* smoothing window is assumed and the power spectrum is approximated locally as a power law $P(k)\propto k^n$ (Bernardeau 1995; Munshi et al. 1999b). For other window functions, e.g. Gaussian window generic results are not possible for arbitrary N. However, an order-by-order approach can be adopted to obtain the lower order cumulants (Matsubara 2003). Notice that the smoothed power law depends on the spectral index while unsmoothed vertices depend solely on the gravitational collapse in 3D spherical or cylindrical in 2D. The smoothed vertices can be recovered by Taylor-expanding the smoothed generating function $\mathcal{G}^s$. Using these vertices, it is possible to now compute the 2D skewness Σ_3 and kurtosis Σ_4 can be computed (Munshi et al. 1999b)

$$\Sigma_3 = \frac{36}{7} - \frac{3}{2}(n+2); \tag{26a}$$

$$\Sigma_4 = \frac{2540}{49} - 33(n+2) + \frac{21}{4}(n+2)^2. \tag{26b}$$

These expressions are derived using 2D where gravitational collapse with cylindrical symmetry is relevant as is the case for projected surveys. However, the underlying statistics for the 3D density field is linked to spherical collapse which we have not considered here but may be relevant for a 3D weak lensing scenario where photometric data is used. However, there is a crucial difference between 2D and 3D statistics. For large separations in 3D we can factorize $C_{pq} = C_{p1}C_{q1}$, while in 2D this approximation is not valid. Thus, we will consider the family of statistics C_{p1} for arbitrary p

$$\mathcal{S}_\ell^{(21)} = R_2\Sigma_{21}P^\kappa(l)\sigma_L^2 = R_2\left[\frac{24}{7} - \frac{1}{2}(n+2)\right]P_\delta(l)\sigma_L^2; \quad \sigma_L^2 = \langle\kappa^2\rangle; \tag{27a}$$

$$R_2 = \int_0^{r_s} \mathrm{d}\,r\,\frac{w^3(r)}{\mathrm{d}_A^{4+2n}(r)} \Big/ \left(\int_0^{r_s} \mathrm{d}\,r\,\frac{w^2(r)}{\mathrm{d}_A^{2+n}(r)}\right)^2. \tag{27b}$$

The corresponding result at the fourth-order is given by

$$\mathcal{K}_\ell^{(31)} = R_3\Sigma_{31}P^\kappa(l)\sigma_L^4 = R_3\left[\frac{1473}{49} - \frac{195}{14}(n+2) + \frac{3}{2}(n+2)^2\right]P_\delta(l)\sigma_L^4; \tag{28a}$$

$$R_3 = \int_0^{r_s} \mathrm{d}\,r\,\frac{w^4(r)}{\mathrm{d}_A^{6+3n}(r)} \Big/ \left(\int_0^{r_s} \mathrm{d}\,r\,\frac{w^2(r)}{\mathrm{d}_A^{2+n}(r)}\right)^3. \tag{28b}$$

The dynamical contribution is encoded on Σ_{p1} where as the line-of-sight integration is represented by the pre-factors in R_p. In our notation, the skewness and kurtosis of a convergence map S_3^κ and S_4^κ are expressed as $S_3^\kappa = R_2\Sigma_2$ and $S_4^\kappa = R_3\Sigma_3$. The quantities Σ_2 and Σ_3 are defined, respectively, in equations (26a) and (26b). The corresponding cumulant correlators for κ can now be expressed as $C_{21}^\kappa = R_2\Sigma_{21}$ and $C_{31}^\kappa = R_3\Sigma_{31}$.

Historically, the generating function approach was developed without any reference to perturbative dynamics and the vertices were left undetermined. Many generic predictions were developed coupling scaling Ansätze with the generating function formalism (Balian & Schaeffer 1989). While in the quasi-linear regime the loop corrections to the tree-level results violate the scaling Ansatz, in the highly non-linear regime the vertices are known to become shape independent parameter as encapsulated in Hyper Extended Peturbation Theory (Scoccimarro & Frieman 1999). In recent years some of the results were derived the Large Deviation Principle (Bernardeau & Reimberg 2016; Uhlemann et al. 2016; Reimberg & Bernardeau 2018; Uhlemann et al. 2018).

Previously, many studies have focused on observed and simulated data of one-point cumulants (Bernardeau 1995; Munshi et al. 1998) as well as for the two-point cumulant correlators (Munshi, Coles & Melott 1999a; Bernardeau 1995). Previous studies have focused on galaxy surveys. In this paper, we extend these results to the context of weak lensing.

4 HIGHER ORDER SPECTRA IN 3D

Next, we will consider higher order statistics in 3D. Future surveys such as *Euclid* will go beyond the projection and using photometric redshifts will be able to retain radial information. In 3D, we will compute the higher order spectra as before in the low-ℓ limit. The results have similar characteristics as in projection, which we have discussed in Section 3 but are very different in certain aspect as we discuss below. We will decompose the lensing field in two different eigenmodes (a) Fourier–Bessel decomposition typically used for radially symmetric fields and (b) the generic Fourier–Cartesian decomposition that are most commonly used for perturbative analysis. The results in this section are kept very general. While relating these results to weak lensing observables, complications due to photometric redshift measurements need to be included.

We will follow the same convention for forward and reverse Fourier transformation introduced in equation (2) for the Cartesian co-ordinate. For an arbitrary function $A(\mathbf{r})$ with $\mathbf{r} \equiv (r, \widehat{\Omega}) = (r, \theta, \phi)$ and its Fourier transform $A(\mathbf{k}; r)$ we will use

$$\mathrm{A}(\mathbf{r}; r) = \int \mathrm{d}\mathbf{k}\,\mathrm{A}(\mathbf{k}; r)\,\exp(i\,\mathbf{k}\cdot\mathbf{r}); \quad \mathrm{A}(\mathbf{k}; r) = \int \mathrm{d}\mathbf{r}\,\mathrm{A}(\mathbf{r}; r)\,\exp(i\,\mathbf{k}\cdot\mathbf{r}). \tag{29}$$

In spherical-Bessel coordinates the eigenfunctions of the Laplacian operators are the products of spherical harmonics $Y_{\ell m}(\widehat{\Omega})$ and spherical Bessels function $j_\ell(r)$ i.e. $j_\ell(kr)Y_{\ell m}(\widehat{\Omega})$ the transforms take the following form:

$$\mathrm{A}_{\ell m}(k) \equiv \sqrt{\frac{2}{\pi}}\int \mathrm{d}^3\mathbf{r}\mathrm{A}(\mathbf{r})kj_\ell(kr)Y^*_{\ell m}(\widehat{\Omega}); \quad \mathrm{A}(\mathbf{r}) \equiv \sqrt{\frac{2}{\pi}}\int k\mathrm{d}k\sum_{\ell=0}^{\infty}\sum_{m-\ell}^{\ell}\mathrm{A}_{\ell m}(k)j_\ell(kr)Y_{\ell m}(\widehat{\Omega}). \tag{30}$$

Using the well-known Rayleigh expansion that expands the plane-wave in a spherical wave basis

$$\exp(i\,\mathbf{k}\cdot\mathbf{r}) = 4\pi\sum_\ell\sum_{m=-\ell}^{m=\ell} i^\ell j_\ell(kr)Y_{\ell m}(\widehat{\Omega}_k)Y_{\ell m}(\widehat{\Omega}); \qquad \widehat{\Omega}_k = (\theta_k, \phi_k). \tag{31}$$

We can relate the spherical harmonic coefficients $\mathrm{A}_{\ell m}$ with their Fourier counterpart A

$$\mathrm{A}_{\ell m}(k;r) = \frac{1}{(2\pi)^{3/2}}\, k\, i^{\ell} \int \mathrm{d}\widehat{\Omega}_k\, \mathrm{A}(\mathbf{k};r) Y_{\ell m}(\widehat{\Omega}_k). \tag{32}$$

The 3D power spectrum $P^{\mathrm{AA}}(k)$ defined, respectively, in Cartesian coordinates and $\mathcal{C}_\ell^{\mathrm{AA}}$ in spherical coordinates are

$$\langle \mathrm{A}(\mathbf{k})\mathrm{A}^*(\mathbf{k})\rangle = (2\pi)^3 P^{AA}(k); \quad \langle \mathrm{A}_{\ell m}(k,l)\mathrm{A}^*_{\ell' m'}(k',l')\rangle = (2\pi)^2 \mathcal{C}_\ell^{\mathrm{AA}}(k;r)\delta_{1\mathrm{D}}(k-k')\delta_{\ell\ell'}\delta_{mm'}. \tag{33}$$

In general, in the absence of any mask, it can be shown that: $\mathcal{C}_\ell = P(k)$ i.e. the 3D power spectrum in spherical co-ordinate is independent of ℓ and is actually same as the power spectrum in Cartesian co-ordinates (Castro, Heavens & Kitching 2005). Next, for the construction of the higher order 3D spectra we will define the following cross-spectra between two arbitrary 3D fields A(**r**) and B(**r**) in spherical co-ordinates

$$\langle \mathrm{A}(\mathbf{k})\mathrm{B}^*(\mathbf{k})\rangle = (2\pi)^3 P^{AB}(k); \quad \langle \mathrm{A}_{\ell m}(k;r)\mathrm{B}^*_{\ell m}(k';r)\rangle = (2\pi)^2 \mathcal{C}_\ell^{\mathrm{AB}}(k;r)\delta_{1\mathrm{D}}(k-k')\delta_{\ell\ell'}\delta_{mm'}. \tag{34}$$

Using this identity, for the 3D density field δ we can derive the following expressions for the higher order spectra of the density field

$$P^{\delta}(k;r,r') = \langle \delta(\mathbf{k};r)\delta^*(\mathbf{k}',r')\rangle_c; \qquad \mathcal{C}_\ell^{\delta}(k;r,r') = \langle \delta_{\ell m}(k;r)\delta^*_{\ell m}(k;r')\rangle_c; \quad P^{\delta}(k;r,r') = \mathcal{C}_\ell^{\delta}(k;r,r'). \tag{35a}$$

$$S^{21,\delta}(k;r,r') = \langle \delta^2(\mathbf{k};r)\delta^*(\mathbf{k}';r')\rangle_c\,; \qquad S_\ell^{21,\delta}(k;r,r') = \langle \delta^2_{\ell m}(k;r)\delta^*_{\ell m}(k;r')\rangle_c; \quad S^{21,\delta}(k,r,r') = S_\ell^{21,\delta}(k;r,r'). \tag{35b}$$

$$T^{31,\delta}(k;r,r') = \langle \delta^3(\mathbf{k};r)\delta^*(\mathbf{k}';r')\rangle_c\,; \quad T_\ell^{31,\delta}(k,r) = \langle \delta^2_{\ell m}(k;r)\delta^*_{\ell m}(k;r')\rangle_c; \quad T^{31,\delta}(k,r,r') = T_\ell^{31,\delta}(k,r,r'). \tag{35c}$$

$$T^{22,\delta}(k,r,r') = \langle \delta^2(\mathbf{k};r)\delta^{2*}(\mathbf{k}';r')\rangle_c\,; \quad T_\ell^{22,\delta}(k;r,r') = \langle \delta^2_{\ell m}(k;r)\delta^{2*}_{\ell m}(k;r')\rangle_c; \quad T^{22,\delta}(k,r,r') = T_\ell^{22,\delta}(k;r,r'). \tag{35d}$$

In our notation, $\delta^p(k)$ is the FT of δ^p. Notice these expressions are non-perturbative and are valid irrespective of detailed modelling and are valid to an arbitrary order i.e. when cross-correlating pth power of δ against the qth power i.e. in $\langle \delta^p(\mathbf{k})\delta^q(\mathbf{k})\rangle$ in spherical or Cartesian co-ordinate. In the Cartesian co-ordinate, the normalized cumulant correlators C_{pq} are defined as follows:

$$\langle \delta^p(\mathbf{k})\delta^{q*}(\mathbf{k})\rangle_c = C_{pq}\langle\delta^2\rangle_c^{p+q-2} P(k) = C_{pq}\langle\delta^2\rangle_c^{p+q-2}\mathcal{C}_\ell(k). \tag{36}$$

The second step relies on equation (35a). In the real-space equation (36) this is equivalent to:

$$\langle \delta^p(\mathbf{r}_1)\delta^{q*}(\mathbf{r}_2)\rangle_c = C_{pq}\langle\delta^2\rangle_c^{p+q-2}\langle\delta(\mathbf{r}_1)\delta(\mathbf{r}_2)\rangle_c. \tag{37}$$

The results equations (35a)–(35d) are non-perturbative and do not depend on any simplifying assumptions. However, in case of studies of galaxy clustering, it is more natural to study high-order statistics in the redshift space. Similarly, for 3D weak lensing, line-of-sight integration will need to be taken into account. Such extensions will be presented separately. The coefficients C_{pq} defined in equation (36) can be computed using perturbative calculations. In 3D, the smoothed and unsmoothed vertex generating functions are related through an implicit expression (Bernardeau 1996a) which is analogous to equation (25)

$$\mathcal{G}^s_\delta(\tau) = \frac{3}{2}\mathcal{G}_\delta(\tau[1+\mathcal{G}^s_\delta]^{-(3+n)/6}). \tag{38}$$

The power spectrum is assumed to be approximated locally by a power law with power-law index n i.e. $P(k) \propto k^n$. On Taylor expanding the 3D (unsmoothed) generating function $\mathcal{G}(\tau)$, we can recover the lower order vertices ν_N in 3D (Bernardeau et al. 2002a)

$$\mathcal{G}_\delta(\tau) = \sum_n \frac{\nu_N}{N!}\tau^N = -\tau + \frac{34}{21}\tau^2 - \frac{682}{189}\tau^3 + \cdots \tag{39}$$

Using these vertices it is possible to compute the normalized lower order moments i.e. skewness S_3 and kurtosis S_4 in 3D (Bernardeau et al. 2002a)

$$S_3 = \frac{34}{7} - (n+3); \qquad S_4 = \frac{60712}{1323} - \frac{62}{3}(n+3) + \frac{7}{3}(n+3)^3. \tag{40a}$$

The lower order cumulant correlators have the following form (Bernardeau et al. 2002a):

$$C_{21} = \frac{68}{21} - \frac{(n+3)}{3}; \qquad C_{31} = \frac{11710}{441} - \frac{61}{7}(n+3) + \frac{2}{3}(n+3)^3. \tag{40b}$$

Detailed derivations regarding construction of one- and two-point probability distribution functions are detailed in Bernardeau et al. (2002a). The 3D vertices defined in equation (39) assume a different numerical value though the formal structure remains the same. In addition a more generic result $C_{pq} = C_{p1}C_{q1}$ gives a much needed consistency check. The results in a 3D collapse are related to a spherical window and the dynamics relate to the 3D spherical collapse.

5 PSEUDO-$\mathcal{C}_\ell$ (PCL) ESTIMATORS

Maximum likelihood (ML) estimators or quadratic maximum likelihood (QML) estimators are most obvious choices for analysing cosmological data sets. However, these estimators require inverse covariance weighting which clearly is not practical for large cosmological data sets though

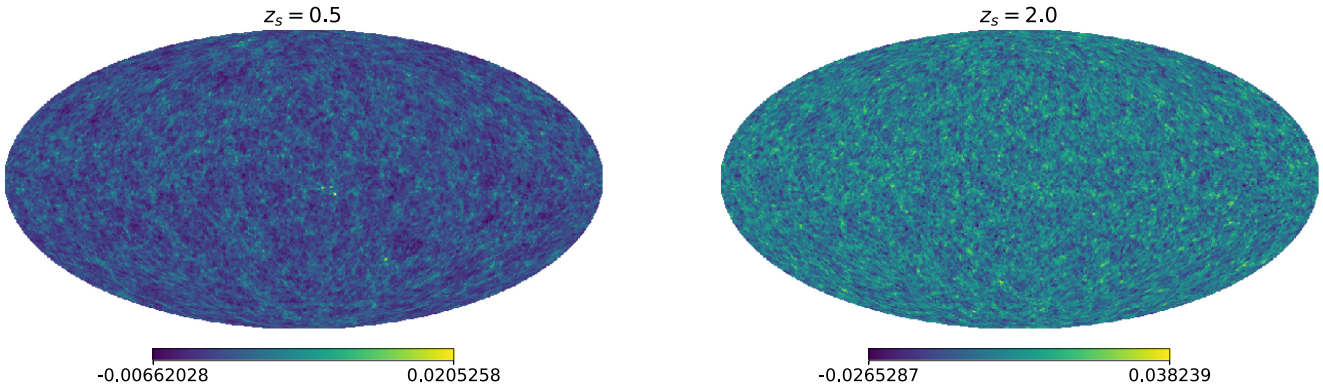

Figure 1. Examples of simulated κ maps used in our study. The left-hand panel corresponds to $z_s = 0.5$ while the right-hand panel corresponds to $z_s = 2.0$. The maps were generated at a resolution of $N_{\rm side} = 4096$. See Section 6 for more detail discussion about construction of maps used in our study.

various clever algorithmic techniques have been considered (Bernardeau et al. 2002a). This has resulted in the development of many sub-optimal estimators which uses heuristic weighting schemes. The so-called pseudo-$\mathcal{C}_\ell$ (PCL) technique was introduced in Hivon et al. (2002); see Szapudi et al. (2001) for a related method. These estimators are unbiased but sub-optimal. Various weighting schemes depending on sky coverage as well as noise characteristic as well as various hybridization schemes to combine large angular scale (equivalently the low ℓ) estimates using QML with small angular scale (high ℓ) PCL estimates were considered in (Efstathiou 2004)

$$M_{\ell\ell'} = (2\ell'+1)\sum_{\ell''}\begin{pmatrix}\ell & \ell' & \ell''\\ 0 & 0 & 0\end{pmatrix}^2 \frac{(2\ell''+1)}{4\pi}|w^2_{\ell''}|; \tag{41a}$$

$$\hat{\mathcal{S}}^{(21)}_\ell = \sum_{\ell'} M^{-1}_{\ell\ell'}\tilde{\mathcal{S}}^{(21)}_{\ell'}. \tag{41b}$$

Here, $\tilde{\mathcal{S}}^{(21)}_{\ell'}$ denotes the skew-spectrum computed from a map in the presence of a mask $w(\widehat{\Omega})$, $\hat{\mathcal{S}}^{(21)}_{\ell'}$ is the all-sky estimate and $w_\ell = 1/(2\ell+1)\sum_m w_{\ell m}w^*_{\ell m}$ is the power spectrum of the mask constructed from the harmonic-coefficient $w_{\ell m}$ of the map. The *coupling* matrix $M_{\ell\ell'}$ represents the mode mixing due to the presence of a mask. The theoretical expectation for the skew-spectrum is denoted as $S^{(21)}_\ell$ which is also identical to the ensemble average of its all-sky estimates i.e. $S^{(21)}_\ell = \langle\hat{S}^{(21)}_\ell\rangle$. The generalization of the PCL method to estimate higher order spectra were developed in Munshi et al. (2011a, 2020) for spin-0 fields and in higher spin fields in (Munshi et al. 2020) as well as in 3D in (Munshi et al. 2011b). Exactly same result holds for higher order spectra, e.g. for all-sky estimate of kurt-spectrum $\hat{\mathcal{K}}^{(31)}_\ell$ and its masked counterpart $\tilde{\mathcal{K}}^{(31)}_{\ell'}$ are related through a similar expression $\hat{\mathcal{K}}^{(31)}_\ell = \sum_{\ell'} M^{-1}_{\ell\ell'}\tilde{\mathcal{K}}^{(31)}_{\ell'}$. Following the discussion above on skew-spectrum, we also introduce the all-sky ensemble average $\mathcal{K}^{(31)}_\ell$ defined as $\mathcal{K}^{(31)}_\ell = \langle\hat{\mathcal{K}}^{(31)}_\ell\rangle$. This has also been generalized to reconstruct the Minkowski Functionals in an order-by-order manner (Munshi et al. 2012). Two equivalent techniques for flat-sky PCLs are developed here Asgari et al. (2018) and Hikage et al. (2011).

6 NUMERICAL SIMULATIONS

We use the publicly available all-sky weak-lensing maps generated by Takahashi et al. (2017)[10] that were generated using ray tracing through N-body simulations. Multiple lens planes were used to generate convergence (κ) as well as shear (γ) maps. Many recent studies were performed using these maps, e.g. Namikawa et al. (2018) and Munshi et al. (2020). In these simulations, the source redshifts used were in the range $z_{\rm s} = 0.05$–5.30 at interval $\Delta z_{\rm s} = 0.05$. In this study, we have used the maps with $z_{\rm s} = 0.5$, 1.0, 1.5, and 2.0. The maps do include post-Born corrections (Lewis & Pratten 2016). Though at the low source redshift such corrections only play a negligible role. Indeed, they do play significant role in CMB lensing. The convergence maps were generated using an equal area pixelization scheme in HEALPIX[11] format (Gorski et al. 2005). In this pixelization scheme, the number of pixels scale as $N_{\rm pix} = 12N^2_{\rm side}$ where $N_{\rm side}$ is the resolution parameter which can take values $N_{\rm side} = 2^N$ with $N = 1, 2, \cdots$. The set of maps we use are generated at $N_{\rm side} = 4096$ which were also cross-checked using higher resolution maps that were constructed at a resolution $N_{\rm side} = 8192, 16384$. These maps were found to be consistent with each other up to the angular harmonics $\ell \leq 3600$. In addition detailed tests were performed by using a electric/magnetic (E/B) decomposition of shear maps for the construction of κ maps (Takahashi et al. 2017). Though we have used high resolution maps $N_{\rm side} = 4096$, we have degraded them to low-resolution maps at $N_{\rm side} = 1024$ as we are primarily interested in the perturbative regime. The following set of cosmological parameters $\Omega_{\rm CDM} = 0.233$, $\Omega_b = 0.046$, $\Omega_{\rm M} = \Omega_{\rm CDM} + \Omega_b$, $\Omega_\Lambda = 1 - \Omega_{\rm M}$, and $h = 0.7$ were used to generate the maps assuming a ΛCDM background cosmology. The amplitude of density fluctuations $\sigma_8 = 0.82$ and the spectral index $n_{\rm s} = 0.97$. Examples of κ maps used in our study are presented in Fig. 1. We will focus on the large separation or the small ℓ regime in our study and we do not expect the baryonic feedback to

[10] http://cosmo.phys.hirosaki-u.ac.jp/takahasi/allsky_raytracing/
[11] https://healpix.jpl.nasa.gov/

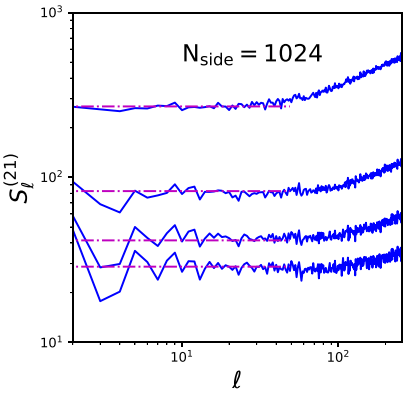

Figure 2. The skew-spectrum $S_\ell^{(21)}$ defined in equations (10a) and (10b) is shown as a function of the harmonics ℓ. From top to bottom, the curves correspond to source redshifts $z_s = 0.5$, 1.0, 1.5, and 2.0, respectively. A total of 10 simulations were used to compute the $S_\ell^{(21)}$. The straight lines at the left correspond to predictions from perturbation theory encapsulated in equations (27a) and (27b). We have assumed a power-law power spectrum $P_\delta(k) \propto k^n$. We have chosen $n = -2.0$ (dot–dashed lines). See text for details.

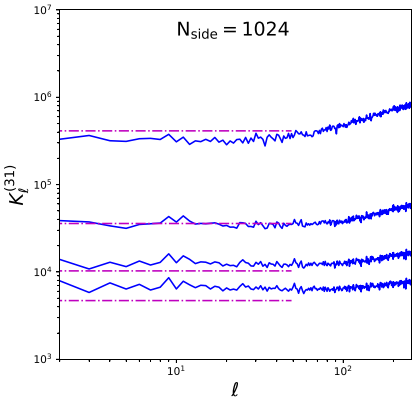

Figure 3. The kurt-spectrum $\mathcal{K}_\ell^{(31)}$ defined in equation (15a) is shown as a function of the harmonics ℓ. From top to bottom, the curves correspond to source redshifts $z_s = 0.5$, 1.0, 1.5, and 2.0, respectively. A total of 10 simulations were used to compute the $\mathcal{K}_\ell^{(31)}$. The straight lines at the left correspond to predictions from perturbation theory encapsulated in equations (28a) and (28b). We have assumed a power-law power spectrum $P_\delta(k) \propto k^n$. We have chosen $n = -2$. See text for details.

play a significant role (Weiss et al. 2019). It is worth mentioning here that these maps were also used to recently analyse the bispectrum the context of CMB lensing (Namikawa et al. 2018).

7 TESTS AGAINST NUMERICAL SIMULATIONS

The skew-spectrum $\mathcal{S}_\ell^{(21)}$ is shown as a function of the harmonics ℓ in Fig. 2. From top to bottom, the curves represent the source redshifts $z_s = 0.5$, 1.0, 1.5, and 2.0, respectively. The results are from maps with $N_{\rm side} = 1024$. We have analysed these maps for $\ell_{\rm max} = 2N_{\rm side}$. The straight lines correspond to perturbative results computed using tree-level perturbation theory equations (27a) and (27b). We have used an ensemble of 10 realizations to compute the mean that is being plotted. We use all-sky maps without an observational mask. The effect of mask can be incorporated using equations (41a) and (41b).

The $\mathcal{K}_\ell^{(31)}$ is shown as a function of the harmonics ℓ in Fig. 3. From top to bottom, the curves represent the source redshifts $z_s = 0.5$, 1.0, 1.5, and 2.0, respectively. The maps used are $N_{\rm side} = 1024$ and as before we have analysed for $\ell_{\rm max} = 2N_{\rm side}$. The straight lines correspond to perturbative results computed using tree-level perturbation theory equations (28a) and (28b). We have used an ensemble of 10 realizations to compute the mean that is being plotted. We use all-sky maps without an observational mask.

Our results for the skew- and kurt-spectra are derived in the large separation limit i.e. the cumulant correlators defined, e.g. in equation (12) and in equation (19b) $|\xi_{12}|/\bar{\xi}_2 \ll 1$. In real-space this limit was seen to be reached very fast as soon as the two neighbouring cells are not overlapping. In harmonic domain, the scale ℓ represents the separation of two beam-smoothed pixels for which the skew-spectrum is being measured. Thus, large separation in our case corresponds to low ℓ, and typical size of the pixels corresponds to the ℓ at which the beam can no longer be approximated as unity. This is the scale where the correction to the skew-spectrum starts to be non-negligible. These corrections, which are of order $\xi_{ij}/\bar{\xi}_2 \ll 1$, are difficult to compute analytically. Though, entire skew-spectrum can be computed with fitting functions.

Clearly, such a computation will not be possible beyond third order i.e. skew-spectrum due to lack of such a fitting function at the fourth order. Thus, the techniques developed here are valuable as their predictions can be made at *all-orders*.

The results we have computed are based on spherical top-hat window. However, many previous studies have shown that the actual shape of the window is not important, and replacing the circle with square can be very good approximation (Munshi et al. 1999b). However, profile of the smoothing beam or window as opposed to its shape can change the theoretical predictions. The predictions for a Gaussian window was worked out in detail in Matsubara (2003). However, results can be derived only in order-by-order manner and approaches based on generating functions are not applicable.

A few comments about going beyond kurt-spectrum are in order. Extraction and interpretation of higher order statistics can be rather complex from any cosmological data sets. Estimators of the cumulants and cumulant correlators are typically known to be biased and elaborate scheme was developed in estimating and correcting such bias as well as scatter in estimators typically used for evaluating these quantities mainly in real space in the context of galaxy clustering. Such corrections are expected to be more dominant role with increasingly higher order (Munshi et al. 1999b). Such corrections and their accurate calibration against simulations are lacking in the literature. Though, for lower order statistics we probed here, such corrections are expected to be negligible, a better understanding of such effects is needed before we can interpret the statistics beyond kurt-spectra (equivalently trispectrum).

An alternative approach considered by various authors was to consider the one-point and two-point probability functions which encode cumulants an their correlators to an arbitrary order (Munshi 2000; Munshi & Jain 2000). These results are applicable in real-space which make them useful for surveys with low-sky coverage. The results derived here will be relevant for surveys with high sky-coverage where harmonics decomposition would mean less correlated measurements for individual ℓ.

By their very nature, projected or 2D surveys unlike their 3D counterparts mixes scale which makes assigning exact spectral index with an angular scales or in our case the harmonic ℓ. We have shown how much variation we should expect for a range of feasible spectral index n. Finally, the redshift dependence of the skew- and kurt-spectra is encoded in the coefficients R_2 and R_3. It is however important to point out that these pre-factors are rather sensitive to the lower limit of the integration $z_{\rm min}$ i.e. in equations (27b) and (27b). Numerical implementation of simulation of ray tracing to generate convergence maps may introduce slight modification in $z_{\rm min}$ which may lead to a bias in the theoretical predictions.

The excellent match between the theoretical predictions and simulations we have found here is encouraging for computing such corrections.

8 MODIFIED THEORIES OF GRAVITY: COMPUTATION OF C_{21}^{κ}

The theoretical modelling of the bispectrum in modified gravity scenarios is more challenging than the power spectrum calculation. Typically a perturbative approach is adopted in the quasi-linear regime (Bernardeau & Brax 2011). In addition, a quasi-static approximation is used, i.e. metric perturbations are varying slowly with time that they can be ignored. Many extensions of the perturbative approach were considered in the literature in recent years (Bose & Taruya 2018; Namikawa, Bouchet & Taruya 2018). Typically, this is achieved by introducing more freedom to the kernels and validating or calibrating them using numerical simulations. Indeed, others including variants of halo model predictions too have been proposed that can reproduce the simulation results with varying degree of success.

In the literature, typically, two main families of modified gravity theories are considered. (A) models with Vainshtein-screening mechanism which include the DGP model as well as the Horndeski (Hordenski 1974) and beyond Horndeski theories (Gleyzes et al. 2015a, b; Langois & Noui 2016b) and (B) the models with Chameleon-screening that includes the Hu–Sawicki f(R) model (Hu & Sawicki 2007). In the DGP model (Dvali, Gabadadze & Porati 2000), the bispectrum from simulations can be reproduced using the GR expression by suitably modifying the power spectrum. The situation is somewhat more complicated for $f(R)$ theories. The numerical modelling is more important at small scales where analytical results start to fail.

8.1 Bernardeau and Brax models

Next, we first turn to different phenomenological toy models of modified gravity presented by Bernardeau & Brax (2011).

(i) **(a) Gamma γ model:** This model is generated by modifying the Euler equation of the Euler-Continuity-Poisson equation. In this model, the gravitational field seen by massive particles (denoted as $\phi^{\rm eff}$) is different from the gravitational potential that solved the Poisson equation ϕ. These two potentials are different and related by $\phi^{\rm eff} = (1+\epsilon)\phi$ through parameter $\epsilon(t)$ in the sub-horizon scale.
In this parametrization, the kernel F_2 in equation (1b) is modified to the following form:

$$F_2(\mathbf{k}_1, \mathbf{k}_2) = \frac{1}{2}(1+\epsilon) + \frac{1}{2}\frac{\mathbf{k}_1\cdot\mathbf{k}_2}{k_1k_2}\left(\frac{k_1}{k_2}+\frac{k_2}{k_1}\right) + \frac{1}{2}(1-\epsilon)\left[\frac{\mathbf{k}_1\cdot\mathbf{k}_2}{k_1k_2}\right]^2 \tag{42}$$

In general the parameter ϵ can be a function of scale factor a or the wavelength k. For $\epsilon = 3/7$ recover the expression given in equation (1c). The Lagrangian perturbation theory is often used to model quasi-linear evolution of gravitational clustering. The Zel'dovich approximation is the linear order in Lagrangian perturbation theory. The bispectrum in the Zel'dovich approximation can be recovered from equation (42) $\epsilon = 0$ (Munshi, Sahni & Starobinsky 1994)

$$\langle F_2\rangle_{\rm 3D} = \frac{\epsilon+2}{3}; \quad \langle F_2\rangle_{\rm 2D} = \frac{\epsilon+3}{4}. \tag{43}$$

The actual value of the parameter ϵ can be computed using the linearized Euler-Continuity-Poisson equation, and assuming a parametric form for the growth factor $f = \mathrm{d}\ln D_+/\mathrm{d}\ln a \approx \Omega_{\rm M}^{\gamma}$. A convenient form of a fitting function can be obtained for values not too far from General Relativistic (ΛCDM) values. This model can be considered as a special case of equation (66b) with $\kappa = 1$ and $\epsilon = 1 - 4/7\lambda$. The smoothing includes a dependence on spectral index. In 2D we have

$$C_{21}^{\kappa} = R_2\left[4\langle F_2\rangle_{\rm 2D} - \frac{1}{2}(n+2)\right]. \tag{44}$$

(ii) **(b) Beta (β) model:** In the β model proposed by Bernardeau & Brax (2011) where the expression for the kernel $F_2(\mathbf{k}_1, \mathbf{k}_2)$ we have

$$F_2(\mathbf{k}_1,\mathbf{k}_2) = \left(\frac{3\nu_{\rm s}}{4} - \frac{1}{2}\right) + \frac{1}{2}\frac{\mathbf{k}_1\cdot\mathbf{k}_2}{k_1k_2}\left[\frac{k_1}{k_2} + \frac{k_2}{k_1}\right] + \left(\frac{3}{2} - \frac{3\nu_{\rm s}}{2}\right)\left[\frac{\mathbf{k}_1\cdot\mathbf{k}_2}{k_1k_2}\right]^2, \tag{45}$$

where the parameter ν_2 can be related to the ϵ parameter in equation (42) $\epsilon = \frac{3}{2}\nu_{\rm s} - 2$. The parametric value for ν_2 can be obtained in a manner similar to the γ model. However, we would leave them unspecified. The angular average gives $\langle F_2\rangle_{\rm 3D} = \nu_{\rm s}/2$ and similarly $\langle F_2\rangle_{\rm 2D} = (3\nu_{\rm s}/2 + 1)/4$ for 2D and is independent of t. In these models, the ν_2 can in general be a function of z as well as wavenumber k. This model was also recently used in Munshi (2017) for computation of a related statistics known as integrated bispectrum. In general, the parameter can also be a k-dependent parameter. The expression for C_{21} has the following form:

$$C_{21}^{\kappa} = R_2\left[\frac{3}{2}\nu_{\rm s} + 1 - \frac{1}{2}(n+2)\right]. \tag{46}$$

The power spectrum too gets modified due to changes in the kernel $F_2(\mathbf{k}_1, \mathbf{k}_2)$ at one loop. The loop corrections to the linear power spectrum depends on the $F_2(\mathbf{k}_1, \mathbf{k}_2)$ and thus can also be used to constrain any departure from GR.

8.2 Normal-branch of Dvali, Gabadadze, Porrati (nDGP) model

The normal branch of Dvali et al. (2000) model also known as the nDGP is a prototypical example that involve Vainshtein screening. The model of bispectrum that is known to accurately reproduce the bispectrum was computed by Koyama, Taruya & Hiramatsu (2009) which correspond to the case $\kappa = 1$ in equation (66b)

$$\kappa_{\rm s}(z) = 1; \qquad \lambda_{\rm s}(z) = \left(1 - \frac{7}{2}\frac{D_2(z)}{D_+^2(z)}\right). \tag{47}$$

Here, $D_2(z)$ and $D_+(z)$ are the first-order and second-order growth factors that can be computed by numerically solving the equations governing growth of perturbations (Bose & Taruya 2018).

We would like to point out here that the nDGP model, Gamma γ model, and the Beta (β) model considered here are all special cases of the general scalar tensor theories we will consider later.

8.3 Horndeski and beyond Horndeski in the perturbative regime

Horndeski theories are scalar–tensor theories with a single propagating degree of freedom and are free from Ostrogradsky-type instabilities. The Horndeski theories have also been extended by considering what are also known as the degenerate higher order scalar–tensor (DHOST) theories. The simplest extensions in the context of non-degenerate scenarios are also known as the Galeyzes-Langlois–Piazza–Venizzi or GPLV theories (Gleyzes et al. 2015a, b). The second-order kernel in these scenario include a scale-dependent additional term that changes the bispectrum (Hirano et al. 2018) that can be constrained using the staistics discussed here

$$F_2(\mathbf{k}_1,\mathbf{k}_2,z) = \kappa_{\rm s}(z)\alpha_{\rm s}(\mathbf{k}_1,\mathbf{k}_2) - \frac{2}{7}\lambda_{\rm s}(z)\gamma_{\rm s}(\mathbf{k}_1,\mathbf{k}_2); \tag{48a}$$

$$\alpha_{\rm s}(\mathbf{k}_1,\mathbf{k}_2) = 1 + \frac{1}{2}(\mathbf{k}_1\cdot\mathbf{k}_2)\frac{(k_1^2+k_2^2)}{k_1^2k_2^2}; \quad \gamma_{\rm s}(\mathbf{k}_1,\mathbf{k}_2) = 1 - \frac{(\mathbf{k}_1\cdot\mathbf{k}_2)^2}{k_1^2k_2^2}. \tag{48b}$$

Throughout in this paper $\langle.\rangle_{\rm 2D}$ and $\langle.\rangle_{\rm 3D}$ will denote the angular averages in 2D and 3D. By θ we will denote the angle between the wave vectors $\mathbf{k}_1$ and $\mathbf{k}_2$ in 3D, i.e. $\mathbf{k}_1\cdot\mathbf{k}_2 = k_1k_2\cos\theta$. Similarly in 2D it will denote the angle between the projected components, i.e. $\mathbf{k}_{1\perp}$ and $\mathbf{k}_{2\perp}$. So we can write $\mathbf{k}_{1\perp}\cdot\mathbf{k}_{2\perp} = k_{1\perp}k_{2\perp}\cos\theta$. The angular average in equation (48b) then reduces to averaging $\cos\theta$ and $\cos^2\theta$ in 2D (circle) or in 3D (sphere). Notice that both in 2D as well as in 3D $\langle\cos\theta\rangle = 0$. On the other hand $\langle\cos^2\theta\rangle$ takes the value of 1/2 in 2D and 1/3 in 3D. Using these results we can see that in 2D $\langle\alpha_{\rm s}(\mathbf{k}_1,\mathbf{k}_2)\rangle_{\rm 2D} = 1$ and $\langle\gamma_{\rm s}(\mathbf{k}_1,\mathbf{k}_2)\rangle_{\rm 2D} = 1/2$ which leads us, respectively, in 2D to

$$\langle F_2\rangle_{\rm 2D} = \kappa_{\rm s}(z) - \frac{1}{7}\lambda_{\rm s}(z); \quad \langle F_2\rangle_{\rm 3D} = \kappa_{\rm s}(z) - \frac{4}{21}\lambda_{\rm s}(z). \tag{49}$$

Similar calculation in the effective field theory (EFT) of dark energy framework can be found in (Cusina, Lewandowskyi & Vernizzi 2018)

$$C_{21}^{\kappa} = 2\int_0^{r_{\rm s}} \mathrm{d}r\, D_+^4(z)\frac{w^3(r)}{\mathrm{d}_A^{4+2n}(r)}\left[\kappa_{\rm s}(z) - \frac{1}{7}\lambda_{\rm s}(z) - \frac{1}{2}\kappa_{\rm s}(z)(n+2)\right] \Big/ \left(\int_0^{r_{\rm s}} \mathrm{d}r\, D_+^2(z)\frac{w^2(r)}{\mathrm{d}_A^{2+n}(r)}\right)^2. \tag{50}$$

In Crisostomi, Lewandoski & Vernizzi (2020), the following equivalent parametrization for the kernel F_2 was introduced:

$$F_2(\mathbf{k}_1, \mathbf{k}_2) = A_\alpha(z)\alpha_s(\mathbf{k}_1, \mathbf{k}_2) + A_\gamma(z)\gamma_s(\mathbf{k}_1, \mathbf{k}_2). \tag{51}$$

In terms of the parameters $A_\alpha(z)$ and $A_\gamma(z)$

$$C_{21}^{\kappa} = 2\int_0^{r_s} \mathrm{d}r\, D_+^4(z)\frac{w^3(r)}{\mathrm{d}_A^{4+2n}(r)}\left[A_\alpha(z) + \frac{1}{2}A_\gamma(z) - \frac{1}{2}A_\alpha(z)(n+2)\right] \Big/ \left(\int_0^{r_s} \mathrm{d}\, r\, D_+^2(z)\frac{w^2(r)}{\mathrm{d}_A^{2+n}(r)}\right)^2. \tag{52}$$

In general, the parameters $A_\alpha(z) = \kappa_s(z)$, $A_\gamma(z) = -2/7\lambda_s(z)$ are time dependent. For this model, we have $\langle F_2\rangle_{3\mathrm{D}} = A_\alpha + \frac{2}{3}A_\gamma$ and $\langle F_2\rangle_{2\mathrm{D}} = A_\alpha + \frac{1}{2}A_\gamma$. It is important to notice that these theories have an important difference with GR and Horndeski theories. The Horndeski theories are invariant under time-dependent spatial co-ordinate transformations. The form for the F_2 kernels is fixed by existence of such symmetry. Many modified gravity theories fall under this category. In beyond Horndeski theories, the fluid equations and the equations of gravity possess very different symmetry properties and have a kernel F_2 that is structurally different. This is related to violation of these theories from the so-called consistency relation which are respected in GR (Peloso & Pietroni 2008). Future surveys such as the *Euclid* survey will be able to probe such theories beyond the consistency relations using the statistics developed here.

A detailed study for the skew-spectrum (Munshi & Heavens 2010) Minkowski functionals (Munshi et al. 2012) for these models and the integrated bispectrum (Munshi et al. 2020) as well as the related consistency relations will be presented elsewhere (Munshi et al. in preparation).

8.4 Massive neutrinos

A small but non-negligible fraction of the cosmological matter density is provided by massive neutrinos (Lesgourgues & Pastor 2006). The massive neutrinos are known to have significant thermal distribution and a different cosmological evolutionary history in comparison to the CDM. The thermal dispersion in velocity results in a damping of perturbation below a length-scale also known as the free-streaming length-scale. This will be probed by future surveys with a very high degree of accuracy. In the long run cosmological surveys are expected to provide an upper limit to the sum of the neutrino masses. This will be very useful when jointly considered with the lower limits from neutrino-oscillation experiments.

The neutrinos decouple and free-stream with a large thermal velocities. The redshift z_{nr} at which neutrinos become relativistic depend on their mass eigenstate m_i, $1 + z_{\mathrm{nr}} = 1980[m_{\nu,i}/1\,\mathrm{eV}]$ The fractional contribution to the total matter density is denoted as f_ν which can be expressed as

$$f_\nu \equiv \frac{\Omega_\nu}{\Omega_M} = \frac{1}{\Omega_{M,0}h^2}\frac{\sum_i M_{\nu,i}}{93.14\mathrm{eV}}. \tag{53}$$

In future it will also be interesting to consider the effect of neutrino mass on bispectrum when simulated all-sky lensing maps for such cosmologies will be available (Coulton et al. 2019; Liu et al. 2018). The total matter distribution thus can be written in terms of the CDM perturbation δ_{cdm} and the fluctuations in the neutrino density distribution δ_ν

$$\delta_{\mathrm{m}} = f_c\delta_c + f_\nu\delta_\nu; \qquad f_c + f_\nu = 1. \tag{54}$$

The resulting matter power spectrum $P_{\mathrm{mm}}(k)$ and bispectrum $B_{\mathrm{mmm}}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)$ can be expressed as (Ruggeri 2018)

$$P_{\mathrm{mm}}(k) = f_c^2 P_{cc}(k) + 2f_\nu f_c P_{\nu c}(k) + f_\nu^2 P_{\nu\nu}(k) \tag{55a}$$

$$B_{\mathrm{mmm}} = f_c^3 B_{ccc} + f_c^2 f_\nu B_{cc\nu} + f_c f_\nu^2 B_{\nu\nu c} + f_\nu^3 B_{\nu\nu\nu}. \tag{55b}$$

Here, P_{cc} and $P_{\nu\nu}$ represent the power spectrum CDM and the neutrino component where as the $P_{\nu c}$ is the cross spectra between them. We will drop the suffix 3D to avoid cluttering. We will only consider the linear order perturbation in δ_ν and ignore all higher order contributions which implies $B_{\nu\nu\nu} = 0$. For B_{ccc}, the expression in the squeezed limit is exactly same as derived before

$$B_{ccc}^{\mathrm{2D,sq}} = \left[\frac{24}{7} - \frac{1}{2}\frac{\mathrm{d}k^2\, P_{cc}(k)}{d\ln k}\right] P_{cc}(k_\perp)P_{cc}(q_{3\perp}). \tag{56a}$$

We will next consider the mixed terms $B_{\nu\nu c}$. These contributions in terms of δ_c and δ_ν can be expressed as

$$B_{cc\nu}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = \langle\delta_c(\mathbf{k}_1)\delta_c(\mathbf{k}_2)\delta_\nu(\mathbf{k}_3)\rangle + \mathrm{cyc.perm.}; \tag{57a}$$

$$B_{\nu\nu c}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = \langle\delta_\nu(\mathbf{k}_1)\delta_\nu(\mathbf{k}_2)\delta_c(\mathbf{k}_3)\rangle. + \mathrm{cyc.perm.}. \tag{57b}$$

In the above equations the cyc. perm. represents cyclic permutations of the wave vectors $\mathbf{k}_1$, $\mathbf{k}_2$, and $\mathbf{k}_3$.

To evaluate $B_{\nu\nu c}$, we expand the terms perturbatively. Employing tree level perturbation theory, the contributions from $B_{\nu\nu c,\,112}$ are from these terms

$$B_{\nu\nu c} = B_{\nu\nu c,112}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) + B_{\nu\nu c,112}(\mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_1) + B_{\nu\nu c,112}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3). \tag{58}$$

In our notation, $B_{vvc,112}(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3) \equiv \langle \delta_v^{(1)}(\mathbf{k}_1)\delta_v^{(1)}(\mathbf{k}_2)\delta_c^{(2)}(\mathbf{k}_3)\rangle$ and similarly for the other terms. In tems of the second-order kernels $F_2(\mathbf{k}_1, \mathbf{k}_2)$ we have

$$B_{vvc,112}(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3) = 2F_2(\mathbf{k}_1,\mathbf{k}_2)P_{vc}(k_1)P_{vc}(k_2)\,; \tag{59}$$

The other terms can be recovered by cyclic permutation of the wavenumber. In the squeezed limit we have

$$B^{\rm 2D,sq}_{vvc} = \left[\frac{24}{7} - \frac{1}{2}\frac{{\rm d}\ln k^2\, P_{vc}(k_\perp)}{d\ln k_\perp}\right] P_{vc}(k_\perp)P_{vc}(q_{3\perp}). \tag{60}$$

Finally, we turn to B_{ccv}. The perturbative contributions are as follows:

$$B^{\rm 2D,sq}_{vvc} = 2[F_2(\mathbf{k}_1,\mathbf{k}_2)P_{cc}(k_1)P_{cv}(k_2) + {\rm cyc.perm.}]. \tag{61}$$

Going through an elaborate algebraic manipulation we arrive at the squeezed limit

$$B^{\rm 2D,sq}_{vvc} = \left[\frac{24}{7} - \frac{1}{2}\frac{{\rm d}\ln k^2 P_{cc}(k_\perp)}{{\rm d}\ln k_\perp}\right] P_{cc}(k_\perp)P_{cc}(q_{3\perp}) + \left[\frac{24}{7} - \frac{1}{2}\frac{{\rm d}k_\perp^2 P_{cv}(k_\perp)}{{\rm d}\ln k_\perp}\right] P_{cv}(k_\perp)P_{cc}(q_{3\perp}). \tag{62}$$

In future it will be interesting to study the effect of neutrino mass on bispectrum using simulations when all-sky lensing maps for such cosmologies will be available (Coulton et al. 2019; Liu et al. 2018).

8.5 Clustering quintessence

Quintessence (Tsujikawa 2013) is the most popular dynamics of dark energy in which the potential energy of a single scalar field drives the accelerated expansion of the Universe. The quintessence model is different from the cosmological constant scenario allowing for a different temporal dependence of the observables. The scalar field in most quintessence models is considered homogeneous and is typically minimally coupled. The sound speed of the scalar field in these models equals the speed of light which prevents any clustering below the horizon scale. However, extensions of such models with vanishing or lower than speed of light have also been considered. These models are known as the clustering quintessence models (Sefusatti & Vernizzi 2011; Basse, Bjalde & Wong 2011). The future LSS surveys can be used to differentiate between these two scenarios. We use our formalism to derive the changes in the bispectrum in the squeezed limit in these models. We quote the expression of the kernel F_2 from (Sefusatti & Vernizzi 2011)

$$\frac{D_+}{a} = \frac{5}{2}\Omega_{\rm M}\left[{\Omega_{\rm M}}^{4/7} + \frac{3}{2}\Omega_{\rm M} + \left(\frac{1}{70} - \frac{1+w}{4}\right)\Omega_{\rm Q}\left(1+\frac{\Omega_{\rm M}}{2}\right)\right]^{-1}. \tag{63a}$$

Here, $\Omega_{\rm Q}$ and $\Omega_{\rm M}$ are the density parameter related to Quintessence and dark matter. The corresponding linear growth rates are denoted by D_{Q+} and D_+. The parameters $\epsilon_s = \frac{\Omega_Q}{\Omega_M}\frac{D_{Q,+}}{D_+}$ and ν_2 can also be expressed in terms of $\Omega_{\rm Q}$ and $\Omega_{\rm M}$ and depend of redshift z

$$F_2(\mathbf{k}_1,\mathbf{k}_2,\eta) = \frac{\nu_{\rm s}}{2} + \frac{1}{2}(1-\epsilon_{\rm s})\frac{\mathbf{k}_1\cdot\mathbf{k}_2}{k_1k_2}\left(\frac{k_1}{k_2}+\frac{k_2}{k_1}\right) - \frac{1}{2}\left(1-\epsilon_{\rm s}-\frac{\nu_{\rm s}}{2}\right)\left[1-3\left(\frac{\mathbf{k}_1\cdot\mathbf{k}_2}{k_1k_2}\right)^2\right]. \tag{64a}$$

Thus, two different parameters $\epsilon_{\rm s}(z)$ and $\nu_{\rm s}(z)$ to describe the tree-level bispectrum in this model

$$C_{21}^{\kappa} = \int_0^{r_{\rm s}} {\rm d}r\,\frac{w^3(r)}{{\rm d}^{4+2n}(r)}D_+^4(z)\left[\frac{1}{4}(1-\epsilon_{\rm s}) + \frac{3}{8}\nu_{\rm s} - \frac{1}{2}(1-\epsilon_{\rm s})(n+2)\right] \Big/ \left(\int_0^{r_{\rm s}} {\rm d}r\,\frac{w^2(r)}{{\rm d}^{2+n}(r)}D_+^2(z)\right)^2. \tag{65a}$$

Typically at low redshift for some values of w the parameter ϵ can reach upto 10 per cent which can lead to roughly an order of 10 per cent correction to the bispectrum which can be accounted for high-precision measurements from future surveys.

8.6 Bispectrum in general scalar–tensor theories

Next, we consider a phenomenological fitting function. The second-order perturbative analysis of the general scalar tensor theories were initially performed by Hirano et al. (2018) which was later extended to smaller non-perturbative scales using a fitting function (Namikawa et al. 2018). Using the fitting function proposed in Namikawa et al. (2018), we can compute the C_{21}^{κ} in a class of models which are represented by the following expression for $F_2(\mathbf{k}_1,\mathbf{k}_2,z)$ replacing $F_2(\mathbf{k}_1,\mathbf{k}_2)$ in equation (1c):

$$F_2(\mathbf{k}_1,\mathbf{k}_2,z) = \left[\kappa_{\rm s}(z) - \frac{2}{7}\lambda_{\rm s}(z)\right]a(k_1,z)a(k_2,z) + \frac{1}{2}\kappa_{\rm s}(z)\left[\frac{\mathbf{k}_1\cdot\mathbf{k}_2}{k_1k_2}\right]\left(\frac{k_1}{k_2}+\frac{k_2}{k_1}\right)b(k_1,z)b(k_2,z)$$
$$+\frac{2}{7}\lambda_{\rm s}(z)\left[\frac{\mathbf{k}_1\cdot\mathbf{k}_2}{k_1k_2}\right]^2 c(k_1,z)c(k_2,z); \tag{66a}$$

$$\lambda_{\rm s}(z) = [\Omega_{\rm M}(z)]^{\xi_\lambda}; \quad \kappa_{\rm s}(z) = [\Omega_{\rm M}(z)]^{\xi_\kappa}; \quad \Omega_{\rm M}(z) = \Omega_{\rm M,0}(1+z)^3/((1+z)^3\Omega_{\rm M,0} + \Omega_\Lambda). \tag{66b}$$

The functions $\kappa_{\rm s}(z)$ and $\lambda_c(z)$ are approximated using the above functional forms and ξ_λ and are free parameters that can be estimated from observational data. The functional forms for a, b, and c are assumed to be same as that of their ΛCDM form (Scoccimarro & Couchman

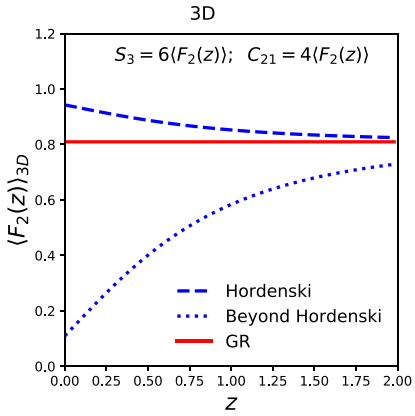

Figure 4. The second-order tree-level perturbative vertex $\langle F_2(z)\rangle_{3D}$ is plotted as a function of redshift z for 3D surveys as given in equation (49). Three different cases shown correspond to the Horndeski, beyond Horndeski, and GR as indicated. See text for more details. The results are shown for unsmoothed field i.e. $n = -3$. We have used the parametrizations in equation (66b) for various models. The Horndeski model is given by $\xi_\kappa = 1$, $\xi_\lambda = 0$ and the beyond Horndeski theories are given by $\xi_\kappa = 1$, $\xi_\lambda = 1$. For the GR we have $\xi_\kappa = 0$, $\xi_\lambda = 0$.

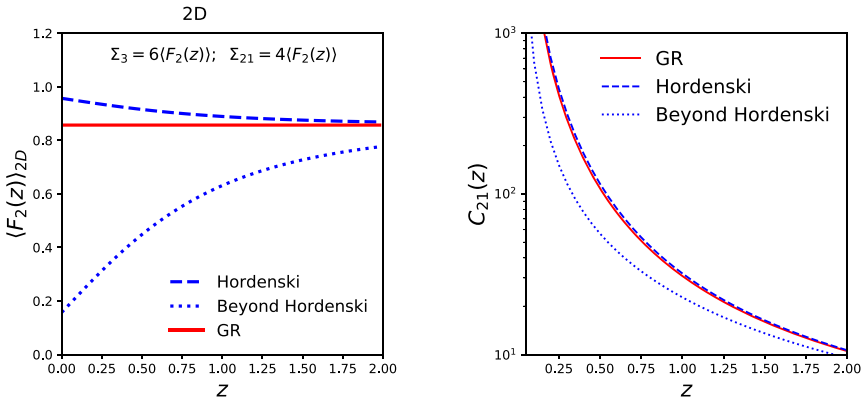

Figure 5. The left-hand panel shows $\langle F_2(z)\rangle_{2D}$ for GR, Horndeski and beyond Horndeski as a function of redshift z in 2D. The results are plotted for $n = -2$ which represents the unsmoothed field. The right-hand panel corresponds to $C_{21}(z)$ as a function for these models.

2001; Gil-Marin et al. 2012) which interpolates the perturbative regime and highly non-linear assumed to be described by Hyper-Extended-Perturbation-Theory (Scoccimarro & Frieman 1999). In the perturbative regime that we are interested we have $a = b = c = 1$. To be consistent with the literature we have used κ_s to denote one of the parameters which should not be confused with the weak lensing convergence κ as their meaning would be obvious from the context. For $\kappa_s(a) = 1$ and $\lambda_s(a) = 1$ or, equivalently, $\xi_\kappa = 0$ and $\xi_\lambda = 0$ we recover the case of GR presented in equation (1c). As discussed before, the Horndeski theories are the most general class of scalar–tensor theories which are non-degenerate that leads second-order equations of motion in 4D. In these models, $\lambda_s \neq 1$ though $\kappa = 1$ still remains valid. A generalization of Horndeski (Hordenski 1974) theory leads to a class of models that are known as 'beyond Horndeski' models (Gleyzes et al. 2015a, b; Langois & Noui 2016b). In these models both κ_s and λ_s can deviate from unity. At high-z the both theories converge to GR as expected. The Horndeski theories violate the Vaishentein mechanism to recover GR at non-linear scale has also been considered. In these scenarios, the parameter both κ and λ deviates from unity. Thus testing GR that correspond to $\lambda = \kappa = 1$ reduces to constrain deviation of λ and κ from unity. The functional form for κ and λ is adopted from Namikawa et al. (2018) and converges to GR at high-z as expected.

We will next focus on computing the second-order vertex ν_2 as defined in equation (21) for both 3D and 3D. Unlike in case of GR, in general these vertices have a redshift dependence. To compute these we start by noticing that in both 3D and 2D we have $\langle \mathbf{k}_1 \cdot \mathbf{k}_2/k_1k_2\rangle = 0$ and in 2D we have $\langle(\mathbf{k}_1 \cdot \mathbf{k}_2/k_1k_2)^2\rangle = 1/2$. In the following, we will ignore smoothing as the correction terms involved will be exactly same as the one presented.

In the quasi-linear regime the functions a, b, and c tend to unity. In this limiting case, the departure from GR is encoded only in the redshift-dependent factors and the expression for C_{21} is identical to equation (50) with the specific form for κ_s and λ_s are given by equation (66b). Substituting $\kappa_s(z) = 1$ and $\lambda_s(z) = 1$ we recover the unsmoothed results for GR. The smoothing in 3D and 2D will introduce terms involving factors of $(n + 3)$ in equations (40a) and (40b) and $(n + 2)$ in equations (26a) and (27b). The results for specific models for 3D and 2D are, respectively, shown in Figs 4 and 5. While for GR the $\langle F_2\rangle$ is independent of redshift z, for Horndeski and beyond Horndeski theories $\langle F_2\rangle$ depends on redshift. At higher z they become identical to that of GR as expected. In Fig. 4, the $\langle F_2\rangle$ for the 2D cylindrical collapse is

plotted as function of z and their pattern of evolution is same as in 3D. The effect of line-of-sight projection is encoded in the factor $R_2(z)$ which is shown in right-hand panel.

Although, the results for higher order spectra are known to an arbitrary order in GR, similar results for most of the modified gravity theories are known mostly to second order. Going beyond third-order in general requires order-by-order calculation. While we have considered the statistics of 3D density field δ and resulting convergence κ similar results can be obtained for the divergence of peculiar velocity.

The tests involving bispectrum related statistics presenetd here can further tighten the constraints obtained using linear growth rate alone. This is particularly important as no strong constraint on λ_s and κ_s exist currently. Indeed, there are no upper or lower limits for κ_s based on theoretical expectation.

Before we conclude this section, we would like to point out that the two parameters used in defining the clustering quintessence i.e. ν_s and ϵ_s (or α_s and β_s for the case of DHOST theories) can independently be constrained using 3D and 2D measurements. This is due to the fact that the statistics C_{21} depends on ν_s and ϵ in a different manner in 3D and 2D. We have concentrated on projected or 2D surveys in this paper but similar results will be presented for 3D surveys in a separate article.

9 CONCLUSIONS AND FUTURE PROSPECTS

We have computed the skew-spectrum (see equation 10a) and kurtosis-spectrum equation (15a) at low ℓ for the analysis of weak lensing convergence or κ maps. These spectra generalizes the one-point cumulants, e.g. the skewness and kurtosis defined in Section 2.3, and are often used in the literature for analysing higher order non-Gaussianity of cosmological maps. They capture some of the essential properties of the full bispectrum or trispectrum which are more difficult to estimate. In the real space, these spectra correspond to cumulant correlators that can be computed in the leading-order using tree-level perturbations in the large-separation limit. In this limit, these spectra can be computed to arbitrary order using tree-level perturbative calculations without any need for any phenomenological fitting functions or extensions of perturbative calculation. We use the flat-sky approximation and Eulerian perturbative results based on generating function approach we show how to compute high-order spectra to arbitrary order. We test these results for lower order spectra namely the skew- and kurt-spectra against state-of-the-art all-sky weak lensing maps. We find the our results are in good agreement. These results will be valuable in analysing higher order statistics from future all-sky weak lensing surveys such as the *Euclid* survey. The presence of mask generated from near all-sky surveys introduces mode mixing. Unless corrected, the mode mixing introduced by a mask can be a source of confusion while analysing the higher order spectra as they encode information about gravity induced mode coupling. We have presented a generalization of existing method typically used in the study of ordinary power spectrum to construct an unbiased estimates of higher order spectra equations (41a) and (41b).

The parameters C_{pq} computed for 3D weak lensing will be important when photo-z information is available. The statistics introduced here will be useful in analysing non-Gaussianity in such context. We will present results of such analysis in future work. The results presented here can be generalized using a 3D Fourier–Bessel transform or a 3D flat sky formalism. As noted before the 3D analysis allows factorization $C_{pq} = C_{pq}C_{q1}$ and their dependence on the spectral index n are different so 2D and 3D results will provide independent information as well as much needed consistency checks and test for possible systematics.

Any modification of gravity leaves detectable signature at the level of bispectrum. Though such signatures are less prominent than any modification at the level of power spectrum, it has recently attracted a lot of attention in the context of CMB lensing bispectrum (Namikawa et al. 2018). Similar investigations in the context of weak lensing are currently being pursued using various statistical estimators. Various techniques were adopted to extend perturbative results derived in the context of GR. Extensions to modified gravity scenarios were implemented by introducing more freedom to the kernels and calibrating then using numerical simulations (Bose & Taruya 2018). The expressions for bispectrum exist for both type of modified gravity scenarios i.e. models with Vainshtein-screening mechanism which includes the DGP model as well as the Horndeski (Hordenski 1974) and beyond Horndeski theories (Gleyzes et al. 2015a, b; Langois & Noui 2016b). In the other class of models i.e. models with Chameleon-screening that includes the Hu–Sawicki f(R) model (Hu & Sawicki 2007), the bispectrum from simulations can be successfully reproduced using the GR expression but with suitable modification of the power spectrum. We will extend our results derived here to the modifying gravity scenarios as well as scenarios involving massive neutrinos.

The position-dependent bispectrum and its higher order generalization at the level of trispectrum has exact one-to-one correspondence with the statistics studied in this paper. Indeed the expressions for integrated bispectrum and the skew-spectrum at low-ℓ are identical. However, the physical interpretation is different. The expressions at the level of fourth order are not the same. The integrated bispectrum or equivalently the position-dependent power spectrum probes the influence of large-scale modes on small-scale structure. The cumulant correlators at large separation limit as well as their harmonic counterparts namely the skew-spectrum and kurt-spectra probe dynamics mainly at scales of smoothing. Comparing results from these two statistics can provide useful cross-checks at each order.

Finite sky coverage can introduce bias in our estimators. The scatter and bias introduced by finite survey size have been studied in great detail for galaxy surveys and to a lesser extent for weak lensing surveys (Munshi & Coles 2003). These are less dominant in the quasi-linear regime where the variance is small in the limiting case which we have studied here.

In our study, we have assumed that the bispectrum is of even parity. Many studies in the recent past have pointed out existence of an off-parity bispectrum (Munshi et al. 2013). Such a bispectrum do not arise from 3D density perturbations. However, signatures of contributions can be used to test possible existence of systematics.

In a recent work (Barthelemy 2020), it was shown that *nulling* can be used effectively to improve the accuracy of perturbative calculations by reducing the cross-talk between quasi-linear and non-linear scales. These calculations were performed in the real-space focusing primarily

on one-point cumulants and PDF. In contrast our results here concern primarily on two-point correlators and their associated spectra in the Fourier domain. Applying the nulling before computing the spectra is expected to improve the validity of the perturbative results.

Last but not least, the next generation of CMB Stage-IV experiments will be able to map the projected lensing potential all the way to the surface of last scattering. It is expected that the results obtained in this paper will be valuable in analysing higher order statistics of maps obtained from such experiments (Abajazajian et al. 2018). However, in this case the estimator described here will have to be optimized to tackle low signal-to-noise ratio for higher order statistics of CMBR. The post-born corrections (Lewis & Pratten 2016) play an important role in higher order statistics of CMBR. For realistic comparison against observations such corrections should be included.

ACKNOWLEDGEMENTS

DM is supported by a grant from the Leverhulme Trust at MSSL. It is a pleasure for DM to thank F. Bouchet, T. D. Kitching, T. Namikawa, R. Takahashi, A. Taruya, and F. Vernizzi for many useful discussions. We would like to also thank R. Takahashi for making the lensing maps publicly available. We would like to also thank R. Schoenrich for careful reading of the draft and many suggestions that greatly improved the presentation. DM would also like to organizers of the *Euclid* Theory Working Group Meeting (2019 8 April to 9 April) in Oxford.

DATA AVAILABILITY

No new data were generated in support of this research. The simulations used in this work are publicly available.

REFERENCES

Abajazajian K. et al., 2018, preprint(arXiv:1907.04437)
Abbott T. et al., 2016, Phys. Rev. D, 94, 022001
Abbott B. P. et al., 2017, Phys. Rev. Lett., 119, 161101
Amendola L. et al., 2013, Living Rev. Relativ., 16, 6
Asgari M., Taylor A., Joachimi B., Kitching T. D., 2018, MNRAS, 479, 454
Baker T., Bellini E., Ferreira P. G., Lagos M., Noller J., Sawicki I., 2017, Phys. Rev. Lett., 119, 251301
Balian R., Schaeffer R, 1989, A&A, 220, 1
Barthelemy A., Codis S., Uhlemann C., Bernardeau F., Gavazzi R., 2020, MNRAS, 492, 3420
Bartolo N., Komatsu E., Matarrese S., Riotto A., 2004, Phys. Rep., 402, 103
Basse T., Bjalde O. E., Wong Y. Y. Y., 2011, JCAP, 10, 038
Bernardeau F., 1992, ApJ, 392, 1
Bernardeau F., 1994b, A&A, 291, 697
Bernardeau F., 1995, A&A, 301, 309
Bernardeau F., 1996a, A&A, 312, 11
Bernardeau F., 1996b, A&A, 312, 11
Bernardeau F., Brax P., 2011, J. Cosmol. Astropart. Phys., 1106, 019
Bernardeau F., Reimberg P., 2016, Phys. Rev. D, 94, 063520
Bernardeau F., Colombi S., Gaztanaga E., Scoccimarro R., 2002a, Phys. Rep., 367, 1
Bernardeau F., Mellier Y., van Waerbeke L., 2002b, A&A, 389, L28
Bernardeau F., van Waerbeke L., Mellier Y., 2003, A&A, 397, 405
Bose B., Taruya A., 2018, J. Cosmol. Astropart. Phys., 1810, 19
Calabrese E., Smidt J., Amblard A., Cooray A., Melchiorri A., Serra P., Heavens A., Munshi D., 2010, Phys. Rev. D, 81, 3529
Castro P. G., Heavens A. F., Kitching T. D., 2005, Phys. Rev. D, 72, 023516
Clifton T., Ferreira P. G., Padilla A., Skordis C., 2012, Phys. Rep., 513, 1
Cooray A., Sheth R., 2002, Phys. Rep., 372, 1
Coulton W. R., Liu J., Madhavacheril M. S., B'ohm V., Spergel D. N., 2019, JCAP, 05, 043
Creminelli P., Vernizzi F., 2017, Phys. Rev. Lett., 119, 251302
Crisostomi M., Lewandowski M., Vernizzi F., 2020, Phys. Rev. D, 101, 123501
Cusina G., Lewandowski M., Vernizzi F., 2018, J. Cosmol. Astropart. Phys., 04, 005
Deffayet C., Gao X., Steer D. A., Zahariade G., 2011, Phys. Rev. D, 84, 064039
Dvali G., Gabadadze G., Porrati M., 2000, Phys. Rev. B, 485, 208
Efstathiou G., 2004, MNRAS, 349, 603
Eggemeier A., Smith R. E., 2017, MNRAS, 486, 2496
Eisenstein D. J. et al., 2011, AJ, 142, 72
Gaztanaga E., Bernardeau F., 1998, A&A, 331, 829
Gil-Marin H., Wagner C., Fragkoudi F., Jimenez R., Verde L., 2012, JCAP, 02, 047
Gleyzes J., Langlois D., Piazza F., Vernizzi F., 2015a, J. Cosmol. Astropart. Phys., 2, 018
Gleyzes J., Langlois D., Piazza G., Vernizzi F., 2015b, Phys. Rev. Lett., 114, 211101
Goldstein A. et al., 2017, ApJ, 848, L14
Gorski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M., Bartelman M., 2005, ApJ, 622, 759
Harrison I., Coles P., 2011, MNRAS, 418, L20
Heavens A. F., Gupta S., 2001, MNRAS, 324, 960
Hikage C., Takada M., Hamana T., Spergel D., 2011, MNRAS, 412, 65
Hirano S., Kobayashi T., Tashiro H., Yokoyama S., 2018, Phys. Rev. D, 97, 103517

Hivon E., Gorski K. M., Netterfield C. B., Crill B. P., Prunet S., Hansen F., 2002, ApJ, 567, 2
Horndeski G. W., 1974, Int. J. Theor. Phys., 10, 363
Hu W., 2001, Phys. Rev. D, 64, 083005
Hu W., Sawicki I., 2007, Phys. Rev. D, 76, 064004
Joyce A., Jain B., Khoury J., Trodden M., 2015, Phys. Rep., 568, 1
Jurek R. J. et al., 2010, MNRAS, 401, 14
Kobayashi T., Yamaguchi M., Yokoyama J., 2011, Prog. Theor. Phys., 126, 511
Koyama K., Taruya A., Hiramatsu T., 2009, Phys. Rev. D, 79, 123512
Kuijken K. et al., 2015, MNRAS, 454, 3500
Langlois D., Noui K., 2016a, J. Cosmol. Astropart. Phys., 7, 16
Langlois D., Noui K., 2016b, J. Cosmol. Astropart. Phys., 02, 034
Laureijs R. et al., 2011, ESA/SRE(2011)12preprint (arXiv: 1110.3193)
Lesgourgues J., Pastor S., 2006, Phys. Rep., 429, 307
Liu J., Bird S., Matilla J. M. Z., Hill J. C., Haiman Z., Madhavacheril M. S., Spergel D. N., Petri A., Spergel D. N., 2018, J. Cosmol. Astropart. Phys., 03, 049
Lombriser L, Lima N. A., 2017, Phys. Lett. B, 765, 382
Matsubara T., 2003, ApJ, 584, 1
Matsubara T., 2007, ApJS, 170, 1
Munshi D., 2000, MNRAS, 318, 145
Munshi D., 2017, J. Cosmol. Astropart. Phys., 01, 049
Munshi D., Coles P., 2003, MNRAS, 338, 846
Munshi D., Heavens A., 2010, MNRAS, 401, 2406
Munshi D., Jain B., 2000, MNRAS, 318, 109
Munshi D., Sahni V., Starobinsky A. A., 1994, ApJ, 436, 517
Munshi D., Coles P., Melott A. L., 1999a, MNRAS, 310, 892
Munshi D., Bernardeau F., Melott A. L., Schaeffer R., 1999b, MNRAS, 303, 433
Munshi D., Valageas P., Van Waerbeke L., Heavens A., 2008, Phys. Rep., 462, 67
Munshi D., Heavens A., Cooray A., Smidt J., Coles P., Serra P., 2011a, MNRAS, 412, 1993
Munshi D., Kitching T., Heavens A., Coles P., 2011b, MNRAS, 416, 629
Munshi D., Valageas P., Cooray A., Heavens A., 2011c, MNRAS, 414, 3173
Munshi D., Coles P., Cooray A., Heavens A., Smidt J., 2011, MNRAS, 410, 1295
Munshi D., van Waerbeke L., Smidt J., Coles P., 2012, MNRAS, 419, 536
Munshi D., Smidt J., Cooray A., Renzi A., Heavens A., Coles P., 2013, MNRAS, 434, 2830
Munshi D., McEwen J. D., Kitching T., Fosalba P., Teyssier R., Stadel J., 2020, JCAP, 05, 043
Munshi D., Namikawa T., Kitching T. D., McEwen J. D., Takahashi R., Bouchet F. R., Taruya A., Bose B., 2020, MNRAS, 493, 3985
Namikawa T., Bouchet F. R., Taruya A., 2018, Phys. Rev. D, 98, 043530
Namikawa T., Bose B., Bouchet F. R., Takahashi R., Taruya A., 2019, Phys. Rev. D, 99, 063511
Oh S. P., Spergel D. N., Hinshaw G., 1990, ApJ, 510, 551
Okamoto T., Hu W., 2002, Phys. Rev. D, 66, 063008
Peel A., Lin C. A., Lanusse F., Leonard A., Starck J. L., Kilbinger M., 2017, A&A, 599, 79
Peloso M., Pietroni M., 2014, J. Cosmol. Astropart. Phys., 04, 011
Planck Collaboration I, 2020, A&A, 641, A1
Planck Collaboration XIII, 2016a, A&A, 594, A13
Planck Collaboration XVI, 2014, A&A, 571, A16
Planck collaboration XVII, 2016b, A&A, 594, A17
Pratten G., Lewis A., 2016, J. Cosmol. Astropart. Phys., 08, 047
Reimberg P., Bernardeau F, 2018, Phys. Rev. D, 97, 023524
Riquelme M. A., Spergel D. N., 2007, ApJ, 661, 672
Ruggeri R., Castorina E., Carbone C., Sefusatti E., 2018, J. Cosmol. Astropart. Phys., 03, 003
Sakstein J., Jain B., 2017, Phys. Rev. Lett., 119, 251303
Scoccimarro R., Couchman H. M. P., 2001, MNRAS, 325, 4
Scoccimarro R., Frieman J. A., 1999, ApJ, 520, 35
Sefusatti E., Vernizzi F., 2011, J. Cosmol. Astropart. Phys. , 1103, 047
Smidt J., Amblard A., Byrnes C. T., Cooray A., Heavens A., Munshi D., 2010, Phys. Rev. D, 81, 123007
Szapudi I., Szalay A. S., 1999, ApJ, 515, L43
Szapudi I, Prunet S., Pogosyan D., Szalay A. S., Bond J. R., 2001, ApJ, 548, 115
Takahashi R., Hamana T., Shirasaki M., Namikawa T., Nishimichi T., Osato K., Shiroyama K., 2017, ApJ, 850, 24
Tsujikawa S., 2013, Class. Quantum Gravity, 30, 214003
Tyson J. A., Wittman D. M., Hennawi J. F., Spergel D. N., 2003, Nucl. Phys. B Proc. Suppl., 124, 21
Uhlemann C., Codis S., Pichon C., Bernardeau F., Reimberg P., 2016, MNRAS, 460, 1529
Uhlemann C., Pichon C., Codis S., L'Huillier B., Kim J., Bernardeau F., Park C., Prunet S., 2018, MNRAS, 477, 2772
Weiss A. J., Schneider A., Sgier R., Kacprzak T., Amara A., Refregier A., 2019, JCAP, 10, 011

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