A&A 508, 1193–1204 (2009)
DOI: 10.1051/0004-6361/200912906
© ESO 2009

Astronomy & Astrophysics

# Bispectrum covariance in the flat-sky limit

B. Joachimi[1,2], X. Shi[1], and P. Schneider[1]

[1] Argelander-Institut für Astronomie (AIfA), Universität Bonn, Auf dem Hügel 71, 53121 Bonn, Germany
e-mail: joachimi@astro.uni-bonn.de
[2] Department of Physics and Astronomy, University College London, London WC1E 6BT, UK

Received 16 July 2009 / Accepted 20 October 2009

**ABSTRACT**

*Aims.* To probe cosmological fields beyond the Gaussian level, three-point statistics can be used, all of which are related to the bispectrum. Hence, measurements of CMB anisotropies, galaxy clustering, and weak gravitational lensing alike have to rely upon an accurate theoretical background concerning the bispectrum and its noise properties. If only small portions of the sky are considered, it is often desirable to perform the analysis in the flat-sky limit. We aim at a formal, detailed derivation of the bispectrum covariance in the flat-sky approximation, focusing on a pure two-dimensional Fourier-plane approach.

*Methods.* We define an unbiased estimator of the bispectrum, which takes the average over the overlap of annuli in Fourier space, and compute its full covariance. The outcome of our formalism is compared to the flat-sky spherical harmonic approximation in terms of the covariance, the behavior under parity transformations, and the information content. We introduce a geometrical interpretation of the averaging process in the estimator, thus providing an intuitive understanding.

*Results.* Contrary to foregoing work, we find a difference by a factor of two between the covariances of the Fourier-plane and the spherical harmonic approach. We argue that this discrepancy can be explained by the differing behavior with respect to parity. However, in an exemplary analysis it is demonstrated that the Fisher information of both formalisms agrees to high accuracy. Via the geometrical interpretation we are able to link the normalization in the bispectrum estimator to the area enclosed by the triangle configuration at consideration as well as to the Wigner symbol, which leads to convenient approximation formulae for the covariances of both approaches.

**Key words.** methods: statistical – cosmology: theory – cosmological parameters

## 1. Introduction

As the concordance model of cosmology becomes more and more consolidated, the focus increasingly turns towards probing effects beyond the standard paradigm, such as non-Gaussian initial conditions or the evolution of the large-scale structure in the highly non-linear regime. To lowest order, these effects can be measured by three-point statistics of the underlying fields, all of which are related to the bispectrum. Hence, work in both theory and observations concerning the bispectrum and its noise properties has been undertaken for CMB measurements (e.g. Hu 2000; Cooray et al. 2008), galaxy clustering surveys (e.g. Scoccimarro 2000; Scoccimarro et al. 2001; Sefusatti et al. 2006), or, more recently, weak gravitational lensing on cosmological scales (e.g. Bernardeau et al. 2002b; Jarvis et al. 2004; Takada & Jain 2004).

While theoretical computations at the bispectrum level are already considerably more demanding than for second-order statistics, this does apply even more so to the bispectrum covariance, which is a six-point statistic. On the full sky calculations are done by expanding the signal into spherical harmonics. If only small angular scales are considered, it is often more convenient to use a flat-sky approximation and work in terms of Fourier amplitudes. In the case of weak lensing the flat-sky limit is appropriate for practically all applications because signal correlations can only be measured up to separations of a few degrees.

Although other approaches exist in the literature (e.g. Matarrese et al. 1997; Sefusatti et al. 2006), a lot of work is done within a flat-sky spherical harmonic formalism (Hu 2000), which suffers – at least formally – from drawbacks. For instance, the resulting flat-sky expressions are valid only for integer arguments and thus for a bin width of unity, whereas it is desirable to evaluate the bispectrum and its covariance at real-valued angular frequencies and e.g. a logarithmic binning. The formulae still contain Wigner symbols whose physical meaning within a flat-sky consideration remain obscure. As the spherical harmonic expansion can only be done on the full unit sphere, the finite size of the survey at consideration is usually accounted for by multiplying a factor, containing the sky coverage, by hand. Moreover, the accuracy of some of the approximations in the transition between full sky and two-dimensional plane (see Hu 2000) is uncertain.

This work aims at clarifying the derivation of bispectrum covariances in the flat-sky limit. We attempt to do so by presenting a detailed calculation which is purely based on the two-dimensional Fourier formalism, followed by a comparison of this approach with the flat-sky spherical harmonic results in terms of their covariance, the behavior under parity transformations, and the information content. Moreover, we provide further insight and illustration by establishing relations between Wigner symbols, the averaging process in the bispectrum estimator, and a geometrical view.

The outline of this note is as follows: in Sect. 2 a bispectrum estimator is defined and shown to be unbiased. Section 3 introduces a geometrical interpretation, which is then applied to deal with the issue of degenerate triangle configurations. In Sect. 4 the covariance of the estimator defined beforehand is computed. The result is compared with the spherical harmonics approach and demonstrated

to be equivalent in terms of information content in Sect. 5. To explain the differences between the covariances, we also discuss the treatment of parity in both formalisms. We summarize our findings and conclude in Sect. 6. To avoid confusion, we refrain from using the term "flat sky" in the following, but refer to our formalism as "Fourier-plane" and to the approach as e.g. given in Hu (2000) as "spherical harmonic" (both are flat-sky approximations).

## 2. Bispectrum estimator

We consider a continuous, two-dimensional random field $g$ with mean zero, which is characterized by its complex Fourier amplitudes $g(\boldsymbol{\ell})$, where $\boldsymbol{\ell}$ denotes the angular frequency vector. Throughout, it will be assumed that this field is statistically homogeneous, i.e. invariant under translations, and statistically isotropic, i.e. invariant under rotations. In a cosmological context $g$ could for instance represent the temperature fluctuations of the CMB, the number density contrast of galaxy surveys, or the weak lensing convergence.

In what follows we will largely follow the approach of Joachimi et al. (2008), assuming likewise measurements in a compact, contiguous survey of size $A$. We will restrict our considerations to an angular extent much smaller than the size of the survey, i.e. to $\ell \gg \pi/\theta_{\rm max}$, where $\theta_{\rm max}$ is the maximum separation allowed by the survey geometry. Boundary effects due to the finite field size, as e.g. discussed in Joachimi et al. (2008) for the second-order level, can then be safely neglected.

Furthermore, we will not explicitly consider additional noise terms due to the discrete sampling of the continuous field $g$, for ease of notation. To account for these shot noise or, in the case of weak lensing, shape noise terms in the covariance, they can simply be added to the second-order measures, so in this Fourier space approach, to the power spectra (e.g. Kaiser 1998; Hu 1999). Note that the galaxy ellipticity, and not the convergence $\kappa$, is the direct observable in weak lensing. However, in absence of shape noise and for $\ell \gg 1$, the estimators in terms of the galaxy ellipticity, as given in Joachimi et al. (2008), can be re-written directly in terms of $\kappa$. Thus, without loss of generality, one can consider the convergence as the observable that the estimator is based on.

For a statistically homogeneous and isotropic random field one defines the bispectrum as

$$\langle g(\boldsymbol{\ell}_1)\, g(\boldsymbol{\ell}_2)\, g(\boldsymbol{\ell}_3)\rangle = (2\pi)^2\, \delta_{\rm D}^{(2)}(\boldsymbol{\ell}_1+\boldsymbol{\ell}_2+\boldsymbol{\ell}_3)\, B(\ell_1,\ell_2,\ell_3), \tag{1}$$

where $\delta_{\rm D}^{(2)}(\boldsymbol{\ell})$ is the two-dimensional Dirac delta-distribution. It ensures in (1) that the three angular frequency vectors form a triangle. For the assumed properties of $g$ the bispectrum has three independent components, for which we have chosen the triangle side lengths $|\boldsymbol{\ell}_i| \equiv \ell_i$. For the absolute values of $\boldsymbol{\ell}$ the triangle condition translates into the requirement $|\ell_1-\ell_2| \leq \ell_3 \leq \ell_1+\ell_2$ or equivalently for any permutation of the $\ell_i$.

Similarly to Joachimi et al. (2008), we construct an estimator of the bispectrum by averaging configurations over annuli, where here one has the complication of allowing only those combinations of angular frequency vectors that form a triangle. The area of an annulus with mean radius $\bar{\ell}_i$ is given by

$$A_R(\bar{\ell}_i) = 2\pi\bar{\ell}_i\Delta\ell_i \tag{2}$$

with the bin size $\Delta\ell_i$. Then we define the estimator

$$\hat{B}(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3) := \frac{(2\pi)^2}{A}\Lambda^{-1}\left(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3\right)\int_{A_R(\bar{\ell}_1)}\frac{{\rm d}^2\ell_1}{A_R(\bar{\ell}_1)}\int_{A_R(\bar{\ell}_2)}\frac{{\rm d}^2\ell_2}{A_R(\bar{\ell}_2)}\int_{A_R(\bar{\ell}_3)}\frac{{\rm d}^2\ell_3}{A_R(\bar{\ell}_3)}\,\delta_{\rm D}^{(2)}(\boldsymbol{\ell}_1+\boldsymbol{\ell}_2+\boldsymbol{\ell}_3)\, g(\boldsymbol{\ell}_1)\, g(\boldsymbol{\ell}_2)\, g(\boldsymbol{\ell}_3)\,, \tag{3}$$

where $\Lambda$ is a function that is related to the fraction of angular frequency combinations allowed by the triangle condition. It is defined such that (3) is unbiased, its explicit form being calculated below. Note that this bispectrum estimator is invariant under any permutation of its arguments since $\Lambda$ is symmetric as will be shown below.

In the following, we demonstrate that (3) is unbiased by computing the ensemble average,

$$\begin{aligned}\left\langle\hat{B}(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3)\right\rangle &= \frac{(2\pi)^2}{A}\Lambda^{-1}\left(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3\right)\int_{A_R(\bar{\ell}_1)}\frac{{\rm d}^2\ell_1}{A_R(\bar{\ell}_1)}\int_{A_R(\bar{\ell}_2)}\frac{{\rm d}^2\ell_2}{A_R(\bar{\ell}_2)}\int_{A_R(\bar{\ell}_3)}\frac{{\rm d}^2\ell_3}{A_R(\bar{\ell}_3)}\,(2\pi)^2\left(\delta_{\rm D}^{(2)}(\boldsymbol{\ell}_1+\boldsymbol{\ell}_2+\boldsymbol{\ell}_3)\right)^2 B(\ell_1,\ell_2,\ell_3) \\ &= (2\pi)^2\Lambda^{-1}\left(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3\right)\int_{A_R(\bar{\ell}_1)}\frac{{\rm d}^2\ell_1}{A_R(\bar{\ell}_1)}\int_{A_R(\bar{\ell}_2)}\frac{{\rm d}^2\ell_2}{A_R(\bar{\ell}_2)}\int_{A_R(\bar{\ell}_3)}\frac{{\rm d}^2\ell_3}{A_R(\bar{\ell}_3)}\,\delta_{\rm D}^{(2)}(\boldsymbol{\ell}_1+\boldsymbol{\ell}_2+\boldsymbol{\ell}_3)\, B(\ell_1,\ell_2,\ell_3).\end{aligned} \tag{4}$$

In the first step the definition of the bispectrum (3) was inserted. The appearance of a squared delta-distribution requires taking into account the finite survey size. As shown in Joachimi et al. (2008), one can identify

$$\left(\delta_{\rm D}^{(2)}(\boldsymbol{\ell})\right)^2 \to \frac{A}{(2\pi)^2}\,\delta_{\rm D}^{(2)}(\boldsymbol{\ell}), \tag{5}$$

which results in the second equality of (4).

Since the bispectrum only depends on the magnitudes of the angular frequency vectors we can perform the integrations over the polar angles of the $\boldsymbol{\ell}$-integrals. If $\varphi_{\ell_i}$ denotes the polar angle of $\boldsymbol{\ell}_i$, one gets

$$\begin{aligned}&\int_0^{2\pi}{\rm d}\varphi_{\ell_1}\int_0^{2\pi}{\rm d}\varphi_{\ell_2}\int_0^{2\pi}{\rm d}\varphi_{\ell_3}\,\delta_{\rm D}^{(2)}(\boldsymbol{\ell}_1+\boldsymbol{\ell}_2+\boldsymbol{\ell}_3) = \int_0^{2\pi}{\rm d}\varphi_{\ell_1}\int_0^{2\pi}{\rm d}\varphi_{\ell_2}\int_0^{2\pi}{\rm d}\varphi_{\ell_3}\int\frac{{\rm d}^2\theta}{(2\pi)^2}\,{\rm e}^{{\rm i}(\boldsymbol{\ell}_1+\boldsymbol{\ell}_2+\boldsymbol{\ell}_3)\cdot\boldsymbol{\theta}} \\ &= \int\frac{{\rm d}^2\theta}{(2\pi)^2}\,(2\pi)^3 J_0(\ell_1\theta)\, J_0(\ell_2\theta)\, J_0(\ell_3\theta) \;=\; (2\pi)^2\int{\rm d}\theta\,\theta\, J_0(\ell_1\theta)\, J_0(\ell_2\theta)\, J_0(\ell_3\theta) \;=\; 2\pi\,\Lambda\,(\ell_1,\ell_2,\ell_3)\,.\end{aligned} \tag{6}$$

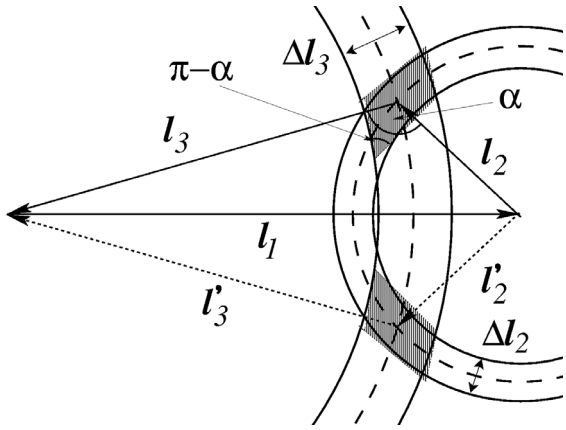


**Fig. 1.** Sketch of the annuli and their overlap for fixed $\boldsymbol{\ell}_1$. The region of overlap is approximated by the shaded parallelograms. Note that due to mirror symmetry a second shaded area, related to the triangle $\boldsymbol{\ell}_1, \boldsymbol{\ell}'_2, \boldsymbol{\ell}'_3$, contributes as well.

After inserting one possible representation of the delta-distribution in the first equality, we have made use of the definition of the Bessel function of the first kind of order 0,

$$J_0(x) = \int_0^{2\pi} \frac{\mathrm{d}\varphi}{2\pi}\, \mathrm{e}^{\mathrm{i}x\cos\varphi}. \tag{7}$$

The result of the integral over three Bessel functions is taken from Gradshteyn et al. (2000), formula No. 6.578.9, where we have defined

$$\Lambda(\ell_1,\ell_2,\ell_3) \equiv \begin{cases} \left\{\frac{1}{4}\sqrt{2\ell_1^2\ell_2^2 + 2\ell_1^2\ell_3^2 + 2\ell_2^2\ell_3^2 - \ell_1^4 - \ell_2^4 - \ell_3^4}\right\}^{-1} & \text{if } |\ell_1-\ell_2| < \ell_3 < \ell_1+\ell_2 \\ 0 & \text{else,} \end{cases} \tag{8}$$

i.e. if $\ell_1, \ell_2, \ell_3$ are chosen such that they can form the sides of a triangle, then $\Lambda^{-1}$ is the area of this triangle. Hence, (6) represents the defining equation for $\Lambda$. The set of integrations (6) is also performed within the spherical harmonic approach, see the appendix of Hu (2000), with a different result, which will be investigated in Sect. 5.1. Note furthermore that $\Lambda(\ell_1,\ell_2,\ell_3) = 0$ in case the angular frequency vectors are collinear or equivalently, if $\ell_i + \ell_j = \ell_k$ for some combination $i, j, k \in \{1, 2, 3\}$. At the same time, the bispectrum is non-zero for these degenerate triangle configurations, see (1). For the time being, we exclude degenerate triangles from the derivation, but develop a treatment for these cases in Sect. 3.2.

Inserting (6) into (4), one obtains

$$\left\langle \hat{B}(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3)\right\rangle = (2\pi)^3\Lambda^{-1}\left(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3\right) \int_{\bar{\ell}_1-1/2\Delta\ell}^{\bar{\ell}_1+1/2\Delta\ell} \frac{\mathrm{d}\ell_1\ell_1}{A_R(\bar{\ell}_1)} \int_{\bar{\ell}_2-1/2\Delta\ell}^{\bar{\ell}_2+1/2\Delta\ell} \frac{\mathrm{d}\ell_2\ell_2}{A_R(\bar{\ell}_2)} \int_{\bar{\ell}_3-1/2\Delta\ell}^{\bar{\ell}_3+1/2\Delta\ell} \frac{\mathrm{d}\ell_3\ell_3}{A_R(\bar{\ell}_3)}\, \Lambda(\ell_1,\ell_2,\ell_3)\, B(\ell_1,\ell_2,\ell_3). \tag{9}$$

Analogous to the derivation at the level of second-order statistics (Joachimi et al. 2008) we assume now that the annuli are thin enough such that $\Lambda$ within the integral, evaluated at the average $\ell$-values, can be taken out of the integration. Applying in addition (2), one arrives at

$$\left\langle \hat{B}(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3)\right\rangle \approx \int_{\bar{\ell}_1-1/2\Delta\ell}^{\bar{\ell}_1+1/2\Delta\ell} \frac{\mathrm{d}\ell_1\ell_1}{\bar{\ell}_1\Delta\ell_1} \int_{\bar{\ell}_2-1/2\Delta\ell}^{\bar{\ell}_2+1/2\Delta\ell} \frac{\mathrm{d}\ell_2\ell_2}{\bar{\ell}_2\Delta\ell_2} \int_{\bar{\ell}_3-1/2\Delta\ell}^{\bar{\ell}_3+1/2\Delta\ell} \frac{\mathrm{d}\ell_3\ell_3}{\bar{\ell}_3\Delta\ell_3}\, B(\ell_1,\ell_2,\ell_3) \equiv B(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3), \tag{10}$$

where in the last step the bin-averaged bispectrum was defined. Hence, (3) defines an unbiased estimator of the bispectrum. Following the restrictions on (6), this estimator is non-zero if the condition $|\bar{\ell}_1 - \bar{\ell}_2| < \bar{\ell}_3 < \bar{\ell}_1 + \bar{\ell}_2$, or likewise for all permutations, holds.

## 3. Averaging over triangles

A central step in the construction of the bispectrum estimator (3) is the correct treatment of the averaging over annuli, given the triangle condition. This section provides an illustrative, geometrical interpretation of the averaging process and applies this view to a practical treatment of degenerate triangle configurations.

### 3.1. Geometrical interpretation

Without loss of generality consider $\boldsymbol{\ell}_1$ to be fixed. Due to the assumed statistical isotropy of the underlying random field the angular integration over $\varphi_{\ell_1}$ is expected to simply reduce to an average over all directions of $\boldsymbol{\ell}_1$. Then the geometric situation in the Fourier plane can be seen as in Fig. 1. For a given triangle, composed of the mean vectors $\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, \boldsymbol{\ell}_3$ with lengths $\bar{\ell}_i$ for $i = \{1, 2, 3\}$, the annuli for $\bar{\ell}_2$ and $\bar{\ell}_3$ are shown. Due to the triangle condition, the average is not taken over the whole area of the annuli, but merely over the region that the annuli have in common. This area of overlap is well approximated by a parallelogram of size $A_\parallel = \Delta\ell_2\Delta\ell_3/\sin\alpha$, where $\alpha$ is the internal angle of the triangle opposite $\bar{\ell}_1$. This relation can readily be computed from the geometry of the sketch and by noting $\sin\alpha = \sin(\pi - \alpha)$.

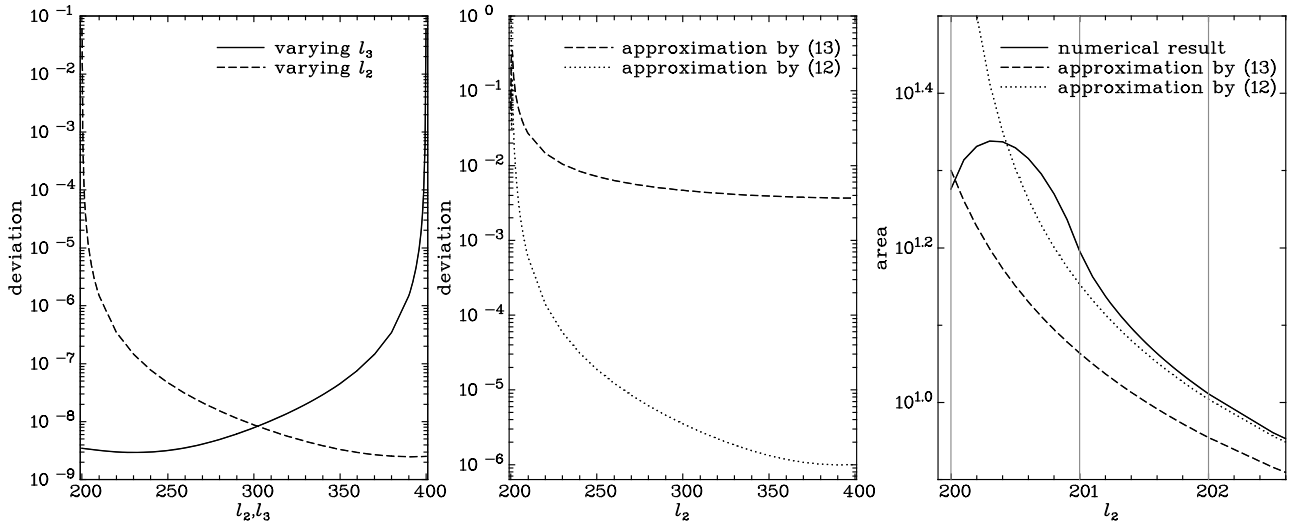


**Fig. 2.** Comparison of expressions for the overlap area of annuli. *Left panel*: relative deviation of (12) from the overlap area of the annuli as a function of angular frequency. The bin width is kept constant at $\Delta = 0.05$. The solid curve shows results for $\bar{\ell}_1, \bar{\ell}_2 = 200$ and varying $\bar{\ell}_3$, while the dashed curve corresponds to $\bar{\ell}_1 = 200, \bar{\ell}_3 = 400$ and varying $\bar{\ell}_2$. *Center panel*: same as above for the case $\bar{\ell}_1 = 200, \bar{\ell}_3 = 400$ and varying $\bar{\ell}_2$, but with $\Delta = 1$. The dotted line illustrates the deviation of (12), the dashed line the deviation of (13). *Right panel*: area of overlap for the case $\bar{\ell}_1 = 200, \bar{\ell}_3 = 400$ and varying $\bar{\ell}_2$, with $\Delta = 1$. The solid curve corresponds to the actual area, the dotted curve to (12), and the dashed curve to (13). Note that due to $\Delta = 1$ only the values at integer values of $\bar{\ell}_2$ are relevant for the covariance calculation.

The configuration is mirror-symmetric with respect to an axis through $\boldsymbol{\ell}_1$. Correspondingly, another area of overlap of the same size, which is connected to the triangle $\boldsymbol{\ell}_1$, $\boldsymbol{\ell}'_2$, $\boldsymbol{\ell}'_3$, contributes as well. Noting that axis reflection is in two dimensions equivalent to the parity transformation, the averaging is performed over triangles of both parities. A detailed discussion on this issue is given in Sect. 5.2.

As the angle $\alpha$ can also be related to the size of the triangle at consideration, $\Lambda^{-1} = (1/2)\,\bar{\ell}_2\bar{\ell}_3 \sin\alpha$, one finds the following correspondence of expressions:

$$\int_0^{2\pi} \frac{\mathrm{d}\varphi_{\ell_1}}{2\pi} \int_0^{2\pi} \frac{\mathrm{d}\varphi_{\ell_2}}{2\pi} \int_0^{2\pi} \frac{\mathrm{d}\varphi_{\ell_3}}{2\pi}\, \delta_{\mathrm{D}}^{(2)}(\boldsymbol{\ell}_1 + \boldsymbol{\ell}_2 + \boldsymbol{\ell}_3) \;=\; (2\pi)^{-2}\; \Lambda\left(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3\right) \;=\; \frac{2\,A_{\|}}{A_R(\bar{\ell}_2)\,A_R(\bar{\ell}_3)}, \tag{11}$$

where the first equality is an immediate consequence of (6). To arrive at the last expression, we used (2). Hence, the angular integration over the delta-distribution yields the ratio of the area of overlap $A_{\mathrm{overlap}}$, approximated by $2\,A_{\|}$, and the product of the area of the annuli the $\ell$-integrations (excluding the fixed $\boldsymbol{\ell}_1$) run over. This ratio is in turn proportional to the inverse of the area of the triangle spanned by the angular frequency vectors. Therefore, by placing a prefactor of $\Lambda^{-1}$ in the estimator (3), one replaces the normalization by the area of the annuli with the effective area, over which the average is actually performed.

Two approximations are involved in this picture. First, the shaded regions in Fig. 1 are approximated as parallelograms, which is a good assumption if the angle, at which the two annuli intersect, does not become too small. Moreover, the narrower the annuli, the less discrepancy between the area of the parallelogram and the actual overlap is expected. If the triangle approaches the degenerate case, where $\bar{\ell}_2$ and $\bar{\ell}_3$ eventually come to lie on $\bar{\ell}_1$, the area of overlap attains a more complex shape. In particular, the correspondence to the area of the triangle, whose inverse is divergent, does not hold anymore. Second, reconsidering (9), we have replaced the average of $\Lambda$ over triangle side lengths by $\Lambda$, evaluated at the average side lengths. This approximation similarly breaks down for thick annuli and configurations in which a small change in the length of an angular frequency vector causes a strong change in the size of the overlap region, as is the case near degeneracy.

In Fig. 2, we have plotted the relative deviation of

$$A_{\mathrm{overlap}}^{\mathrm{n.d.}} = \bar{\ell}_2\,\bar{\ell}_3\,\Delta\ell_2\,\Delta\ell_3\,\Lambda\left(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3\right) \tag{12}$$

from the actual area of the overlap region, which we calculated numerically. For simplicity, we assume a constant bin width $\Delta\ell_2 = \Delta\ell_3 \equiv \Delta$ for all computations related to Fig. 2. For a small bin width $\Delta = 0.05$, given integer steps in $\ell$, we find for the two configurations considered in the top panel that the approximation of the overlap area by parallelograms is excellent for the vast majority of triangle configurations. However, as expected, the deviation rises sharply when approaching the degenerate case. Changing to $\Delta = 1$, i.e. the maximum meaningful bin width in this setup, the relative deviation is larger, but still very small except for triangles close to degeneracy.

### *3.2. Degenerate triangles*

As discussed in the foregoing section, the approximations made in the course of the construction of the bispectrum estimator break down for degenerate triangle configurations. Equation (11) becomes invalid, the inverse area of the triangle $\Lambda$ diverging. Yet, to be of practical use, it is necessary to extend the validity of (3) to the case of degenerate triangles. We do so by making use of the geometrical interpretation of the averaging process.

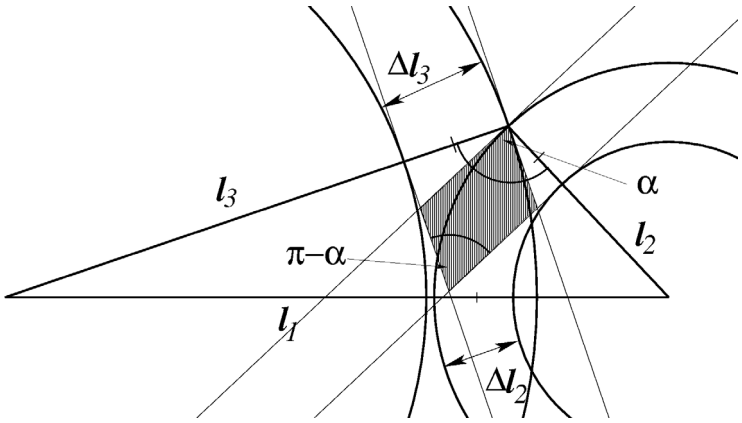


**Fig. 3.** Sketch of the region averaged over in case of a degenerate triangle, again for fixed $\boldsymbol{\ell}_1$. The depicted triangle has side lengths $\bar{\ell}_1$, $\bar{\ell}_2 + \Delta\ell_2/2$, and $\bar{\ell}_3 + \Delta\ell_3/2$. The shaded parallelogram approximates the region of overlap, the mirror-symmetric counterpart not being shown.

Still keeping $\boldsymbol{\ell}_1$ fixed, consider the situation of a degenerate triangle as sketched in Fig. 3. Here, $\bar{\ell}_1 = \bar{\ell}_2 + \bar{\ell}_3$, while the depicted triangle has side lengths $\bar{\ell}_1$, $\bar{\ell}_2 + \Delta\ell_2/2$, and $\bar{\ell}_3 + \Delta\ell_3/2$. Again, we identify a parallelogram that serves as an approximation for the overlap of the annuli, although, as the sketch suggests, with considerably lower accuracy. The relation between the internal angle $\alpha$ of the triangle to the internal angle of the parallelogram $\pi - \alpha$ holds as before, so that one can derive an analogous formula to (12), but with modified triangle side lengths. Symmetrizing this argument for all three angular frequency vectors, we propose the following formula to compute the area of overlap in the degenerate case:

$$A_{\rm overlap}^{\rm deg.} := \bar{\ell}_2\ \bar{\ell}_3\ \Delta\ell_2\ \Delta\ell_3\ \Lambda\left(\bar{\ell}_1 + \frac{\Delta\ell_1}{2}, \bar{\ell}_2 + \frac{\Delta\ell_2}{2}, \bar{\ell}_3 + \frac{\Delta\ell_3}{2}\right). \tag{13}$$

As is evident from Fig. 2, center panel, the relative deviation of (13) from the true overlap area is still fairly small, but – unsurprisingly – noticeably stronger than for (12). The right-hand panel gives the size of the overlap area for values of $\bar{\ell}_2$ close to 200, which is the degenerate case. Note that since this plot was determined for $\Delta = 1$, the values relevant for the covariance calculation are only those at integer $\ell$. While the true overlap area curbs down to a finite value at $\bar{\ell}_2 = 200$, (12) diverges. Still, for $\bar{\ell}_2 = 201$ it produces a fair and for $\bar{\ell}_2 \geq 202$ an excellent approximation. In the degenerate case (13) is indeed capable of reproducing the size of the overlap area to good accuracy.

Thus, we suggest to incorporate degenerate triangle configurations into our formalism by replacing $\bar{\ell}_i \rightarrow \bar{\ell}_i + \Delta\ell_i/2$ in all arguments of $\Lambda$ for these cases. This way, we heuristically correct for the breakdown of approximations in the assignment of the actual area, over which triangle configurations are averaged. While the modification is at this stage only motivated by the geometrical interpretation, we will establish a more strict foundation of (13) by relating it to Wigner symbols in Sect. 5.1.

## 4. Bispectrum covariance

The covariance of the bispectrum is defined as

$$\begin{aligned}\mathrm{Cov}\left(\hat{B}(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3),\ \hat{B}(\bar{\ell}_4,\bar{\ell}_5,\bar{\ell}_6)\right) &\equiv \left\langle\left(\hat{B}(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3) - \left\langle\hat{B}(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3)\right\rangle\right)\left(\hat{B}(\bar{\ell}_4,\bar{\ell}_5,\bar{\ell}_6) - \left\langle\hat{B}(\bar{\ell}_4,\bar{\ell}_5,\bar{\ell}_6)\right\rangle\right)\right\rangle \\ &= \left\langle\hat{B}(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3)\ \hat{B}(\bar{\ell}_4,\bar{\ell}_5,\bar{\ell}_6)\right\rangle - B(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3)\ B(\bar{\ell}_4,\bar{\ell}_5,\bar{\ell}_6).\end{aligned} \tag{14}$$

The computation of the correlator of two bispectrum estimators involves a 6-point correlator of $g$, which can be expanded into its connected parts as e.g. outlined in Bernardeau et al. (2002a). Denoting the connected correlators by a subscript c, which will only be done in this paragraph to avoid confusion, we obtain

$$\begin{aligned}&\langle g(\boldsymbol{\ell}_1)\ g(\boldsymbol{\ell}_2)\ g(\boldsymbol{\ell}_3)\ g(\boldsymbol{\ell}_4)\ g(\boldsymbol{\ell}_5)\ g(\boldsymbol{\ell}_6)\rangle = \\ &\quad\langle g(\boldsymbol{\ell}_1)\ g(\boldsymbol{\ell}_2)\rangle_{\rm c}\,\langle g(\boldsymbol{\ell}_3)\ g(\boldsymbol{\ell}_4)\rangle_{\rm c}\,\langle g(\boldsymbol{\ell}_5)\ g(\boldsymbol{\ell}_6)\rangle_{\rm c} + \langle g(\boldsymbol{\ell}_1)\ g(\boldsymbol{\ell}_2)\rangle_{\rm c}\,\langle g(\boldsymbol{\ell}_3)\ g(\boldsymbol{\ell}_5)\rangle_{\rm c}\,\langle g(\boldsymbol{\ell}_4)\ g(\boldsymbol{\ell}_6)\rangle_{\rm c} + (13\ \text{perm.}) \\ &+\ \langle g(\boldsymbol{\ell}_1)\ g(\boldsymbol{\ell}_2)\ g(\boldsymbol{\ell}_3)\rangle_{\rm c}\,\langle g(\boldsymbol{\ell}_4)\ g(\boldsymbol{\ell}_5)\ g(\boldsymbol{\ell}_6)\rangle_{\rm c} + \langle g(\boldsymbol{\ell}_1)\ g(\boldsymbol{\ell}_2)\ g(\boldsymbol{\ell}_4)\rangle_{\rm c}\,\langle g(\boldsymbol{\ell}_3)\ g(\boldsymbol{\ell}_5)\ g(\boldsymbol{\ell}_6)\rangle_{\rm c} + (8\ \text{perm.}) \\ &+\ \langle g(\boldsymbol{\ell}_1)\ g(\boldsymbol{\ell}_2)\ g(\boldsymbol{\ell}_3)\ g(\boldsymbol{\ell}_4)\rangle_{\rm c}\,\langle g(\boldsymbol{\ell}_5)\ g(\boldsymbol{\ell}_6)\rangle_{\rm c} + \langle g(\boldsymbol{\ell}_1)\ g(\boldsymbol{\ell}_2)\ g(\boldsymbol{\ell}_3)\ g(\boldsymbol{\ell}_5)\rangle_{\rm c}\,\langle g(\boldsymbol{\ell}_4)\ g(\boldsymbol{\ell}_6)\rangle_{\rm c} + (13\ \text{perm.}) \\ &+\ \langle g(\boldsymbol{\ell}_1)\ g(\boldsymbol{\ell}_2)\ g(\boldsymbol{\ell}_3)\ g(\boldsymbol{\ell}_4)\ g(\boldsymbol{\ell}_5)\ g(\boldsymbol{\ell}_6)\rangle_{\rm c}\,,\end{aligned} \tag{15}$$

where the permutations are to be taken with respect to the indices of the angular frequencies such that for each correlator, no combination of indices is repeated (as the individual correlators are invariant under permutations of the indices within that correlator). The resulting connected parts are related to spectra via

$$\left\langle\prod_{i=1}^{N} g(\boldsymbol{\ell}_i)\right\rangle_{\rm c} = (2\pi)^2\ \delta_{\rm D}^{(2)}\left(\sum_{i=1}^{N}\boldsymbol{\ell}_i\right)\ P_N(\boldsymbol{\ell}_1,\ \ldots,\boldsymbol{\ell}_N), \tag{16}$$

where we identify $P_2(\boldsymbol{\ell}_1,\boldsymbol{\ell}_2) \equiv P(\ell_1)$ as the power spectrum and $P_3(\boldsymbol{\ell}_1,\boldsymbol{\ell}_2,\boldsymbol{\ell}_3) \equiv B(\ell_1,\ell_2,\ell_3)$ as the bispectrum. As the random field $g$ vanishes on average, $\langle g(\boldsymbol{\ell})\rangle = 0$, only $P_4$ (the trispectrum) and $P_6$ (the pentaspectrum) will appear in addition in the covariance formula, see (15).

Introducing a shorthand notation $\int_{A_R(\bar{\ell}_i)} \mathrm{d}^2\ell_i/A_R(\bar{\ell}_i) \equiv \int_i$, one can write the correlator of the bispectrum estimators by using (3) as

$$\left\langle \hat{B}(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3)\,\hat{B}(\bar{\ell}_4,\bar{\ell}_5,\bar{\ell}_6)\right\rangle = \frac{(2\,\pi)^4}{A^2}\,\Lambda^{-1}\left(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3\right)\Lambda^{-1}\left(\bar{\ell}_4,\bar{\ell}_5,\bar{\ell}_6\right)\int_1\int_2\int_3\int_4\int_5\int_6 \times\ \delta_{\mathrm{D}}^{(2)}(\boldsymbol{\ell}_1+\boldsymbol{\ell}_2+\boldsymbol{\ell}_3)\,\delta_{\mathrm{D}}^{(2)}(\boldsymbol{\ell}_4+\boldsymbol{\ell}_5+\boldsymbol{\ell}_6)\,\langle g(\boldsymbol{\ell}_1)\,g(\boldsymbol{\ell}_2)\,g(\boldsymbol{\ell}_3)\,g(\boldsymbol{\ell}_4)\,g(\boldsymbol{\ell}_5)\,g(\boldsymbol{\ell}_6)\rangle\,, \tag{17}$$

which then allows us to insert (15) and (16). The resulting terms contain products of several delta-distribution. Concerning the terms containing three two-point correlators, one obtains e.g.

$$\begin{aligned}&\delta_{\mathrm{D}}^{(2)}(\boldsymbol{\ell}_1+\boldsymbol{\ell}_2+\boldsymbol{\ell}_3)\,\delta_{\mathrm{D}}^{(2)}(\boldsymbol{\ell}_4+\boldsymbol{\ell}_5+\boldsymbol{\ell}_6)\,\delta_{\mathrm{D}}^{(2)}(\boldsymbol{\ell}_1+\boldsymbol{\ell}_2)\,\delta_{\mathrm{D}}^{(2)}(\boldsymbol{\ell}_3+\boldsymbol{\ell}_4)\,\delta_{\mathrm{D}}^{(2)}(\boldsymbol{\ell}_5+\boldsymbol{\ell}_6)P(\ell_1)\,P(\ell_3)\,P(\ell_5) = \\ &\delta_{\mathrm{D}}^{(2)}(\boldsymbol{\ell}_3)\,\delta_{\mathrm{D}}^{(2)}(\boldsymbol{\ell}_4+\boldsymbol{\ell}_5+\boldsymbol{\ell}_6)\,\delta_{\mathrm{D}}^{(2)}(\boldsymbol{\ell}_1+\boldsymbol{\ell}_2)\,\delta_{\mathrm{D}}^{(2)}(\boldsymbol{\ell}_3+\boldsymbol{\ell}_4)\,\delta_{\mathrm{D}}^{(2)}(\boldsymbol{\ell}_5+\boldsymbol{\ell}_6)P(\ell_1)\,P(0)\,P(\ell_5)\ =\ 0,\end{aligned} \tag{18}$$

and likewise for all other terms in which the correlators do not contain one angular frequency each out of the sets $\{\boldsymbol{\ell}_1,\boldsymbol{\ell}_2,\boldsymbol{\ell}_3\}$ and $\{\boldsymbol{\ell}_4,\boldsymbol{\ell}_5,\boldsymbol{\ell}_6\}$. A similar argument holds for the terms composed of power spectrum and trispectrum, where the trispectrum is readily shown to vanish if the two-point correlator contains both angular frequencies out of the same of the sets mentioned above. This way, the number of terms with three power spectra reduces to 6, the number of terms with trispectrum and power spectrum to 9.

To proceed, we demonstrate the treatment of some exemplary terms in the covariance, for instance

$$\begin{aligned}&\int_1\int_2\int_3\int_4\int_5\int_6 \delta_{\mathrm{D}}^{(2)}(\boldsymbol{\ell}_1+\boldsymbol{\ell}_2+\boldsymbol{\ell}_3)\,\delta_{\mathrm{D}}^{(2)}(\boldsymbol{\ell}_4+\boldsymbol{\ell}_5+\boldsymbol{\ell}_6)\,\langle g(\boldsymbol{\ell}_1)\,g(\boldsymbol{\ell}_4)\rangle\langle g(\boldsymbol{\ell}_2)\,g(\boldsymbol{\ell}_5)\rangle\langle g(\boldsymbol{\ell}_3)\,g(\boldsymbol{\ell}_6)\rangle = \\ &(2\pi)^6\int_1\int_2\int_3\int_4\int_5\int_6 \delta_{\mathrm{D}}^{(2)}(\boldsymbol{\ell}_1+\boldsymbol{\ell}_2+\boldsymbol{\ell}_3)\,\delta_{\mathrm{D}}^{(2)}(\boldsymbol{\ell}_4+\boldsymbol{\ell}_5+\boldsymbol{\ell}_6)\,\delta_{\mathrm{D}}^{(2)}(\boldsymbol{\ell}_1+\boldsymbol{\ell}_4)\,\delta_{\mathrm{D}}^{(2)}(\boldsymbol{\ell}_2+\boldsymbol{\ell}_5)\,\delta_{\mathrm{D}}^{(2)}(\boldsymbol{\ell}_3+\boldsymbol{\ell}_6)\,P(\ell_1)\,P(\ell_2)\,P(\ell_3) \\ =\ &(2\pi)^6\frac{\delta_{\bar{\ell}_1\bar{\ell}_4}\,\delta_{\bar{\ell}_2\bar{\ell}_5}\,\delta_{\bar{\ell}_3\bar{\ell}_6}}{A_R(\bar{\ell}_1)\,A_R(\bar{\ell}_2)\,A_R(\bar{\ell}_3)}\int_1\int_2\int_3\left(\delta_{\mathrm{D}}^{(2)}(\boldsymbol{\ell}_1+\boldsymbol{\ell}_2+\boldsymbol{\ell}_3)\right)^2\,P(\ell_1)\,P(\ell_2)\,P(\ell_3)\,,\end{aligned} \tag{19}$$

where the integrations over $\boldsymbol{\ell}_4$ to $\boldsymbol{\ell}_6$ only yield a non-zero result if the annuli of the angular frequencies in the corresponding delta-distributions, which are integrated over, coincide. Thus, for every such integration a Kronecker symbol is generated. The resulting expression in (19) can now easily be simplified by using (5), producing a factor of $A/(2\,\pi)^2$, and subsequently (6) to execute the remaining angular integrations. Note that again only the delta-distribution depends on the polar angles of the angular frequencies. Therefore, considering only the Gaussian contribution to the covariance, (17) turns into

$$\left\langle \hat{B}(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3)\,\hat{B}(\bar{\ell}_4,\bar{\ell}_5,\bar{\ell}_6)\right\rangle_{\mathrm{Gauss}} = \frac{(2\,\pi)^9}{A\,A_R(\bar{\ell}_1)\,A_R(\bar{\ell}_2)\,A_R(\bar{\ell}_3)}\,\Lambda^{-1}\left(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3\right) \times\ D_{\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3,\bar{\ell}_4,\bar{\ell}_5,\bar{\ell}_6}\int_{A_R(\bar{\ell}_1)}\frac{\mathrm{d}\ell_1\ell_1}{A_R(\bar{\ell}_1)}\,P(\ell_1)\int_{A_R(\bar{\ell}_2)}\frac{\mathrm{d}\ell_2\ell_2}{A_R(\bar{\ell}_2)}\,P(\ell_2)\int_{A_R(\bar{\ell}_3)}\frac{\mathrm{d}\ell_3\ell_3}{A_R(\bar{\ell}_3)}\,P(\ell_3), \tag{20}$$

where we again pulled $\Lambda$, evaluated at the averaged angular frequencies, out of the radial integrations. Besides, we defined the shorthand notation

$$D_{\ell_1,\ell_2,\ell_3,\ell_4,\ell_5,\ell_6} \equiv \delta_{\ell_1\ell_4}\,\delta_{\ell_2\ell_5}\,\delta_{\ell_3\ell_6} + \delta_{\ell_1\ell_5}\,\delta_{\ell_2\ell_4}\,\delta_{\ell_3\ell_6} + \delta_{\ell_1\ell_4}\,\delta_{\ell_2\ell_6}\,\delta_{\ell_3\ell_5} + \delta_{\ell_1\ell_5}\,\delta_{\ell_2\ell_6}\,\delta_{\ell_3\ell_4} + \delta_{\ell_1\ell_6}\,\delta_{\ell_2\ell_4}\,\delta_{\ell_3\ell_5} + \delta_{\ell_1\ell_6}\,\delta_{\ell_2\ell_5}\,\delta_{\ell_3\ell_4} \tag{21}$$

for convenience. By making use of (2) and defining the bin-averaged power spectrum as

$$P(\bar{\ell}_i) \equiv \int_{\bar{\ell}_i-1/2\Delta\ell}^{\bar{\ell}_i+1/2\Delta\ell}\frac{\mathrm{d}\ell_i\ell_i}{\bar{\ell}_i\Delta\ell}\,P(\ell_i), \tag{22}$$

see Joachimi et al. (2008), in analogy to the definition of the bin-averaged bispectrum, one obtains the expression

$$\left\langle \hat{B}(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3)\,\hat{B}(\bar{\ell}_4,\bar{\ell}_5,\bar{\ell}_6)\right\rangle_{\mathrm{Gauss}} = \frac{(2\,\pi)^3}{A\,\bar{\ell}_1\bar{\ell}_2\bar{\ell}_3\Delta\ell_1\Delta\ell_2\Delta\ell_3}\,\Lambda^{-1}\left(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3\right)\,D_{\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3,\bar{\ell}_4,\bar{\ell}_5,\bar{\ell}_6}\,P(\bar{\ell}_1)P(\bar{\ell}_2)P(\bar{\ell}_3). \tag{23}$$

Terms composed of two three-point correlators can be processed as follows,

$$\begin{aligned}&\int_1\int_2\int_3\int_4\int_5\int_6 \delta_{\mathrm{D}}^{(2)}(\boldsymbol{\ell}_1+\boldsymbol{\ell}_2+\boldsymbol{\ell}_3)\,\delta_{\mathrm{D}}^{(2)}(\boldsymbol{\ell}_4+\boldsymbol{\ell}_5+\boldsymbol{\ell}_6)\,\langle g(\boldsymbol{\ell}_1)\,g(\boldsymbol{\ell}_2)\,g(\boldsymbol{\ell}_4)\rangle\langle g(\boldsymbol{\ell}_3)\,g(\boldsymbol{\ell}_5)\,g(\boldsymbol{\ell}_6)\rangle = \\ &(2\pi)^4\int_1\int_2\int_3\int_4\int_5\int_6 \delta_{\mathrm{D}}^{(2)}(\boldsymbol{\ell}_1+\boldsymbol{\ell}_2+\boldsymbol{\ell}_3)\,\delta_{\mathrm{D}}^{(2)}(\boldsymbol{\ell}_4+\boldsymbol{\ell}_5+\boldsymbol{\ell}_6)\,\delta_{\mathrm{D}}^{(2)}(\boldsymbol{\ell}_1+\boldsymbol{\ell}_2+\boldsymbol{\ell}_4)\,\delta_{\mathrm{D}}^{(2)}(\boldsymbol{\ell}_3+\boldsymbol{\ell}_5+\boldsymbol{\ell}_6)\,B(\ell_1,\ell_2,\ell_4)\,B(\ell_3,\ell_5,\ell_6) \\ =\ &(2\pi)^4\,\delta_{\bar{\ell}_3\bar{\ell}_4}\int_1\int_2\int_3\int_5\int_6 \delta_{\mathrm{D}}^{(2)}(\boldsymbol{\ell}_1+\boldsymbol{\ell}_2+\boldsymbol{\ell}_3)\,\delta_{\mathrm{D}}^{(2)}(-\boldsymbol{\ell}_1-\boldsymbol{\ell}_2+\boldsymbol{\ell}_5+\boldsymbol{\ell}_6)\,\delta_{\mathrm{D}}^{(2)}(\boldsymbol{\ell}_3+\boldsymbol{\ell}_5+\boldsymbol{\ell}_6)\,B(\ell_1,\ell_2,|\boldsymbol{\ell}_1+\boldsymbol{\ell}_2|)\,B(\ell_3,\ell_5,\ell_6) \\ =\ &A\,(2\pi)^2\,\delta_{\bar{\ell}_3\bar{\ell}_4}\int_1\int_2\int_3\int_5\int_6 \delta_{\mathrm{D}}^{(2)}(\boldsymbol{\ell}_1+\boldsymbol{\ell}_2+\boldsymbol{\ell}_3)\,\delta_{\mathrm{D}}^{(2)}(\boldsymbol{\ell}_3+\boldsymbol{\ell}_5+\boldsymbol{\ell}_6)\,B(\ell_1,\ell_2,\ell_3)\,B(\ell_3,\ell_5,\ell_6),\end{aligned} \tag{24}$$

where to generate the Kronecker symbol $\delta_{\bar{\ell}_3\bar{\ell}_4}$, we made use of fact that $\boldsymbol{\ell}_1+\boldsymbol{\ell}_2 = -\boldsymbol{\ell}_3$ due to the corresponding delta-distribution. To arrive at the last equality, (5) has been applied after processing the arguments of the delta-distributions similar to (18). The

remaining terms, containing four- and six-point correlators of $g$, can be dealt with in close analogy to (24). We mention the special case

$$\begin{aligned}
&\int_1\int_2\int_3\int_4\int_5\int_6 \delta_{\rm D}^{(2)}(\boldsymbol{\ell}_1+\boldsymbol{\ell}_2+\boldsymbol{\ell}_3)\,\delta_{\rm D}^{(2)}(\boldsymbol{\ell}_4+\boldsymbol{\ell}_5+\boldsymbol{\ell}_6)\;\langle g(\boldsymbol{\ell}_1)\,g(\boldsymbol{\ell}_2)\,g(\boldsymbol{\ell}_3)\rangle\langle g(\boldsymbol{\ell}_4)\,g(\boldsymbol{\ell}_5)\,g(\boldsymbol{\ell}_6)\rangle = \\
&(2\pi)^4\int_1\int_2\int_3\int_4\int_5\int_6 \left(\delta_{\rm D}^{(2)}(\boldsymbol{\ell}_1+\boldsymbol{\ell}_2+\boldsymbol{\ell}_3)\right)^2\,\left(\delta_{\rm D}^{(2)}(\boldsymbol{\ell}_4+\boldsymbol{\ell}_5+\boldsymbol{\ell}_6)\right)^2 B(\ell_1,\ell_2,\ell_3)\,B(\ell_4,\ell_5,\ell_6) \\
=\;& A^2\int_1\int_2\int_3 \delta_{\rm D}^{(2)}(\boldsymbol{\ell}_1+\boldsymbol{\ell}_2+\boldsymbol{\ell}_3)\,B(\ell_1,\ell_2,\ell_3)\int_4\int_5\int_6 \delta_{\rm D}^{(2)}(\boldsymbol{\ell}_4+\boldsymbol{\ell}_5+\boldsymbol{\ell}_6)\,B(\ell_4,\ell_5,\ell_6) \\
=\;& \frac{A^2}{(2\pi)^4}\,\Lambda\left(\bar\ell_1,\bar\ell_2,\bar\ell_3\right)\,\Lambda\left(\bar\ell_4,\bar\ell_5,\bar\ell_6\right)\,B(\bar\ell_1,\bar\ell_2,\bar\ell_3)\,B(\bar\ell_4,\bar\ell_5,\bar\ell_6),
\end{aligned} \tag{25}$$

which, after inserting this expression into (17), cancels the product $B(\bar\ell_1,\bar\ell_2,\bar\ell_3)B(\bar\ell_4,\bar\ell_5,\bar\ell_6)$ in the definition of the covariance (14).

Combining these results, we obtain the total bispectrum covariance

$$\begin{aligned}
{\rm Cov}\left(B(\bar\ell_1,\bar\ell_2,\bar\ell_3),\;B(\bar\ell_4,\bar\ell_5,\bar\ell_6)\right) =\;& \frac{(2\,\pi)^3}{A\;\bar\ell_1\bar\ell_2\bar\ell_3\Delta\ell_1\Delta\ell_2\Delta\ell_3}\,\Lambda^{-1}\left(\bar\ell_1,\bar\ell_2,\bar\ell_3\right)\,D_{\bar\ell_1,\bar\ell_2,\bar\ell_3,\bar\ell_4,\bar\ell_5,\bar\ell_6}\,P(\bar\ell_1)P(\bar\ell_2)P(\bar\ell_3) \\
&+\frac{C}{A}\,\delta_{\bar\ell_3\bar\ell_4}\int_1\int_2\int_3\int_5\int_6 \delta_{\rm D}^{(2)}(\boldsymbol{\ell}_1+\boldsymbol{\ell}_2+\boldsymbol{\ell}_3)\,\delta_{\rm D}^{(2)}(\boldsymbol{\ell}_3+\boldsymbol{\ell}_5+\boldsymbol{\ell}_6)\,B(\ell_1,\ell_2,\ell_3)\,B(\ell_3,\ell_5,\ell_6) + (8\text{ perm.}) \\
&+\frac{C}{A}\,\delta_{\bar\ell_3\bar\ell_6}\int_1\int_2\int_3\int_4\int_5 \delta_{\rm D}^{(2)}(\boldsymbol{\ell}_1+\boldsymbol{\ell}_2+\boldsymbol{\ell}_3)\,\delta_{\rm D}^{(2)}(\boldsymbol{\ell}_4+\boldsymbol{\ell}_5-\boldsymbol{\ell}_3)\,P_4(\boldsymbol{\ell}_1,\boldsymbol{\ell}_2,\boldsymbol{\ell}_4,\boldsymbol{\ell}_5)\,P(\ell_3) + (8\text{ perm.}) \\
&+\frac{C}{A}\int_1\int_2\int_3\int_4\int_5\int_6 \delta_{\rm D}^{(2)}(\boldsymbol{\ell}_1+\boldsymbol{\ell}_2+\boldsymbol{\ell}_3)\,\delta_{\rm D}^{(2)}(\boldsymbol{\ell}_4+\boldsymbol{\ell}_5+\boldsymbol{\ell}_6)\,P_6(\boldsymbol{\ell}_1,\boldsymbol{\ell}_2,\boldsymbol{\ell}_3,\boldsymbol{\ell}_4,\boldsymbol{\ell}_5,\boldsymbol{\ell}_6),
\end{aligned} \tag{26}$$

where the prefactor reads $C \equiv (2\pi)^6\,\Lambda^{-1}\left(\bar\ell_1,\bar\ell_2,\bar\ell_3\right)\,\Lambda^{-1}\left(\bar\ell_4,\bar\ell_5,\bar\ell_6\right)$.

The general form of the covariance terms is in agreement with the expressions derived in Sefusatti et al. (2006). As mentioned in Sect. 2, shot or shape noise can readily be included into this covariance by adding a corresponding noise term to the power spectra. Weak lensing or galaxy clustering surveys often have in addition tomographic information, so that the data is binned into (photometric) redshift bins. The covariance can be generalized to this case in a straightforward manner by obeying the practical rule that each photometric redshift "sticks" to the angular frequency it is assigned to, see Takada & Jain (2004). A similar argument holds for the generalization to CMB polarization bispectrum covariances (Hu 2000).

## 5. Equivalence to spherical harmonics approach

In this section we demonstrate that both our and the spherical harmonic approach are equivalent in the sense that they measure the same information in a survey. Moreover, we investigate the behavior with respect to parity, and the relation between the covariances of both approaches, considering for the remainder of this work only the Gaussian part of (26).

### 5.1. Comparison of covariances

On the celestial sphere one can decompose the random field $g$ into spherical harmonics, which produces a set of coefficients $g_{\ell m}$ with $\ell, m$ integers and $\ell \geq 0$, $-\ell \leq m \leq \ell$. In terms of the $g_{\ell m}$ one can define a bispectrum estimator as (e.g. Hu 2000)

$$\hat B_{\ell_1,\ell_2,\ell_3} = \sum_{m_1,m_2,m_3}\begin{pmatrix}\ell_1 & \ell_2 & \ell_3\\ m_1 & m_2 & m_3\end{pmatrix} g_{\ell_1 m_1}\,g_{\ell_2 m_2}\,g_{\ell_3 m_3}, \tag{27}$$

where the object in parentheses is the Wigner-3$j$ symbol. Properties of the Wigner symbol are reviewed in Hu (2000); most importantly, it obeys the triangle condition, i.e. it is non-zero only for $|\ell_1-\ell_2| \leq \ell_3 \leq \ell_1+\ell_2$ and permutations thereof. For this estimator Hu (2000) derived the simple Gaussian covariance

$${\rm Cov}\left(B_{\bar\ell_1,\bar\ell_2,\bar\ell_3},\;B_{\bar\ell_4,\bar\ell_5,\bar\ell_6}\right) = D_{\bar\ell_1,\bar\ell_2,\bar\ell_3,\bar\ell_4,\bar\ell_5,\bar\ell_6}\;P_{\bar\ell_1}P_{\bar\ell_2}P_{\bar\ell_3}, \tag{28}$$

where $P_\ell$ denotes the full-sky power spectrum, and where $D_{\bar\ell_1,\bar\ell_2,\bar\ell_3,\bar\ell_4,\bar\ell_5,\bar\ell_6}$ is used as defined in (21). Moreover, he gives approximate relations between the spherical harmonic and Fourier-plane spectra,

$$P_\ell \approx P(\ell); \qquad B_{\ell_1,\ell_2,\ell_3} \approx \begin{pmatrix}\ell_1 & \ell_2 & \ell_3\\ 0 & 0 & 0\end{pmatrix}\sqrt{\frac{(2\ell_1+1)(2\ell_2+1)(2\ell_3+1)}{4\pi}}\,B(\ell_1,\ell_2,\ell_3), \tag{29}$$

valid for $\ell_1,\ell_2,\ell_3 \gg 1$. These equations can only hold for integer $\ell$. In addition, the Wigner symbol with $m_1 = m_2 = m_3 = 0$ vanishes for $L \equiv \ell_1+\ell_2+\ell_3$ odd, see the following section for details. Making use of the standard procedure of multiplying (28)

by an ad hoc factor of $f_{\rm sky}^{-1} = 4\pi/A$ to account for finite sky coverage of the survey, one can derive a flat-sky spherical harmonic covariance with (29) as (Hu 2000; Takada & Jain 2004)

$$\left\langle \hat{B}(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3)\,\hat{B}(\bar{\ell}_4,\bar{\ell}_5,\bar{\ell}_6)\right\rangle \approx \frac{(4\pi)^2\,D_{\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3,\bar{\ell}_4,\bar{\ell}_5,\bar{\ell}_6}}{A\,(2\bar{\ell}_1+1)\,(2\bar{\ell}_2+1)\,(2\bar{\ell}_3+1)}\begin{pmatrix}\bar{\ell}_1 & \bar{\ell}_2 & \bar{\ell}_3\\ 0 & 0 & 0\end{pmatrix}^{-2} P(\bar{\ell}_1)\,P(\bar{\ell}_2)\,P(\bar{\ell}_3), \tag{30}$$

where still the angular frequencies are required to be integer, and $L$ even. As is true for our approach, (30) holds for $\ell \gg 1$ only. To be able to compare this widely used formula to our results, a relation between the Wigner symbol and $\Lambda$ has to be found.

When comparing the spherical harmonics and the Fourier-plane approach, Hu (2000) already came across integrals of the form (6). We reproduce his computation,

$$\int \mathrm{d}^2\ell_1 \int \mathrm{d}^2\ell_2 \int \mathrm{d}^2\ell_3\,\delta_{\rm D}^{(2)}(\boldsymbol{\ell}_1+\boldsymbol{\ell}_2+\boldsymbol{\ell}_3) \;=\; \int \mathrm{d}^2\ell_1 \int \mathrm{d}^2\ell_2 \int \mathrm{d}^2\ell_3 \int \frac{\mathrm{d}^2\theta}{(2\pi)^2}\,\mathrm{e}^{\mathrm{i}(\boldsymbol{\ell}_1+\boldsymbol{\ell}_2+\boldsymbol{\ell}_3)\cdot\boldsymbol{\theta}} \tag{31}$$

$$\approx \int \mathrm{d}\ell_1\ell_1 \int \mathrm{d}\ell_2\ell_2 \int \mathrm{d}\ell_3\ell_3 \sqrt{\frac{(2\pi)^5}{\ell_1\ell_2\ell_3}} \int \mathrm{d}\Omega\; Y^0_{\ell_1}(\boldsymbol{n})\,Y^0_{\ell_2}(\boldsymbol{n})\,Y^0_{\ell_3}(\boldsymbol{n}) \;\approx\; 8\pi^2 \int \mathrm{d}\ell_1\ell_1 \int \mathrm{d}\ell_2\ell_2 \int \mathrm{d}\ell_3\ell_3 \begin{pmatrix}\ell_1 & \ell_2 & \ell_3\\ 0 & 0 & 0\end{pmatrix}^2,$$

where $\int \mathrm{d}\Omega$ is the integral over the unit sphere, and where $Y^m_\ell(\boldsymbol{n})$ denotes the spherical harmonic function with $\boldsymbol{n}$ the unit normal vector on the sphere. We are concerned with the validity of this derivation for the following reasons: Terms with integer and real-valued $\ell$ are mixed, e.g. it remains unclear how the integration over the Wigner symbol squared is to be understood. To get from the second to the third equality, the Fourier base $\mathrm{e}^{\mathrm{i}\boldsymbol{\ell}\cdot\boldsymbol{\theta}}$ is expanded into spherical harmonics, an approximation which Hu (2000) correctly states to be valid for small angles only. However, the integration over angles runs over the full two-dimensional plane or the unit sphere, respectively. Moreover, it is not specified how the non-trivial transition from an integral over the plane to one over the unit sphere is executed. Instead of (31), we propose to use (6), which is an exact and rigorous expression.

To allow for a comparison between (31) and our approach based on (6), we need to establish a relation between the square of the Wigner symbol and (8). We refer to Borodin et al. (1978, see also references therein) who compute approximation formulae of the Wigner symbol in the context of the quasi-continuous limit of quantum states with high angular momenta. The base of their derivation is formed by the exact relation

$$\int_0^{2\pi}\mathrm{d}\varphi \int_0^{2\pi}\mathrm{d}\psi \int_0^{\pi}\mathrm{d}\theta\,\sin\theta\; D^{\ell_1}_{m_1m_1'}(\varphi,\theta,\psi)\;D^{\ell_2}_{m_2m_2'}(\varphi,\theta,\psi)\;D^{\ell_3}_{m_3m_3'}(\varphi,\theta,\psi) = 8\pi^2\begin{pmatrix}\ell_1 & \ell_2 & \ell_3\\ m_1 & m_2 & m_3\end{pmatrix}\cdot\begin{pmatrix}\ell_1 & \ell_2 & \ell_3\\ m_1' & m_2' & m_3'\end{pmatrix}, \tag{32}$$

where $D^\ell_{mm'}$ denotes the $m\times m'$ element of the Wigner D matrix, which in turn is a function of the three Euler angles $\varphi$, $\theta$, and $\psi$. Making use of a quasi-classical approximation of the $D^\ell_{mm'}$, Borodin et al. (1978) compute expressions for the general Wigner symbol in the limit of large and continuous angular frequencies. From these results we extract the approximation

$$\begin{aligned}\begin{pmatrix}\ell_1 & \ell_2 & \ell_3\\ 0 & 0 & 0\end{pmatrix}^2 \;\approx\; & \frac{2}{\pi}\left\{2\left(\ell_1+\frac{1}{2}\right)^2\left(\ell_2+\frac{1}{2}\right)^2 + 2\left(\ell_2+\frac{1}{2}\right)^2\left(\ell_3+\frac{1}{2}\right)^2\right.\\ & \left. +\; 2\left(\ell_3+\frac{1}{2}\right)^2\left(\ell_1+\frac{1}{2}\right)^2 - \left(\ell_1+\frac{1}{2}\right)^4 - \left(\ell_2+\frac{1}{2}\right)^4 - \left(\ell_3+\frac{1}{2}\right)^4\right\}^{-1/2},\end{aligned} \tag{33}$$

which allows us to generalize the Wigner symbol to real-valued arguments. Equation (33) holds only for $\ell_1,\ell_2,\ell_3 \gg 1$, which, in the quantum-mechanical context of Borodin et al. (1978), originates from the use of expressions that are valid for large angular momenta, i.e. the quasi-classical limit, only. This condition on angular frequencies also underlies the approximations in (29) and (31) and can in our context be interpreted as a natural consequence of working in the flat-sky approximation.

As is demonstrated in Fig. 4, we find that (33) constitutes an excellent approximation, whose accuracy over a wide range of $\ell$-values is orders of magnitude better than the approximation given in Takada & Jain (2004), Eq. (A3),

$$\begin{aligned}\begin{pmatrix}\ell_1 & \ell_2 & \ell_3\\ 0 & 0 & 0\end{pmatrix}^2 \;\approx\; & \frac{\mathrm{e}^3}{\sqrt{2}\,\pi}\,(L+2)^{-1/2}\left(\frac{L}{2}-\ell_1+1\right)^{-1/2}\left(\frac{L}{2}-\ell_2+1\right)^{-1/2}\left(\frac{L}{2}-\ell_3+1\right)^{-1/2}\\ & \times\left(\frac{L/2-\ell_1+1/2}{L/2-\ell_1+1}\right)^{L-2\ell_1+1/2}\left(\frac{L/2-\ell_2+1/2}{L/2-\ell_2+1}\right)^{L-2\ell_2+1/2}\left(\frac{L/2-\ell_3+1/2}{L/2-\ell_3+1}\right)^{L-2\ell_3+1/2}.\end{aligned} \tag{34}$$

Only for triangle configurations close to degeneracy does the latter formula perform slightly better. Both approximation formulae are least accurate in the case of a degenerate triangle configuration with fractional errors around 10% or slightly above, but improve quickly to very small percentage deviations when the configuration approaches a more equilateral form. In Fig. 4 we also plot the fractional errors as a function of the triangle area enclosed by the three angular frequency vectors and as a function of the internal angle $\Phi_3$ opposite $\ell_3$, being the longest side of the triangle in the configurations considered. In terms of these quantities we observe a more universal behavior of the errors, in particular in the regime where the approximations are less accurate. We find to good approximation that, when approaching the degenerate case, relative errors increase exponentially with decreasing triangle area and increasing $\Phi_3$.

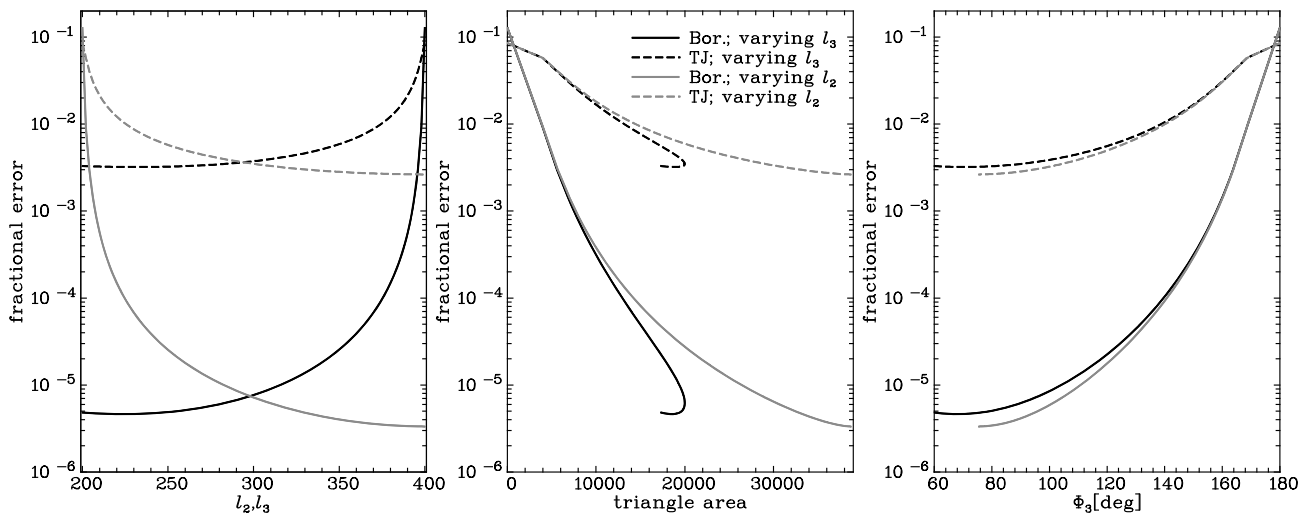


**Fig. 4.** Fractional error of the approximation formulae for the Wigner symbol. *Left panel*: shown are the relative deviations of (33) and (34) from the true absolute value of the Wigner symbol. The same triangle configurations as in Fig. 2 are used. Results for $\ell_1 = 200, \ell_3 = 400$ and varying $\ell_2$ are shown in gray while those corresponding to $\ell_1, \ell_2 = 200$ and varying $\ell_3$ are plotted in black. Solid curves are obtained using (33), dashed curves by employing (34). *Center panel*: same as above, but now plotting on the abscissa the corresponding triangle area enclosed by the three angular frequency vectors. Note that before reaching the equilateral configuration, the area has a maximum and starts to decrease again. *Right panel*: same as above, but now as a function of the internal angle $\Phi_3$ opposite $\ell_3$, which is the longest side of the triangle in both configurations considered. Hence, $\Phi_3 = 60°$ corresponds to the equilateral case, and $\Phi_3 = 180°$ to the degenerate case.

For $\ell \gg 1$, and if the triangle configuration is not too close to the degenerate case, one may approximate $\ell_i + 1/2 \approx \ell_i$, so that one finds from (8) and (33)

$$\begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ 0 & 0 & 0 \end{pmatrix}^2 \approx \frac{\Lambda(\ell_1, \ell_2, \ell_3)}{2\pi} \,. \tag{35}$$

Remarkably, since for integer angular frequencies we have $\Delta\ell_1 = \Delta\ell_2 = \Delta\ell_3 = 1$, (33) exactly reproduces our earlier conjecture (13), which strongly supports its validity. If one replaces the Wigner symbol in (31) by (35), however, one obtains a result which is a factor of 2 larger compared to (6).

Inserting (35) into (30), and using $2\ell + 1 \approx 2\ell$ for $\ell \gg 1$, we get

$$\left\langle \hat{B}(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3)\, \hat{B}(\bar{\ell}_4, \bar{\ell}_5, \bar{\ell}_6) \right\rangle \approx \frac{2\pi^2\, D_{\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3,\bar{\ell}_4,\bar{\ell}_5,\bar{\ell}_6}}{A\, \bar{\ell}_1\, \bar{\ell}_2\, \bar{\ell}_3} \begin{pmatrix} \bar{\ell}_1 & \bar{\ell}_2 & \bar{\ell}_3 \\ 0 & 0 & 0 \end{pmatrix}^{-2} P(\bar{\ell}_1)\, P(\bar{\ell}_2)\, P(\bar{\ell}_3), \tag{36}$$

which is equivalent to (23) if the latter equation is specified to $\Delta\ell_1 = \Delta\ell_2 = \Delta\ell_3 = 1$, and integer $\ell$ with $L$ even – except for (36) being a factor of 2 smaller. In the following, we are going to elaborate on this apparent discrepancy.

### 5.2. Parity

To elucidate the different noise properties of the Fourier-plane and spherical harmonic bispectrum estimators, we investigate their behavior with respect to parity. In two dimensions the parity transformation corresponds to an axis reflection, or equivalently, the reversal of the polar angle of all spatial vectors. To flip the parity of a triangle, one can do an odd permutation of its sides, see e.g. the two triangles sketched in Fig. 1. Hence, to test the behavior of estimators for triangles of different parity, it is sufficient to flip any two of its angular frequency arguments.

Consulting (27), we find

$$\hat{B}_{\ell_1,\ell_3,\ell_2} = (-1)^{\ell_1+\ell_2+\ell_3}\, \hat{B}_{\ell_1,\ell_2,\ell_3} \tag{37}$$

because of the behavior of the Wigner symbol under change of parity,

$$\begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ m_1 & m_2 & m_3 \end{pmatrix} = (-1)^{\ell_1+\ell_2+\ell_3} \begin{pmatrix} \ell_1 & \ell_3 & \ell_2 \\ m_1 & m_3 & m_2 \end{pmatrix}, \tag{38}$$

and likewise for all odd permutations of the columns in the Wigner symbol. Thus, the spherical harmonics estimator is parity-invariant for $L$ even and changes sign for $L$ odd. Most cosmological theories predict parity-invariant large-scale structures and CMB anisotropies. If parity symmetry is built into the cosmological model at consideration, measures that vary under parity transformations do not have any predictive power, wherefore they are usually not considered in a data analysis. Accordingly, (27) is only used for arguments that have $L$ even. Note that parity invariance is also incorporated into the relation between the spherical harmonics and Fourier-plane bispectra, see the second equality of (29), via the Wigner symbol which vanishes for $L$ odd (this behavior is a direct consequence of (38) for $m_1 = m_2 = m_3 = 0$).

The Fourier-plane estimator is by design parity-invariant, which can be seen mathematically from swapping arguments of (3), or illustratively by inspecting Fig. 1. From the sketch it is evident that triangle configurations of different parity are averaged over with equal weight. For a more formal argument, we can explicitly construct estimators that average only over triangle configurations of the same parity. To this end, consider the two-dimensional cross product $\boldsymbol{a}\times\boldsymbol{b} = a_x b_y - a_y b_x$ (Schneider & Lombardi 2003) of the angular frequency vectors $\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, \boldsymbol{\ell}_3$. If they form a triangle, one finds $\boldsymbol{\ell}_1\times\boldsymbol{\ell}_2 = \boldsymbol{\ell}_2\times\boldsymbol{\ell}_3 = \boldsymbol{\ell}_3\times\boldsymbol{\ell}_1$, which follows from $\boldsymbol{\ell}_1+\boldsymbol{\ell}_2+\boldsymbol{\ell}_3 = 0$. A change in the parity of the triangle implies a sign flip in these cross products.

Noting that $\boldsymbol{\ell}_i\times\boldsymbol{\ell}_j = \ell_i\ell_j\,\sin(\varphi_{\ell_j}-\varphi_{\ell_i})$, we compute a condition on the polar angles,

$$\varphi_{\ell_2}-\varphi_{\ell_1}\in[0,\pi]\,;\quad \varphi_{\ell_3}-\varphi_{\ell_2}\in[0,\pi]\,;\quad \varphi_{\ell_1}-\varphi_{\ell_3}\in[0,\pi]\,. \tag{39}$$

To obtain the parity transformed triangle, swap the signs of the polar angles in (39). Under the premise that the vectors do form a triangle, one of the conditions in (39) is redundant, the remaining ones restricting the angular integrations in the averaging of (3). For instance, the integration ranges could be modified to $\varphi_{\ell_1}\in[0,2\pi]$, $\varphi_{\ell_2}\in[\varphi_{\ell_1},\pi+\varphi_{\ell_1}]$, and $\varphi_{\ell_3}\in[\varphi_{\ell_1}-\pi,\varphi_{\ell_1}]$. Due to rotational symmetry, which still holds, the inner integrals have to yield the same result for all possible values of $\varphi_{\ell_1}$. Therefore, we can set the ranges of the inner integrals to $\varphi_{\ell_2}\in[0,\pi]$ and $\varphi_{\ell_3}\in[-\pi,0]$ without loss of generality. To maintain the symmetry, we keep the integral over $\varphi_{\ell_1}$ in our notation. These findings are reflected in the shorthand notation

$$\int \mathrm{d}\{\varphi_1,\varphi_2,\varphi_3\} \equiv \int_0^{2\pi}\frac{\mathrm{d}\varphi_{\ell_1}}{2\pi}\int_0^{\pi}\frac{\mathrm{d}\varphi_{\ell_2}}{\pi}\int_{-\pi}^{0}\frac{\mathrm{d}\varphi_{\ell_3}}{\pi}, \tag{40}$$

which we use to define the following bispectrum estimators,

$$\begin{aligned}\hat{B}_\Delta(\bar\ell_1,\bar\ell_2,\bar\ell_3) \;=\; &\frac{2\pi^2}{A}\Lambda^{-1}\left(\bar\ell_1,\bar\ell_2,\bar\ell_3\right)\int_{\bar\ell_1-1/2\Delta\ell}^{\bar\ell_1+1/2\Delta\ell}\frac{\mathrm{d}\ell_1\ell_1}{\bar\ell_1\Delta\ell_1}\int_{\bar\ell_2-1/2\Delta\ell}^{\bar\ell_2+1/2\Delta\ell}\frac{\mathrm{d}\ell_2\ell_2}{\bar\ell_2\Delta\ell_2}\int_{\bar\ell_3-1/2\Delta\ell}^{\bar\ell_3+1/2\Delta\ell}\frac{\mathrm{d}\ell_3\ell_3}{\bar\ell_3\Delta\ell_3}\\ &\times\frac{1}{3}\left\{\int\mathrm{d}\{\varphi_1,\varphi_2,\varphi_3\}+\int\mathrm{d}\{\varphi_2,\varphi_3,\varphi_1\}+\int\mathrm{d}\{\varphi_3,\varphi_1,\varphi_2\}\right\}\,\delta_{\mathrm{D}}^{(2)}(\boldsymbol{\ell}_1+\boldsymbol{\ell}_2+\boldsymbol{\ell}_3)\,g(\boldsymbol{\ell}_1)\,g(\boldsymbol{\ell}_2)\,g(\boldsymbol{\ell}_3);\\ \hat{B}_\nabla(\bar\ell_1,\bar\ell_2,\bar\ell_3) \;=\; &\frac{2\pi^2}{A}\Lambda^{-1}\left(\bar\ell_1,\bar\ell_2,\bar\ell_3\right)\int_{\bar\ell_1-1/2\Delta\ell}^{\bar\ell_1+1/2\Delta\ell}\frac{\mathrm{d}\ell_1\ell_1}{\bar\ell_1\Delta\ell_1}\int_{\bar\ell_2-1/2\Delta\ell}^{\bar\ell_2+1/2\Delta\ell}\frac{\mathrm{d}\ell_2\ell_2}{\bar\ell_2\Delta\ell_2}\int_{\bar\ell_3-1/2\Delta\ell}^{\bar\ell_3+1/2\Delta\ell}\frac{\mathrm{d}\ell_3\ell_3}{\bar\ell_3\Delta\ell_3}\\ &\times\frac{1}{3}\left\{\int\mathrm{d}\{\varphi_1,\varphi_3,\varphi_2\}+\int\mathrm{d}\{\varphi_2,\varphi_1,\varphi_3\}+\int\mathrm{d}\{\varphi_3,\varphi_2,\varphi_1\}\right\}\,\delta_{\mathrm{D}}^{(2)}(\boldsymbol{\ell}_1+\boldsymbol{\ell}_2+\boldsymbol{\ell}_3)\,g(\boldsymbol{\ell}_1)\,g(\boldsymbol{\ell}_2)\,g(\boldsymbol{\ell}_3).\end{aligned} \tag{41}$$

Here, we have symmetrized the restricted integrations (40) by averaging over all either even or odd permutations of $\{\varphi_{\ell_1},\varphi_{\ell_2},\varphi_{\ell_3}\}$. Consequently, changing parity via any odd permutation of the angular frequencies in the arguments of (41) turns one estimator into the other, as demanded, for instance $\hat{B}_\Delta(\bar\ell_1,\bar\ell_3,\bar\ell_2) = \hat{B}_\nabla(\bar\ell_1,\bar\ell_2,\bar\ell_3)$.

Note that the prefactor of the estimators in (41) is diminished by a factor of 2 with respect to (3), which is necessary to keep them unbiased. This can be shown by computing the expectation value of (41) in close analogy to the procedure outlined in Sect. 2. However, the separate consideration of angular and radial integrals that enabled us to make use of (6) is not possible anymore in this non-symmetric case. For instance, given fixed $\boldsymbol{\ell}_1$, the restricted angular integrations (40) can still produce a triangle of opposite parity by including a triangle with $|\boldsymbol{\ell}_2'| = \ell_3$ and $|\boldsymbol{\ell}_3'| = \ell_2$. This is reflected in the fact that the integration (6), if properly normalized[1], still yields the same result when limiting the length of the integration range to $\pi$.

Instead, one can execute the integral over the angular frequency which is still averaged over the full two-dimensional plane, such as

$$\begin{aligned}&\int_{\bar\ell_1-1/2\Delta\ell}^{\bar\ell_1+1/2\Delta\ell}\frac{\mathrm{d}\ell_1\ell_1}{\bar\ell_1\Delta\ell_1}\int_{\bar\ell_2-1/2\Delta\ell}^{\bar\ell_2+1/2\Delta\ell}\frac{\mathrm{d}\ell_2\ell_2}{\bar\ell_2\Delta\ell_2}\int_{\bar\ell_3-1/2\Delta\ell}^{\bar\ell_3+1/2\Delta\ell}\frac{\mathrm{d}\ell_3\ell_3}{\bar\ell_3\Delta\ell_3}\int_0^{2\pi}\frac{\mathrm{d}\varphi_{\ell_1}}{2\pi}\int_0^{\pi}\frac{\mathrm{d}\varphi_{\ell_2}}{\pi}\int_{-\pi}^{0}\frac{\mathrm{d}\varphi_{\ell_3}}{\pi}\,\delta_{\mathrm{D}}^{(2)}(\boldsymbol{\ell}_1+\boldsymbol{\ell}_2+\boldsymbol{\ell}_3) =\\ &\frac{1}{2\pi\bar\ell_1\Delta\ell_1}\int_{\bar\ell_2-1/2\Delta\ell}^{\bar\ell_2+1/2\Delta\ell}\frac{\mathrm{d}\ell_2\ell_2}{\bar\ell_2\Delta\ell_2}\int_{\bar\ell_3-1/2\Delta\ell}^{\bar\ell_3+1/2\Delta\ell}\frac{\mathrm{d}\ell_3\ell_3}{\bar\ell_3\Delta\ell_3}\int_0^{\pi}\frac{\mathrm{d}\varphi_{\ell_2}}{\pi}\int_{-\pi}^{0}\frac{\mathrm{d}\varphi_{\ell_3}}{\pi}\,\mathbf{1}_{\boldsymbol{\ell}_1,\boldsymbol{\ell}_2,\boldsymbol{\ell}_3},\end{aligned} \tag{42}$$

where $\mathbf{1}_{\boldsymbol{\ell}_1,\boldsymbol{\ell}_2,\boldsymbol{\ell}_3} = 1$ if $\boldsymbol{\ell}_1,\boldsymbol{\ell}_2,\boldsymbol{\ell}_3$ form a triangle, and 0 else. The remaining integrations reproduce the overlapping region of the annuli for $\boldsymbol{\ell}_2$ and $\boldsymbol{\ell}_3$, as depicted in Fig. 1. By limiting the integration to the half plane to one side of an axis collinear to $\boldsymbol{\ell}_1$, the overlap is obviously halved. Since the area of the annuli for $\boldsymbol{\ell}_2$ and $\boldsymbol{\ell}_3$ is also reduced by half each, the value of the integration should double, see (11). Following the geometrical interpretation once again, we thus arrive at

$$\begin{aligned}&\int_{\bar\ell_1-1/2\Delta\ell}^{\bar\ell_1+1/2\Delta\ell}\frac{\mathrm{d}\ell_1\ell_1}{\bar\ell_1\Delta\ell_1}\int_{\bar\ell_2-1/2\Delta\ell}^{\bar\ell_2+1/2\Delta\ell}\frac{\mathrm{d}\ell_2\ell_2}{\bar\ell_2\Delta\ell_2}\int_{\bar\ell_3-1/2\Delta\ell}^{\bar\ell_3+1/2\Delta\ell}\frac{\mathrm{d}\ell_3\ell_3}{\bar\ell_3\Delta\ell_3}\int_0^{2\pi}\frac{\mathrm{d}\varphi_{\ell_1}}{2\pi}\int_0^{\pi}\frac{\mathrm{d}\varphi_{\ell_2}}{\pi}\int_{-\pi}^{0}\frac{\mathrm{d}\varphi_{\ell_3}}{\pi}\,\delta_{\mathrm{D}}^{(2)}(\boldsymbol{\ell}_1+\boldsymbol{\ell}_2+\boldsymbol{\ell}_3)\,B(\ell_1,\ell_2,\ell_3)\approx\\ &\frac{1}{2\pi^2}\,\Lambda(\bar\ell_1,\bar\ell_2,\bar\ell_3)\,B(\bar\ell_1,\bar\ell_2,\bar\ell_3).\end{aligned} \tag{43}$$

Comparing this result to (4), the estimators (41) have indeed to be smaller by a factor of 2 to still be unbiased.

[1] In the derivation of Sect. 2 the proper normalization of $2\pi$ for each angular integral is hidden within $A_R(\ell)$. Note that we have given (6) without this normalization, whereas it is included in (11).

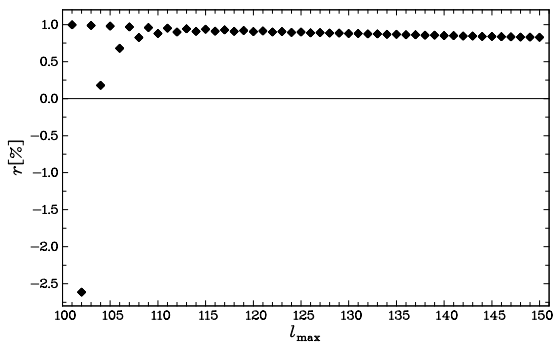


**Fig. 5.** Comparison of the Fisher information as obtained by spherical harmonics and Fourier-plane approach. Given is the relative deviation $r$ as a function of the maximum angular frequency $\ell_{\max}$ used in (45).

To obtain bispectrum estimators that are completely analogous to (27), we define

$$\hat{B}_{\pm}(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3) \equiv \frac{1}{2} \left( \hat{B}_{\Delta}(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3) \pm \hat{B}_{\nabla}(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3) \right). \tag{44}$$

As $\left\langle \hat{B}_{-}(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3) \right\rangle = 0$ for a parity symmetric random field $g$, and $\hat{B}_{-}(\bar{\ell}_1, \bar{\ell}_3, \bar{\ell}_2) = -\hat{B}_{-}(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3)$, this estimator shows identical behavior compared to $\hat{B}_{\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3}$ with $L$ odd. In practice both measures could be used to assess deviations from parity symmetry. The estimators $\hat{B}_{+}(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3)$ and $\hat{B}_{\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3}$ with $L$ even are likewise invariant under parity transformations. After some algebra that closely follows the outline of Sect. 4 we find that the covariance of $\hat{B}_{+}$ is the same as (26), which is not unexpected because we already noted that (3) is also parity-symmetric.

With (44) at hand, one can readily extract the different treatment of even and odd parity measures in the spherical harmonic and Fourier-plane formalisms. Estimators (27) separate the set of possible arguments $\{\ell_1, \ell_2, \ell_3\}$ disjointly into parity even ($L$ even) and parity odd ($L$ odd), whereas $\hat{B}_{+}$ and $\hat{B}_{-}$ are defined on the same full set of angular frequency combinations[2]. In other words, when limiting $\hat{B}_{+}$ to integer angular frequencies only, the same information is contained in "half" the number of measures in the spherical harmonics case, namely those with $L$ even. The latter estimators have a covariance of half the size of the covariance of $\hat{B}_{+}$, so that the overall information content is the same for both approaches – as required.

### 5.3. Information content

We verify the findings of the foregoing section by comparing the information contained in both approaches in terms of the Fisher matrix (Tegmark et al. 1997). For a practical implementation we specialize to a non-tomographic weak lensing survey (see e.g. Bartelmann & Schneider 2001, for an overview), assuming a cosmology-independent covariance that is well approximated by the Gaussian approximation, i.e. using (23) and (30), respectively. To allow for direct comparison, we limit the Fourier-plane approach to integer $\ell$ with all bin sizes set to unity. Due to the symmetry under permutations of the arguments of the bispectra, one can impose the condition $\ell_1 \leq \ell_2 \leq \ell_3$ on both formalisms, rendering a block-wise diagonal covariance matrix. Inspecting (23), the only dependence on the arguments of the second bispectrum, i.e. $\ell_4$ to $\ell_6$, is due to the Kronecker symbols (21), so that the summations over $\ell_4$ to $\ell_6$ become trivial.

Hence, the Fisher matrix can be written as

$$F_{\mu\nu} = \sum_{l_{\min} \leq \ell_1 \leq \ell_2 \leq \ell_3 \leq l_{\max}} D_{\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3,\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3} \, \frac{\partial B(\ell_1, \ell_2, \ell_3)}{\partial p_\mu} \, \frac{A \; \bar{\ell}_1 \bar{\ell}_2 \bar{\ell}_3 \; \Delta\ell_1 \Delta\ell_2 \Delta\ell_3 \; \Lambda\left(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3\right)}{(2\,\pi)^3 \; P(\bar{\ell}_1) P(\bar{\ell}_2) P(\bar{\ell}_3)} \, \frac{\partial B(\ell_1, \ell_2, \ell_3)}{\partial p_\nu}, \tag{45}$$

where $D_{\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3,\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3} = 6$ for equilateral triangles, $D_{\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3,\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3} = 2$ for isosceles, and $D_{\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3,\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3} = 1$ else. The derivatives are taken with respect to a set of cosmological parameters $\boldsymbol{p}$. In this toy example we use only the single parameter $\Omega_{\rm m}$, reducing the Fisher matrix to a scalar $F$. Besides, we restrict the angular frequency values to an unphysically small range between $l_{\min} = 100$ and $l_{\max} = 150$ for computational reasons.

Weak lensing power spectra are computed for a standard ΛCDM cosmology, including non-linear evolution via the fit formula of Smith et al. (2003). The bispectra are obtained via perturbation theory (e.g. Fry 1984), using Scoccimarro & Couchman (2001) with the definition of the non-linear wave vector by Takada & Jain (2004) to account for non-linear evolution. For the projections along the line of sight we assume a redshift probability distribution according to Smail et al. (1994) with $\beta = 1.5$ and a deep survey of 0.9 median redshift. Shape noise is incorporated by replacing the power spectra in the covariances with

$$\bar{P}(\ell) = P(\ell) + \frac{\sigma_\epsilon^2}{2\bar{n}}, \tag{46}$$

where the ellipticity dispersion $\sigma_\epsilon = 0.35$ and the galaxy number density $n = 40\,\mathrm{arcmin}^{-2}$ are set to typical values for planned space-based surveys.

[2] A similar behavior as for the spherical harmonic estimators would have been unexpected since the possible arguments of $\hat{B}_{\pm}$ form a non-countable set.

We calculate the relative deviation of the Fisher information, $r \equiv F_{\text{Fourier}}/F_{\text{sph. harm.}} - 1$, as a function of $l_{\max}$. Note that, since we only consider ratios of $F$, the survey size $A$ drops out. Our results are shown in Fig. 5. For $l_{\max}$ very close to $l_{\min}$ one sees alternating jumps in $r$ which can mostly be traced back to the fact that, due to the condition $\ell_1 \leq \ell_2 \leq \ell_3$, the terms entering (45) do not always split exactly half into $L$ even and odd. After this "burn in" for $l_{\max} \lesssim 120$, $r$ shows only little variation. The remaining offset from zero, which is slowly decreasing, can entirely be assigned to the different prefactors in the covariances, i.e. the terms related to the Wigner symbol and $\Lambda$, respectively. The range of angular frequencies plotted in Fig. 5 is still far from any physically relevant situation, but nonetheless the two approaches agree already better than 99%.

## 6. Conclusions

In this work we intended to give insight into the derivation and the form of the bispectrum covariance in the flat-sky approximation, based exclusively on the two-dimensional Fourier formalism. We defined an unbiased estimator that takes the average over the overlap of annuli in Fourier space, and computed its covariance. To obtain precise normalizations, a case distinction is necessary between degenerate and non-degenerate triangle configurations. However, given that both normalizations become very similar for $\ell \gg 1$, which is assumed in the flat-sky approach anyway, we suggest as a simple and fair approximation to use the expression derived for the degenerate case. Then our result for the Gaussian part of the bispectrum covariance reads

$$\mathrm{Cov}\Big(B(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3),\ B(\bar{\ell}_4,\bar{\ell}_5,\bar{\ell}_6)\Big)_{\text{Gauss}} = \frac{(2\pi)^3\, D_{\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3,\bar{\ell}_4,\bar{\ell}_5,\bar{\ell}_6}}{A\ \bar{\ell}_1\bar{\ell}_2\bar{\ell}_3\ \Delta\ell_1\Delta\ell_2\Delta\ell_3}\ \Lambda^{-1}\left(\bar{\ell}_1+\frac{\Delta\ell_1}{2},\bar{\ell}_2+\frac{\Delta\ell_2}{2},\bar{\ell}_3+\frac{\Delta\ell_3}{2}\right)\ P(\bar{\ell}_1)P(\bar{\ell}_2)P(\bar{\ell}_3). \tag{47}$$

This formula is readily generalized to the total covariance by modifying the arguments of $\Lambda$, appearing in the non-Gaussian terms of (26), accordingly. It is directly applicable to any real values of angular frequencies, to arbitrary binning, and to any compact, finite survey geometry. This formula can be modified to incorporate shot or shape noise, as well as to account for photometric redshift information or CMB polarization in a straightforward manner.

While the general form of our result was in agreement with existing work, we found, contrary to Hu (2000), that the size of the covariance is a factor of 2 larger than the one obtained by the flat-sky spherical harmonic approach. By defining parity-sensitive bispectrum estimators, we discussed the behavior of both formalisms with respect to parity transformations, arguing that the difference in the covariances is indeed to be expected because in the spherical harmonic framework, parity-invariant measures are restricted to a subset of the angular frequency combinations at which the bispectra are evaluated. In a practical example we demonstrated that both approaches indeed contain the same information in terms of the Fisher matrix, with a high level of agreement. As a consequence, we can confirm that studies performed in the flat-sky spherical harmonic approach, such as Takada & Jain (2004), yield correct parameter constraints as long as the analysis is restricted to integer $\ell$ with the sum of the three angular frequencies being even.

We established a relation between the geometrical and intuitive process of averaging over the overlapping regions of annuli in the Fourier plane and the Wigner symbol of the spherical harmonic approach. Both quantities were demonstrated to be in turn connected to a simple measure that is proportional to the size of the area enclosed by the triangle configuration for which the bispectrum is calculated. This resulted in convenient, yet precise approximation formulae for the prefactors of the covariances of both approaches at consideration.

Under the assumption of a compact survey geometry and scales much smaller than the extent of the survey area, (26) provides a cleanly derived bispectrum covariance matrix that naturally incorporates the scaling with survey size, is not restricted to integer angular frequencies, and allows for any appropriate binning.

*Acknowledgements.* The authors would like to thank Bhuvnesh Jain and Masahiro Takada for helpful discussions and the referee for a helpful report. We thank Joel Bergé for comparison tests of our bispectrum codes. B.J. is grateful to Sarah Bridle for kind hospitality at UCL. B.J. acknowledges support by the Deutsche Telekom Stiftung and the Bonn-Cologne Graduate School of Physics and Astronomy. X.S. is supported by the International Max Planck Research School (IMPRS) for Astronomy and Astrophysics. This work was supported by the DFG under the Priority Programme 1177 Galaxy Evolution and within the Transregional Collaborative Research Centre TR33 The Dark Universe.

## References

Bartelmann, M., & Schneider, P. 2001, Phys. Rep., 340, 291
Bernardeau, F., Colombi, S., Gaztañaga, E., et al. 2002a, Phys. Rep., 367, 1
Bernardeau, F., Mellier, Y., & van Waerbeke, L. 2002b, A&A, 389, L28
Borodin, K., Kroshilin, A., & Tolmachev, V. 1978, TMF, 34, 110
Cooray, A., Sarkar, D., & Serra, P. 2008, Phys. Rev. D, 77, 123006
Fry, J. 1984, ApJ, 279, 449
Gradshteyn, I., Ryzhik, I., Jeffrey, A., et al. 2000, Tables of Integrals, Series, and Products (Academic Press)
Hu, W. 1999, ApJ, 522, 21
Hu, W. 2000, Phys. Rev. D, D62, 043007
Jarvis, M., Bernstein, G., & Jain, B. 2004, MNRAS, 352, 338
Joachimi, B., Schneider, P., & Eifler, T. 2008, A&A, 477, 43
Kaiser, N. 1998, ApJ, 498, 26
Matarrese, S., Verde, L., & Heavens, A. 1997, MNRAS, 290, 651
Schneider, P., & Lombardi, M. 2003, A&A, 397, 809
Scoccimarro, R. 2000, ApJ, 544, 597
Scoccimarro, R., & Couchman, H. 2001, MNRAS, 325, 1312
Scoccimarro, R., Feldman, H., Fry, J., et al. 2001, ApJ, 546, 652
Sefusatti, E., Crocce, M., Pueblas, S., et al. 2006, Phys. Rev. D, 74, 023522
Smail, I., Ellis, R., & Fitchett, M. 1994, MNRAS, 270, 245
Smith, R., Peacock, J., Jenkins, A., et al. 2003, MNRAS, 341, 1311
Takada, M., & Jain, B. 2004, MNRAS, 348, 897
Tegmark, M., Taylor, A., & Heavens, A. 1997, ApJ, 480, 22