

Best unbiased estimators for the three-point correlators of the cosmic microwave background radiation

Alejandro Gangui*

*Instituto de Astronomía y Física del Espacio, Ciudad Universitaria, 1428 Buenos Aires, Argentina;
Departamento de Física, Universidad de Buenos Aires, Ciudad Universitaria–Pab. 1, 1428 Buenos Aires, Argentina;
and DARC, Observatoire de Paris-Meudon, UMR 8629 CNRS, 92195 Meudon Cedex, France*

Jérôme Martin†

*DARC, Observatoire de Paris-Meudon, UMR 8629 CNRS, 92195 Meudon Cedex, France
(Received 21 January 2000; published 9 October 2000)*

Measuring the three-point correlators of the cosmic microwave background (CMB) anisotropies could help to get a handle on the level of non-Gaussianity present in the observational data sets and therefore would strongly constrain models of the early Universe. However, typically, the expected non-Gaussian signal is very small. Therefore, one has to face the problem of extracting it from the noise, in particular from the “cosmic variance” noise. For this purpose, one has to construct the best unbiased estimators for the three-point correlators that are needed for concrete detection of non-Gaussian features. In this article, we study this problem for both the CMB third moment and the CMB angular bispectrum. We emphasize that knowledge of the best estimator for the former does not permit one to infer the best estimator for the latter and vice versa. We present the corresponding best unbiased estimators in both cases and compute their corresponding cosmic variances.

PACS number(s): 98.70.Vc

I. INTRODUCTION

The cosmic microwave background (CMB) has been recognized as one of the best tools for studying the early Universe (e.g., [1]). In particular, the statistical properties of the CMB anisotropies are a powerful means to discriminate among the possible scenarios. This is because, in general, different models predict different statistical properties. For example, the simplest models of inflation predict that the temperature anisotropies should obey a Gaussian statistics and therefore any nonvanishing measurement of a three-point correlator (in a sense to be made precise below) would automatically rule out such models, a very interesting result indeed.

From a practical point of view, measuring any non-Gaussianity in the data is a very difficult task since the signal is typically very small. Of course, this signal should be compared to the noise and what really matters is the signal to noise ratio. The noise can have many different origins including instrumental errors, foregrounds contamination, or incomplete sky coverage. Another source of error is the so-called “cosmic variance.” Roughly speaking, it comes from the fact that we only have access to one realization of the temperature anisotropies whereas theoretical predictions are expressed through ensemble averages. In a Gaussian model, for example, the mean value of any three-point correlator has to vanish but this does not guarantee that a concrete detection of a nonzero signal on the sky would be in contradiction with the model [2,3]. The important point is that the cosmic variance can dominate the other sources of error, as this is in

fact the case for the two-point correlators on large angular scales. Therefore, if one wants to unveil non-Gaussianity, it is necessary to address the cosmic variance problem for the three-point correlators. The usual way to deal with this problem is to construct estimators by performing spatial averages on the celestial sphere and to find the one which has the smallest possible variance. The aim of this paper is then to find the best unbiased estimators both for the third moment $\langle a_{l_1}^{m_1} a_{l_2}^{m_2} a_{l_3}^{m_3} \rangle$ and for the angular bispectrum $C_{l_1 l_2 l_3}$ and to display the corresponding cosmic variances.

Recently, there has been a lot of activity in the subject triggered by the finding that non-Gaussianities are present in the 4-yr Cosmic Background Explorer (COBE) Differential Microwave Radiometer (DMR) data [4,5]. Further analyses have confirmed this result (e.g., [6]). However, soon after, it was demonstrated by [7] that the non-Gaussian signal is driven by the 53 GHz data. This systematic artifact in the CMB maps rejects a possible cosmological origin. More generally, it is clear that the presence of foregrounds [8,9] renders difficult the detection of a genuine non-Gaussian signal. Nevertheless, one should expect non-Gaussian features to be present in the CMB anisotropy data sets. These could be produced in the early Universe during inflation either because the initial conditions are non-Gaussian themselves (i.e., the quantum initial state is not the vacuum [10,11]) or owing to the existence of couplings between different perturbation modes at the nonlinear level [12,13,14]. In the context of slow-roll inflation, the CMB bispectrum has recently been studied in [15,16]. Even if non-Gaussianities are not primordial in origin, they will nevertheless arise during later stages of evolution. In this context, the Rees-Sciama effect will build up a small but nonvanishing signal [17,18,19]. Also, cosmic topological defects of the vacuum, such as strings and textures, are among the best motivated sources for non-

*Email address: gangui@irupé.obspm.fr

†Email address: martin@edelweiss.obspm.fr

Gaussian features [20–24]. Regarding secondary sources, Goldberg and Spergel [25,26] have recently calculated the angular bispectrum due to second order gravitational effects such as the correlation of lensing of CMB photons and secondary anisotropies coming from the integrated Sachs-Wolfe effect and thermal Sunyaev-Zel’dovich effect. In the same line, Cooray and Hu [27] have taken into account further additional contributions to the bispectrum in the presence of re-ionization. Other approaches to the study of non-Gaussian features include preferred-direction statistics for sky maps [28], the three-point correlation function [29,30,12,31], lensing statistics [32–34], the genus and Euler-Poincaré statistics [35–37], peak statistics [38–40], correlation function of peaks [41], Minkowski functionals [42], and wavelet analyses [43,44].

This article is organized as follows. In the next section, the general strategy for finding best estimators is exposed. As a warm-up, in Sec. III, we implement this strategy for the two-point correlators. Section IV is the core of the article. There we explicitly derive, for the first time, the best unbiased estimator for the angular bispectrum and show its corresponding variance. Except for an overall normalization factor, this estimator turns out to be the one already employed by [4] and other authors recently. Our result places their choice on a firm basis. Next, we find the expression for the best unbiased estimator for the third moment. An earlier study was performed in [45]; however, our findings go beyond the results obtained in that article and, moreover, are explicit. Moreover, we present its corresponding variance. In addition, we also emphasize that the knowledge of the best estimator for the third moment does not allow one to infer the best estimator for the angular bispectrum and vice versa. In the last section, we briefly present our main conclusions. We finish up with a short Appendix which includes formulas related to the inverse two-point correlation function.

II. GENERAL STRATEGY FOR FINDING THE BEST ESTIMATOR

In this section, we expose the cosmic variance problem from the viewpoint of the theory of cosmological perturbations of quantum-mechanical origin and describe the method of the best unbiased estimators. This theory rests on the principles of general relativity and quantum field theory. At the beginning of the inflationary phase [46–49] the Friedmann-Lemaître-Robertson-Walker background spacetime already behaves classically whereas the excitations of the metric around this background are still quantum mechanical in nature. Technically, this means that the perturbed metric must be considered as a quantum operator. This operator either represents density perturbations or gravitational waves. In each case, the quantization can be carried out in a consistent way [50–57]. Then, the (zero-point) quantum fluctuations, which are the seeds of the cosmological perturbations, are amplified during inflation owing to the particle-creation phenomenon or squeezing effect [58]. Next, these primordial fluctuations give rise to the large scale structures and to the CMB anisotropies observed today in our Universe.

One should also discuss the choice of the quantum state in

which the metric operator is placed. Obviously, it is not possible to prepare the initial state of the Universe and therefore the choice of the quantum state of the perturbations is *a priori* free unless some theory of the initial conditions is provided (for example, quantum cosmology [59]). Usually, it is assumed that the initial state is the vacuum although different hypotheses are possible [60,10,11]. If the initial state is the vacuum, then the corresponding statistical properties are Gaussian. This is because the ground-state wave function of a harmonic oscillator is a Gaussian. It is possible to avoid this general conclusion either by considering nonlinear cosmological perturbations or by assuming that the initial state is a nonvacuum state. We have recently investigated the first possibility in [15]. The second possibility has been studied by [10]. In the latter case, non-Gaussianity is likely to be significant only for relatively small angular scales.

It should be emphasized that the mechanism described previously is deeply rooted in the quantum-mechanical nature of the gravitational field. The observable quantities calculated in this framework are always proportional to the Planck length. In other words, if observations confirm the full set of inflationary predictions then the fact that $\delta T/T \neq 0$ would be a direct observational consequence of quantum gravity.

The quantum-mechanical origin of the anisotropies in the framework of inflation raises also profound problems of interpretation. One should not think that these problems are purely theoretical. On the contrary, they have consequences with regards to the experimental strategy that one should follow in order to extract as much information as possible from the data. The fluctuations in the CMB effective temperature are linked to the perturbed metric as shown for the first time by Sachs and Wolfe. Therefore, the fact that the perturbed metric is an operator implies that the primordial fluctuations in the temperature must also be considered as a quantum operator. The observables are often expressed as n -point correlation functions of the operator $\hat{\Delta}(\mathbf{e}) \equiv \delta\hat{T}/T(\mathbf{e})$ in the arbitrary state $|\Psi\rangle$:

$$\xi_n(\mathbf{e}_1, \dots, \mathbf{e}_n) \equiv \langle \Psi | \hat{\Delta}(\mathbf{e}_1) \cdots \hat{\Delta}(\mathbf{e}_n) | \Psi \rangle, \quad (1)$$

where $\mathbf{e}_i$ ’s are arbitrary directions on the celestial sphere. In the following, we will also use the notation $\xi(\mathbf{e}_1 \cdot \mathbf{e}_2) \equiv \xi_2(\mathbf{e}_1, \mathbf{e}_2)$. According to the postulates of quantum mechanics, the previous theoretical predictions should be confronted to experiment in the following way. The same experiment should be performed N times, giving each time different outcomes q_i . If the quantity $(1/N) \sum_{i=1}^N q_i$ goes to the corresponding quantum expectation value when N goes to infinity, the theoretical prediction is said to be “compatible with experiment.” This is the core of the problem in cosmology: we only have access to one realization, i.e., one map of the CMB sky, and that means that N is fixed and equal to 1. Therefore, the question arises as to how we can verify the theoretical predictions of the theory of quantum-mechanical cosmological perturbations. This is a way of stating the cosmic variance problem. It is a fundamental limita-

tion in the sense that it remains even when other limitations like instrumental errors or low angular resolution have been fully mastered.

The usual method to deal with this problem is to replace quantum averages with spatial averages over the celestial sphere. Suppose we wish to measure $\xi_n(\mathbf{e}_1, \dots, \mathbf{e}_n)$. (Of course, the discussion could also be applied to quantities other than correlation functions.) The first step is to introduce a new operator, the estimator $\hat{\xi}(\xi_n)$ of ξ_n , defined as

$$\hat{\xi}(\xi_n) \equiv \int \dots \int d\Omega_1 \dots d\Omega_n E(\xi_n)(\mathbf{e}_1, \dots, \mathbf{e}_n) \hat{\Delta}(\mathbf{e}_1) \dots \hat{\Delta}(\mathbf{e}_n), \quad (2)$$

where $E(\xi_n)(\mathbf{e}_1, \dots, \mathbf{e}_n)$ is a weight function to be determined. Clearly, $\hat{\xi}(\xi_n)$ is defined through a spatial average. The second step is to require that the estimator is unbiased, i.e.,

$$\langle \Psi | \hat{\xi}(\xi_n) | \Psi \rangle = \xi_n. \quad (3)$$

In general, this restricts the class of functions $E(\xi_n)$ allowed. The fact that the mean value of the estimator be equal to the quantity we are seeking does not guarantee that each outcome will be for sure ξ_n . The third and final step is then to find the function $E(\xi_n)$ such that the variance (squared) of $\hat{\xi}(\xi_n)$, i.e.,

$$\sigma_{\hat{\xi}(\xi_n)}^2 = \langle \hat{\xi}(\xi_n) \hat{\xi}(\xi_n) \rangle - \langle \hat{\xi}(\xi_n) \rangle^2, \quad (4)$$

be as small as possible, taking into account the constraint given by Eq. (3). The corresponding estimator is then called the *best unbiased estimator*. Mathematically, this requirement is expressed through the following variation equation:

$$\delta(\sigma_{\hat{\xi}(\xi_n)}^2 - \lambda[\langle \Psi | \hat{\xi}(\xi_n) | \Psi \rangle - \xi_n]) = 0, \quad (5)$$

where one has introduced a Lagrange multiplier λ which can then be determined from the previous equation and the constraint itself. Once we have λ , we plug it into Eq. (5) and this completely fixes $E(\xi_n)$ and hence the corresponding best estimator. In turn, its variance can now be calculated. If this one vanishes, then we are sure that each outcome is ξ_n and from one realization we can determine the n -point correlation function. In this case, $\hat{\Delta}(\mathbf{e})$ is said to be ergodic; i.e., ensemble or quantum averages coincide with spatial averages. Unfortunately, one can show that this cannot be the case on the two-dimensional sphere [61]. Otherwise, we have found the weight function $E(\xi_n)$ which leads to the smallest nonvanishing variance. If the variance is small enough, each outcome will be concentrated around the mean value and with just one realization we have good chances to get a reasonable estimate of the correlation function. The typical error made in considering that one given outcome is equal to the mean value is characterized by the variance of the estimator; see Fig. 1.

All of the above analysis performed for the best (quantum) estimators can equally well be reproduced in the case where the anisotropies are due to an underlying stochastic

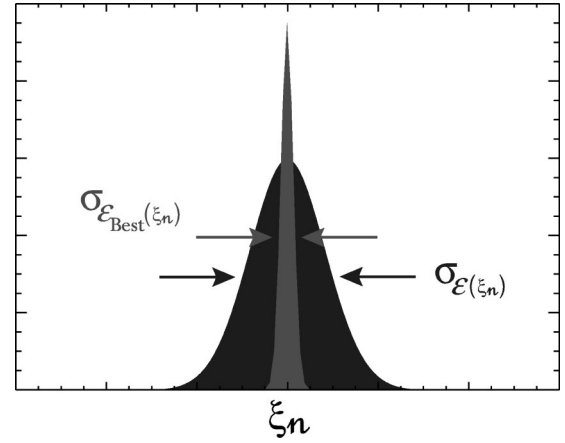


FIG. 1. Sketch comparing the variances of an arbitrary estimator of ξ_n with the best estimator of the same quantity. The widest distribution does not permit an accurate determination of ξ_n due to its large variance while the narrow distribution corresponds to the best (unbiased) estimator and possesses the smallest possible variance. As a consequence, one given realization (i.e., our CMB sky) will most probably be closest to the mean value in the second case rather than in the first one.

process although the former is generally physically best motivated. In that case, quantum averages $\langle \Psi | \dots | \Psi \rangle$ are just replaced with stochastic averages $\langle \dots \rangle$. In the following, we will drop out the caret symbol and consider that the different quantities are either operators or stochastic processes. In the same manner, we will denote an ensemble average by the symbol $\langle \dots \rangle$, having in mind that this means either quantum or classical averages.

Let us now describe the relevant quantities to estimate. It is convenient to expand the temperature fluctuations over the basis of spherical harmonics according to

$$\Delta(e) = \sum_{lm} a_l^m Y_l^m(\mathbf{e}). \quad (6)$$

This equation assumes a complete sky coverage. Implementation of the method for the incomplete (galaxy-cut) sky can be performed by using the basis introduced in [62]. Once a specific model is given, the statistical properties of the a_l^m 's are determined. Since $\Delta(\mathbf{e})$ is real, the a_l^m 's must satisfy $a_l^{m*} = (-1)^m a_l^{-m}$. Without restricting the generality of the underlying physics, the first three moments can be written as

$$\begin{aligned} \langle a_l^m \rangle &= 0, & \langle a_{l_1}^{m_1} a_{l_2}^{m_2*} \rangle &= C_{l_1} \delta_{l_1 l_2} \delta_{m_1 m_2}, \\ \langle a_{l_1}^{m_1} a_{l_2}^{m_2} a_{l_3}^{m_3} \rangle &= \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} C_{l_1 l_2 l_3}, \end{aligned} \quad (7)$$

where $\begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix}$ is a Wigner 3j symbol. The second equation ensures the isotropy of the CMB. The quantity $\langle a_{l_1}^{m_1} a_{l_2}^{m_2*} \rangle$ is the second moment of the a_l^m 's and C_l is usually called the angular spectrum. In the third equation, the quantity $\langle a_{l_1}^{m_1} a_{l_2}^{m_2} a_{l_3}^{m_3} \rangle$ is the third moment while $C_{l_1 l_2 l_3}$ is

called the angular bispectrum. For $l_1=l_2=l_3=l$, this quantity is generally written as $\mathcal{B}_l \equiv \mathcal{C}_{lll}$. The presence of the Wigner $3j$ symbol guarantees that the third moment vanishes unless $m_1+m_2+m_3=0$ and $|l_i-l_j| \leq l_k \leq l_i+l_j$. Moreover, invariance under spatial inversions of ξ_3 implies an additional “selection rule” [63,15] $l_1+l_2+l_3=\text{even}$ in order for the third moment not to vanish. Finally, from this last relation and using standard properties of the $3j$ symbols, it follows that the angular bispectrum is left unchanged under any arbitrary permutation of the indices l_i .

We will need the higher moments as well. Since departures from Gaussianity are expected to be small (especially on large angular scales), higher moments will be calculated

in the mildly non-Gaussian approximation. Within this approximation we can write $a_l^m = a_l^{m(0)} + \epsilon a_l^{m(1)} + \mathcal{O}(\epsilon^2)$ where $a_l^{m(0)}$ is a Gaussian random variable and the expansion parameter ϵ is small. In the following, each moment will be calculated to the first nonvanishing order in ϵ . For example, the fourth moment yields $\langle a_{l_1}^{m_1} a_{l_2}^{m_2} a_{l_3}^{m_3} a_{l_4}^{m_4} \rangle = \langle a_{l_1}^{m_1(0)} a_{l_2}^{m_2(0)} a_{l_3}^{m_3(0)} a_{l_4}^{m_4(0)} \rangle + \mathcal{O}(\epsilon)$. As a consequence, the connected fourth moment can be neglected because it is of higher order than the Gaussian part. The “(0)” label will be dropped out hereafter. Therefore, in the mildly non-Gaussian approximation we can write

$$\begin{aligned} \langle a_{l_1}^{m_1} a_{l_2}^{m_2} a_{l_3}^{m_3*} a_{l_4}^{m_4*} \rangle &\approx \langle a_{l_1}^{m_1} a_{l_2}^{m_2} \rangle \langle a_{l_3}^{m_3*} a_{l_4}^{m_4*} \rangle + \langle a_{l_1}^{m_1} a_{l_3}^{m_3*} \rangle \langle a_{l_2}^{m_2} a_{l_4}^{m_4*} \rangle + \langle a_{l_1}^{m_1} a_{l_4}^{m_4*} \rangle \langle a_{l_2}^{m_2} a_{l_3}^{m_3*} \rangle \\ &= (-1)^{m_2+m_4} C_{l_1} C_{l_3} \delta_{l_1 l_2} \delta_{m_1, -m_2} \delta_{l_3 l_4} \delta_{m_3, -m_4} + C_{l_1} C_{l_2} \delta_{l_1 l_3} \delta_{m_1 m_3} \delta_{l_2 l_4} \delta_{m_2 m_4} + C_{l_1} C_{l_2} \delta_{l_1 l_4} \delta_{m_1 m_4} \delta_{l_2 l_3} \delta_{m_2 m_3}. \end{aligned} \quad (8)$$

The fifth moment could be determined in a similar way but we will not need this quantity in the following. Finally, the sixth moment can be expressed as

$$\langle a_{l_1}^{m_1} a_{l_2}^{m_2} a_{l_3}^{m_3} a_{l_4}^{m_4*} a_{l_5}^{m_5*} a_{l_6}^{m_6*} \rangle \approx \langle a_{l_1}^{m_1} a_{l_2}^{m_2} \rangle \langle a_{l_3}^{m_3} a_{l_4}^{m_4*} \rangle \langle a_{l_5}^{m_5*} a_{l_6}^{m_6*} \rangle + 14 \text{ additional permutations.} \quad (9)$$

Although the explicit expression is not particularly illuminating, the last equation will be useful for the calculation of the variance when dealing with the three-point correlators below. In particular, one can write

$$\begin{aligned} \langle a_{l_1}^{m_1} a_{l_2}^{m_2} a_{l_3}^{m_3} a_{l_1}^{m_1*} a_{l_2}^{m_2*} a_{l_3}^{m_3*} \rangle &= C_{l_1} C_{l_2} C_{l_3} + 2C_{l_1}^3 \delta_{l_1 l_2 l_3} (\delta_{m_1 m_3} + \delta_{m_1 - m_3}) (\delta_{m_1 m_2} + \delta_{m_1 - m_2}) + C_{l_1} C_{l_2}^2 \delta_{l_2 l_3} (\delta_{m_2 m_3} + \delta_{m_2 - m_3}) \\ &\quad + C_{l_2} C_{l_3}^2 \delta_{l_3 l_1} (\delta_{m_1 m_3} + \delta_{m_1 - m_3}) + C_{l_3} C_{l_1}^2 \delta_{l_1 l_2} (\delta_{m_1 m_2} + \delta_{m_1 - m_2}), \end{aligned} \quad (10)$$

where the symbol $\delta_{l_1 l_2 l_3}$ vanishes unless $l_1=l_2=l_3$ in which case it is one. This equation coincides with Eq. (24) of [63] provided the undefined symbol $\delta_{m_1 m_2 m_3, 0}$ written in that work has the meaning $\delta_{m_1 m_2 m_3, 0} \equiv (\delta_{m_1 m_3} + \delta_{m_1 - m_3}) (\delta_{m_1 m_2} + \delta_{m_1 - m_2})$.

As we mentioned in the Introduction, in this article we are mainly interested in finding the best unbiased estimators for the two following quantities: the third moment $\langle a_{l_1}^{m_1} a_{l_2}^{m_2} a_{l_3}^{m_3} \rangle$ and the angular bispectrum $C_{l_1 l_2 l_3}$. These are related to the three-point correlation function ξ_3 . It is clear that, as mentioned above, the very same method could also be utilized for the computation of ξ_n with n arbitrary. Before addressing this question, however, we will first treat the analogous quantities related to the two-point correlation function, namely $\langle a_{l_1}^{m_1} a_{l_2}^{m_2*} \rangle$ and C_l , the main purpose being to illustrate concretely the tactics presented above in a case where everything can be calculated easily. This will be used as a guideline for the case of the three-point correlators.

III. TWO-POINT CORRELATORS

A. Best estimator for the angular spectrum C_l

The definition of the estimator $\mathcal{E}(C_l)$ is given by our general prescription [see Eq. (2)]

$$\mathcal{E}(C_l) \equiv \int \int d\Omega_1 d\Omega_2 E^l(\mathbf{e}_1, \mathbf{e}_2) \Delta(\mathbf{e}_1) \Delta(\mathbf{e}_2), \quad (11)$$

where $E^l(\mathbf{e}_1, \mathbf{e}_2)$ is the weight function. The angular spectrum C_l is a real quantity and its estimator $\mathcal{E}(C_l)$ must also be real. Therefore, the weight function can be taken as real. From the previous definition, it is clear that the antisymmetric part of $E^l(\mathbf{e}_1, \mathbf{e}_2)$ does not contribute to the estimator. Then, we can replace $E^l(\mathbf{e}_1, \mathbf{e}_2)$ in $\mathcal{E}(C_l)$ by its symmetrized expression $E_S^l(\mathbf{e}_1, \mathbf{e}_2) = (1/2)[E^l(\mathbf{e}_1, \mathbf{e}_2) + E^l(\mathbf{e}_2, \mathbf{e}_1)]$. At this stage, two methods can be applied. Either we work directly with the weight function or we expand it over the spherical harmonics basis and try to determine the coefficients of this expansion. Clearly, both paths are equivalent and can be followed for any estimator. Here we employ the second

method, leaving the first one for the determination of the second-moment estimator considered in the next subsection. Therefore, we write the weight function as

$$E_S^l(\mathbf{e}_1, \mathbf{e}_2) = \sum_{l_1 m_1} \sum_{l_2 m_2} d \begin{Bmatrix} l_1 & l_2 \\ m_1 & m_2 \end{Bmatrix} Y_{l_1}^{m_1}(\mathbf{e}_1) Y_{l_2}^{m_2}(\mathbf{e}_2). \quad (12)$$

The reality and symmetry properties of the weight function E_S^l imply that the complex coefficient of the expansion must satisfy, respectively,

$$\begin{aligned} d \begin{Bmatrix} l_1^* & l_2 \\ m_1 & m_2 \end{Bmatrix} &= (-1)^{m_1+m_2} d \begin{Bmatrix} l_1 & l_2 \\ -m_1 & -m_2 \end{Bmatrix}, \\ d \begin{Bmatrix} l_1 & l_2 \\ m_1 & m_2 \end{Bmatrix} &= d \begin{Bmatrix} l_2 & l_1 \\ m_2 & m_1 \end{Bmatrix}. \end{aligned} \quad (13)$$

Inserting the expression of the weight function (12) into the general definition of the estimator (11) and using standard properties of the spherical harmonics, one gets

$$\mathcal{E}(C_l) = \sum_{l_1 m_1} \sum_{l_2 m_2} d \begin{Bmatrix} l_1^* & l_2 \\ m_1 & m_2 \end{Bmatrix} a_{l_1}^{m_1} a_{l_2}^{m_2}. \quad (14)$$

Our first move is now to require that $\langle \mathcal{E}(C_l) \rangle = C_l$. Using the second equation of Eq. (7), we find that the coefficients d must satisfy the following constraints:

$$\sum_{m_1} (-1)^{m_1} d \begin{Bmatrix} l_1 & l_1 \\ -m_1 & m_1 \end{Bmatrix} = \delta_{l_1}^l. \quad (15)$$

All the estimators satisfying this condition are unbiased. However, it is clear that this does not completely determine the estimator but just a class of estimators. Our second move is to calculate the variance. Using Eq. (8), a straightforward computation gives

$$\sigma_{\mathcal{E}(C_l)}^2 = 2 \sum_{l_1 m_1} \sum_{l_2 m_2} d \begin{Bmatrix} l_1^* & l_2 \\ m_1 & m_2 \end{Bmatrix} d \begin{Bmatrix} l_1 & l_2 \\ m_1 & m_2 \end{Bmatrix} C_{l_1} C_{l_2}. \quad (16)$$

This quantity is obviously positive. From this expression, one sees that the imaginary part of the coefficients d only increases the variance. Since a vanishing $\mathcal{I}(d)$ satisfies the constraint equation, we can consider that the d 's are real.

Our third move is to minimize the variance taking into account the constraint. For this purpose we introduce a set of Lagrange multipliers $\lambda_{l_1}^l$ (i.e., one Lagrange multiplier per constraint since l must be seen as a fixed index) and require that

$$\delta \left[\sigma_{\mathcal{E}(C_l)}^2 + \sum_{l_1} \lambda_{l_1}^l \left(\sum_{m_1} (-1)^{m_1} d \begin{Bmatrix} l_1 & l_1 \\ -m_1 & m_1 \end{Bmatrix} - \delta_{l_1}^l \right) \right] = 0. \quad (17)$$

The definition of the variation δ must respect the symmetry properties of the coefficients d ; we take

$$\begin{aligned} \frac{\delta d \begin{Bmatrix} l_1 & l_2 \\ m_1 & m_2 \end{Bmatrix}}{\delta d \begin{Bmatrix} l_1' & l_2' \\ m_1' & m_2' \end{Bmatrix}} &= \frac{1}{2} \left\{ \frac{1}{2} \delta_{l_1 l_1'} \delta_{l_2 l_2'} [\delta_{m_1 m_1'} \delta_{m_2 m_2'} \right. \\ &\quad + (-1)^{m_1+m_2} \delta_{m_1 - m_1'} \delta_{m_2 - m_2'}] \\ &\quad + \frac{1}{2} \delta_{l_1 l_2'} \delta_{l_2 l_1'} [\delta_{m_1 m_2'} \delta_{m_2 m_1'} \\ &\quad \left. + (-1)^{m_1+m_2} \delta_{m_1 - m_2'} \delta_{m_2 - m_1'}] \right\}. \end{aligned} \quad (18)$$

Although it is not compulsory to use this equation, since a naive definition of the variation would lead to the same final result [61], it is nevertheless interesting to utilize it as a warm up for what will be done for the angular bispectrum. The variation leads to the following relation between the coefficients d and the Lagrange multipliers:

$$4d \begin{Bmatrix} l_1 & l_2 \\ m_1 & m_2 \end{Bmatrix} C_{l_1} C_{l_2} + (-1)^{m_1} \lambda_{l_1}^l \delta_{l_1 l_2} \delta_{m_1 - m_2} = 0. \quad (19)$$

We see that the choice of the λ^l 's is not free; it is fixed by the variation itself. Using the constraint in this equation, we find $\lambda_{l_1}^l = -4[C_{l_1}^2/(2l_1+1)]\delta_{l_1}^l$. Having determined what the Lagrange multipliers are, the problem is completely solved. It is now sufficient to reintroduce this value for $\lambda_{l_1}^l$ in Eq. (19), get the coefficients d and, from this, also the best unbiased estimator: the weight function is then given by $E_{S,\text{best}}^l(\mathbf{e}_1, \mathbf{e}_2) = (1/4\pi)P_l(\mathbf{e}_1 \cdot \mathbf{e}_2)$ and the estimator itself by (see also [61])

$$\mathcal{E}_{\text{best}}(C_l) = \frac{1}{2l+1} \sum_{m=-l}^l a_l^m a_l^{m*}. \quad (20)$$

The variance of this estimator is the well-known ‘‘cosmic variance’’

$$\sigma_{\mathcal{E}_{\text{best}}(C_l)}^2 = \frac{2C_l^2}{2l+1}. \quad (21)$$

One remark is in order here. The cosmic variance is usually obtained in the following way: the previous estimator appears naturally from Eqs. (7) and it is usually assumed that the a_l^m 's are Gaussian random variables. In this case, the estimator (20) has a χ^2 probability density function and from this the cosmic variance can be easily recovered. The proof presented above is by no means equivalent to this naive derivation. There are many unbiased estimators and, *a priori*, nothing guarantees that the simplest one is the best one: i.e., the naive derivation is *not* sufficient to prove that the estimator (20) is the best one. This can be proved only along the lines described above. Moreover, we do not need to assume that the a_l^m 's obey a Gaussian statistics. Only the mildly non-Gaussian assumption is necessary for the calculation of the four-point correlators.

B. Best estimator for the second moment C_α

In this section, we discuss the best estimator for $C_\alpha \equiv \langle a_{l_1}^{m_1} a_{l_2}^{m_2*} \rangle$. Greek letters will always be employed for collective-index notation, like

$$\alpha \equiv \begin{Bmatrix} l_1 & l_2 \\ m_1 & m_2 \end{Bmatrix}$$

or

$$\alpha' \equiv \begin{Bmatrix} l'_1 & l'_2 \\ m'_1 & m'_2 \end{Bmatrix}.$$

Unlike in the foregoing subsection, we find the weight function directly without expanding it over the spherical harmonics basis. This method is closer to the one used by [45]. We also show that the best estimator of C_α cannot be deduced from the best estimator of the angular spectrum C_l obtained in the previous subsection as one could naively think.

Let us start with a couple of definitions: first, the quantity

$$R^\alpha(\mathbf{e}_1, \mathbf{e}_2) \equiv Y_{l_1}^{m_1}(\mathbf{e}_1) Y_{l_2}^{m_2*}(\mathbf{e}_2), \quad (22)$$

which is complex and nonsymmetric with regard to the position of the complex conjugate symbol “*.” This last property comes from the definition of $\langle a_{l_1}^{m_1} a_{l_2}^{m_2*} \rangle$ itself and it would be the case for any even-point correlator. Then, regarding complex conjugation, there is a slight difference between the even- and odd-point correlators. The real part of $R^\alpha(\mathbf{e}_1, \mathbf{e}_2)$ will be denoted as

$$R_R^\alpha(\mathbf{e}_1, \mathbf{e}_2) \equiv \frac{1}{2} [Y_{l_1}^{m_1}(\mathbf{e}_1) Y_{l_2}^{m_2*}(\mathbf{e}_2) + Y_{l_1}^{m_1*}(\mathbf{e}_1) Y_{l_2}^{m_2}(\mathbf{e}_2)]. \quad (23)$$

$R_R^\alpha(\mathbf{e}_1, \mathbf{e}_2)$ is not symmetric under a permutation of the two directions. It satisfies $R_R^\alpha(\mathbf{e}_2, \mathbf{e}_1) = R_R^{\bar{\alpha}}(\mathbf{e}_1, \mathbf{e}_2)$, where the index $\bar{\alpha}$ is defined by

$$\bar{\alpha} \equiv \begin{Bmatrix} l_2 & l_1 \\ m_2 & m_1 \end{Bmatrix}$$

The symmetries in the indices and in the directions will play an important role in what follows. Using the previous definitions, we can introduce a quantity which is symmetric *both* under a permutation of the two directions *and* under a permutation of the columns of internal indices in α :

$$\begin{aligned} R_S^\alpha(\mathbf{e}_1, \mathbf{e}_2) \equiv & \frac{1}{4} [Y_{l_1}^{m_1}(\mathbf{e}_1) Y_{l_2}^{m_2*}(\mathbf{e}_2) + Y_{l_1}^{m_1}(\mathbf{e}_2) Y_{l_2}^{m_2*}(\mathbf{e}_1) \\ & + Y_{l_1}^{m_1*}(\mathbf{e}_1) Y_{l_2}^{m_2}(\mathbf{e}_2) + Y_{l_1}^{m_1*}(\mathbf{e}_2) Y_{l_2}^{m_2}(\mathbf{e}_1)]; \end{aligned} \quad (24)$$

i.e., we have $R_S^\alpha(\mathbf{e}_1, \mathbf{e}_2) = R_S^\alpha(\mathbf{e}_2, \mathbf{e}_1) = R_S^{\bar{\alpha}}(\mathbf{e}_1, \mathbf{e}_2) = R_S^{\bar{\alpha}}(\mathbf{e}_2, \mathbf{e}_1)$. Finally, we will also employ a symmetrized Kröcker symbol defined according to

$$\delta_S^{\alpha\alpha'} \equiv \frac{1}{2} (\delta_{l_1 l'_1} \delta_{l_2 l'_2} \delta_{m_1 m'_1} \delta_{m_2 m'_2} + \delta_{l_1 l'_2} \delta_{l_2 l'_1} \delta_{m_1 m'_2} \delta_{m_2 m'_1}), \quad (25)$$

which is left unchanged under a permutation of the indices α and α' and also under permutations of the columns of each collective index separately.

Let us now turn to the computation of the best unbiased estimator; the general definition of an estimator can be expressed as

$$\mathcal{E}(C_\alpha) \equiv \int \int d\Omega_1 d\Omega_2 E^\alpha(\mathbf{e}_1, \mathbf{e}_2) \Delta(\mathbf{e}_1) \Delta(\mathbf{e}_2). \quad (26)$$

The quantity C_α is unchanged if we permute the columns of indices in α , $C_\alpha = C_{\bar{\alpha}}$ with $\bar{\alpha}$ given above. Being an estimator for C_α it is then natural to assume that $\mathcal{E}(C_\alpha)$ possesses the same property, $\mathcal{E}(C_\alpha) = \mathcal{E}(C_{\bar{\alpha}})$. Looking at the definition of $E^\alpha(\mathbf{e}_1, \mathbf{e}_2)$, Eq. (26), we easily see that the weight function will also satisfy $E^\alpha(\mathbf{e}_1, \mathbf{e}_2) = E^\alpha(\mathbf{e}_2, \mathbf{e}_1)$. Moreover, we take $E^\alpha(\mathbf{e}_1, \mathbf{e}_2)$ symmetric under a permutation in the directions $\mathbf{e}_1$ and $\mathbf{e}_2$, i.e., $E_S^\alpha(\mathbf{e}_1, \mathbf{e}_2) = E_S^\alpha(\mathbf{e}_2, \mathbf{e}_1)$.

Now, we require that the estimator $\mathcal{E}(C_\alpha)$ be unbiased, which implies that the following relation must be satisfied:

$$\int \int d\Omega_1 d\Omega_2 E_S^\alpha(\mathbf{e}_1, \mathbf{e}_2) R^{\alpha'}(\mathbf{e}_1, \mathbf{e}_2) = \delta_S^{\alpha\alpha'}. \quad (27)$$

We have required the presence of $\delta_S^{\alpha\alpha'}$ on the right hand side of the previous equation to respect the symmetries in α of the weight function. In the previous equation the E_S^α and $R^{\alpha'}$ are *a priori* complex. However, one can show that it is possible to work only with a real weight function $E_S^{\alpha*}(\mathbf{e}_1, \mathbf{e}_2) = E_S^\alpha(\mathbf{e}_1, \mathbf{e}_2)$ and with the $R_R^\alpha(\mathbf{e}_1, \mathbf{e}_2)$ defined above: the imaginary contributions would just increase the variance. (One could have also chosen to work with a pure imaginary weight function and the imaginary part of R^α .) Therefore, the constraint can be written as

$$\int \int d\Omega_1 d\Omega_2 E_S^\alpha(\mathbf{e}_1, \mathbf{e}_2) R_R^{\alpha'}(\mathbf{e}_1, \mathbf{e}_2) = \delta_S^{\alpha\alpha'}, \quad (28)$$

where E_S^α has been taken real. Let us now calculate the variance of the estimator $\mathcal{E}(C_\alpha)$. Using Eqs. (8), we easily find that

$$\begin{aligned} \sigma_{\mathcal{E}(C_\alpha)}^2 = & 2 \int \int \int \int d\Omega_1 d\Omega_2 d\Omega_3 d\Omega_4 E_S^\alpha(\mathbf{e}_1, \mathbf{e}_2) \\ & \times E_S^\alpha(\mathbf{e}_3, \mathbf{e}_4) \xi(\mathbf{e}_1 \cdot \mathbf{e}_3) \xi(\mathbf{e}_2 \cdot \mathbf{e}_4). \end{aligned} \quad (29)$$

Our next step is to minimize this variance under the constraint given in Eq. (28). For this purpose, we introduce a set of Lagrange multipliers λ^α and require that

$$\begin{aligned} \delta \left[\sigma_{\mathcal{E}(C_\alpha)}^2 + \sum_{\alpha'} \lambda_{\alpha'}^\alpha \left(\int \int d\Omega_1 d\Omega_2 E_S^\alpha(\mathbf{e}_1, \mathbf{e}_2) \right. \right. \\ \left. \left. \times R_R^{\alpha'}(\mathbf{e}_1, \mathbf{e}_2) - \delta_S^{\alpha\alpha'} \right) \right] = 0. \end{aligned} \quad (30)$$

At this point the precise meaning of the variation symbol δ matters. Before performing the variation, let us recall that the symmetries of the weight function must be respected; hence we will have

$$\frac{\delta E_S^\alpha(\mathbf{e}_i, \mathbf{e}_j)}{\delta E_S^\beta(\mathbf{e}_k, \mathbf{e}_l)} = \frac{1}{2} \left[\frac{\delta(\mathbf{e}_i \cdot \mathbf{e}_k - 1)}{2\pi} \frac{\delta(\mathbf{e}_j \cdot \mathbf{e}_l - 1)}{2\pi} + \frac{\delta(\mathbf{e}_i \cdot \mathbf{e}_l - 1)}{2\pi} \frac{\delta(\mathbf{e}_j \cdot \mathbf{e}_k - 1)}{2\pi} \right] \delta_S^{\alpha\beta}. \quad (31)$$

The 2π 's in the denominators come from the fact that, while the direction $\mathbf{e}_i$ is expressed in terms of the corresponding spherical angles like $\mathbf{e}_i \equiv (\theta_i, \varphi_i)$ and then $d\Omega_i = d\cos\theta_i d\varphi_i$, there is an extra 2π factor in $\delta(\mathbf{e}_i \cdot \mathbf{e}_k - 1) = 2\pi\delta(\cos\theta_i - \cos\theta_k)\delta(\varphi_i - \varphi_k)$. As a result of the variation, we obtain the following equation:

$$\int \int d\Omega_1 d\Omega_2 E_S^\alpha(\mathbf{e}_1, \mathbf{e}_2) \xi(\mathbf{e}_1 \cdot \mathbf{e}_3) \xi(\mathbf{e}_2 \cdot \mathbf{e}_4) = \frac{1}{4} \sum_{\alpha'} \lambda_{\alpha'}^\alpha R_S^{\alpha'}(\mathbf{e}_3, \mathbf{e}_4). \quad (32)$$

The quantity $R_S^{\alpha'}(\mathbf{e}_3, \mathbf{e}_4)$ appears naturally as a result of Eq. (31). The previous equation should be compared with Eq. (19) of the previous subsection. The result of the variation is a relation between the weight function and the Lagrange multiplier. Our aim now is to get an explicit expression for the weight function. This can be done by using the inverse two-point correlation function ξ^{-1} which satisfies (see also the Appendix)

$$\int d\Omega_j \xi(\mathbf{e}_i \cdot \mathbf{e}_j) \xi^{-1}(\mathbf{e}_j \cdot \mathbf{e}_k) \equiv \delta(\mathbf{e}_i \cdot \mathbf{e}_k - 1). \quad (33)$$

In the case of the three-point correlator, this definition leads to subtleties which will be examined in detail in the next section. Multiplying Eq. (32) by $\xi^{-1}(\mathbf{e}_1 \cdot \mathbf{e}_3) \xi^{-1}(\mathbf{e}_2 \cdot \mathbf{e}_4)$, integrating over directions $\mathbf{e}_3$ and $\mathbf{e}_4$, and relabeling indices, we arrive at

$$E_S^\alpha(\mathbf{e}_1, \mathbf{e}_2) = \frac{1}{4} \frac{1}{(2\pi)^2} \sum_{\alpha'} \int \int d\Omega_3 d\Omega_4 \lambda_{\alpha'}^\alpha \xi^{-1}(\mathbf{e}_1 \cdot \mathbf{e}_3) \times \xi^{-1}(\mathbf{e}_2 \cdot \mathbf{e}_4) R_S^{\alpha'}(\mathbf{e}_3, \mathbf{e}_4). \quad (34)$$

This equation for the second moment is analogous to Eq. (21) of [45] obtained for the third moment. It is clear from the previous section that, at this stage, our final goal has not yet been reached. The weight function is still expressed in terms of the Lagrange multipliers. The correct way to proceed is to remove the latter using the constraint given by Eq. (28) as was done in the previous subsection. Then, we first multiply Eq. (34) by the quantity $R_S^{\alpha''}(\mathbf{e}_1, \mathbf{e}_2)$ and, after having integrated over solid angles Ω_1 and Ω_2 , this leads to the equation that the Lagrange multipliers must satisfy:

$$16C_{l_1} C_{l_2} \delta_S^{\alpha\alpha''} = \left[(-1)^{m_1''+m_2''} \lambda \begin{Bmatrix} l_1'' & l_2'' \\ -m_1'' & -m_2'' \end{Bmatrix} + (-1)^{m_1''+m_2''} \lambda \begin{Bmatrix} l_2'' & l_1'' \\ -m_2'' & -m_1'' \end{Bmatrix} + \lambda \begin{Bmatrix} l_1'' & l_2'' \\ m_1'' & m_2'' \end{Bmatrix} + \lambda \begin{Bmatrix} l_2'' & l_1'' \\ m_2'' & m_1'' \end{Bmatrix} \right]. \quad (35)$$

This equation is the analogous of the equation $\lambda_{l_1}^l = -4[C_{l_1}^2/(2l_1+1)]\delta_{l_1}^l$ of the previous subsection. Here, the difference is that a *combination* of Lagrange multipliers with different indices appears rather than the Lagrange multiplier itself. However, we can reconstruct exactly this combination on the right hand side of Eq. (34). Indeed, we just have to multiply each side of Eq. (35) by $R_S^{\alpha''}(\mathbf{e}_1, \mathbf{e}_2)$ and perform the sum over α'' . Using the symmetry properties of $R_S^{\alpha''}(\mathbf{e}_1, \mathbf{e}_2)$ we obtain

$$4C_{l_1} C_{l_2} R_S^\alpha(\mathbf{e}_1, \mathbf{e}_2) = \sum_{\alpha''} \lambda_{\alpha''}^\alpha R_S^{\alpha''}(\mathbf{e}_1, \mathbf{e}_2). \quad (36)$$

We now insert this equation on the right hand side of Eq. (34) to remove the Lagrange multipliers and find

$$E_{S,\text{best}}^\alpha(\mathbf{e}_1, \mathbf{e}_2) = \frac{1}{4\pi^2} C_{l_1} C_{l_2} \int \int d\Omega_3 d\Omega_4 \xi^{-1}(\mathbf{e}_1 \cdot \mathbf{e}_3) \times \xi^{-1}(\mathbf{e}_2 \cdot \mathbf{e}_4) R_S^\alpha(\mathbf{e}_3, \mathbf{e}_4). \quad (37)$$

To go further, we express ξ^{-1} in Eq. (37) explicitly in terms of spherical harmonics (see the Appendix) which yields $E_{S,\text{best}}^\alpha(\mathbf{e}_1, \mathbf{e}_2) = R_S^\alpha(\mathbf{e}_1, \mathbf{e}_2)$. Now, we just plug this result into Eq. (26) to get the final explicit expression of the best unbiased estimator $\mathcal{E}_{\text{best}}(C_\alpha)$:

$$\mathcal{E}_{\text{best}}(C_\alpha) = \frac{1}{2} (a_{l_1}^{m_1*} a_{l_2}^{m_2} + a_{l_1}^{m_1} a_{l_2}^{m_2*}). \quad (38)$$

Some remarks are in order here. First, despite appearances, one *cannot* deduce $\mathcal{E}_{\text{best}}(C_\alpha)$ from $\mathcal{E}_{\text{best}}(C_{l_1})$. It is clear that the following equation holds:

$$\langle \mathcal{E}(C_\alpha) \rangle = \langle \mathcal{E}(C_{l_1}) \rangle \delta_{l_1 l_2} \delta_{m_1 m_2}. \quad (39)$$

However, from this equation we are *not* allowed to conclude that $\mathcal{E}_{\text{best}}(C_\alpha) = \mathcal{E}_{\text{best}}(C_{l_1}) \delta_{l_1 l_2} \delta_{m_1 m_2}$. Therefore, knowing one of the best unbiased estimators does not allow us to infer the other best one. To be specific, if one assumed the previous wrong relation, from Eq. (20) one would get

$$\mathcal{E}_{\text{best}}(C_\alpha) = \frac{1}{2l_1+1} \left(\sum_{m=-l_1}^{l_1} a_{l_1}^{m*} a_{l_1}^m \right) \delta_{l_1 l_2} \delta_{m_1 m_2} \quad (\text{false}), \quad (40)$$

and although this estimator is unbiased, it is *not* the best one. The second remark is that although the estimator given by

Eq. (38) could be naively regarded as a trivial one, this is not the case. Indeed, the estimator of Eq. (40) is as trivial as the actual best one. The moral is then that there exist simple choices which lead to the wrong answer. The only reliable method in the problem of minimizing the variance is therefore the one exposed above.

IV. THREE-POINT CORRELATORS

A. Best estimator for the angular bispectrum $C_{l_1 l_2 l_3}$

In this section, our aim is to determine the best unbiased estimator for the angular bispectrum $C_{l_1 l_2 l_3}$ defined in the third equation of Eqs. (7). According to our general prescription, the most general definition reads

$$\mathcal{E}(C_{l_1 l_2 l_3}) \equiv \int \int \int d\Omega_1 d\Omega_2 d\Omega_3 E_S^{l_1 l_2 l_3}(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) \times \Delta(\mathbf{e}_1) \Delta(\mathbf{e}_2) \Delta(\mathbf{e}_3). \quad (41)$$

As in the case of the two-point correlators, the weight function also possesses the properties of being real and symmetric under arbitrary permutations of directions $\mathbf{e}_i$. In addition, like $C_{l_1 l_2 l_3}$, the weight function satisfies $E_S^{l_1 l_2 l_3} = E_S^{l_2 l_1 l_3}$, as well as for any other arbitrary permutation of the indices l_i . We follow similar steps as for the angular spectrum and therefore we choose to expand the weight function on the basis of the spherical harmonics. Then, as in Eq. (12), we write

$$E_S^{l_1 l_2 l_3}(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) = \sum_{l'_1 m'_1} \sum_{l'_2 m'_2} \sum_{l'_3 m'_3} d \left\{ \begin{matrix} l_1 & l_2 & l_3 \\ l'_1 & l'_2 & l'_3 \\ m'_1 & m'_2 & m'_3 \end{matrix} \right\} Y_{l'_1}^{m'_1}(\mathbf{e}_1) Y_{l'_2}^{m'_2}(\mathbf{e}_2) Y_{l'_3}^{m'_3}(\mathbf{e}_3). \quad (42)$$

The properties of the weight function imply that the coefficients d must satisfy equations similar to those given in Eqs. (13):

$$d \left\{ \begin{matrix} l_1 & l_2 & l_3^* \\ l'_1 & l'_2 & l'_3 \\ m'_1 & m'_2 & m'_3 \end{matrix} \right\} = (-1)^{m'_1 + m'_2 + m'_3} d \left\{ \begin{matrix} l_1 & l_2 & l_3 \\ l'_1 & l'_2 & l'_3 \\ -m'_1 & -m'_2 & -m'_3 \end{matrix} \right\}, \quad d \left\{ \begin{matrix} l_1 & l_2 & l_3 \\ l'_1 & l'_2 & l'_3 \\ m'_1 & m'_2 & m'_3 \end{matrix} \right\} = d \left\{ \begin{matrix} l_1 & l_2 & l_3 \\ l'_2 & l'_1 & l'_3 \\ m'_2 & m'_1 & m'_3 \end{matrix} \right\}, \quad (43)$$

where the last relation is in fact valid for arbitrary permutations of any two columns of the collective subindex. Like the weight function, d is also left invariant under arbitrary permutations of indices l_i (not primed). The estimator can be expressed in terms of the coefficients d and the a_l^m 's only: inserting the expansion of the weight function in the above expression for the estimator and using standard properties of the spherical harmonics one obtains

$$\mathcal{E}(C_{l_1 l_2 l_3}) = \sum_{l'_1 m'_1} \sum_{l'_2 m'_2} \sum_{l'_3 m'_3} d \left\{ \begin{matrix} l_1 & l_2 & l_3^* \\ l'_1 & l'_2 & l'_3 \\ m'_1 & m'_2 & m'_3 \end{matrix} \right\} a_{l'_1}^{m'_1} a_{l'_2}^{m'_2} a_{l'_3}^{m'_3}. \quad (44)$$

In practice, CMB observational settings are devised such that both the monopole and the dipole are subtracted from the anisotropy maps. This means that the coefficients d in the last equation are only nonvanishing for indices $l'_i \geq 2$ in the collective subindex. Moreover, the coefficients d satisfy $l_1 + l_2 + l_3 = \text{even}$. We must now require that our general estimator given by Eq. (44) be unbiased, i.e., $\langle \mathcal{E}(C_{l_1 l_2 l_3}) \rangle = C_{l_1 l_2 l_3}$. This forces the coefficients d to satisfy the constraint

$$\sum_{m'_1 m'_2 m'_3} d \left\{ \begin{matrix} l_1 & l_2 & l_3^* \\ l'_1 & l'_2 & l'_3 \\ m'_1 & m'_2 & m'_3 \end{matrix} \right\} \begin{pmatrix} l'_1 & l'_2 & l'_3 \\ m'_1 & m'_2 & m'_3 \end{pmatrix} = \delta_S^{l_1 l_2 l_3}, \quad (45)$$

where we have defined a new symmetrized Kröneckers symbol, this time for the l multipole indices only, as follows:

$$\delta_S^{l_1 l_2 l_3} \equiv \frac{1}{6} (\delta_{l_1 l'_1} \delta_{l_2 l'_2} \delta_{l_3 l'_3} + 5 \text{ additional permutations}). \quad (46)$$

It is easy to check that the constraint equation satisfies the conditions imposed by Eqs. (43) on the coefficients d . In particular, let us justify the presence of the symbol $\delta_S^{l_1 l_2 l_3}$. Using the previous properties for d , relabeling the indices $m'_1 \leftrightarrow m'_2$ in Eq. (45), and finally noting that $l'_1 + l'_2 + l'_3 = l_1 + l_2 + l_3 = \text{even}$, which allows us to permute any two columns of the Wigner $3j$ symbol, one verifies that the left hand side of the constraint is invariant under $l'_1 \leftrightarrow l'_2$. The same applies for any pair of l multipole indices and this explains the presence of the symmetrized $\delta_S^{l_1 l_2 l_3}$ in Eq. (45). We see from this that all coefficients d that do not satisfy $l'_1 + l'_2 + l'_3 = \text{even}$ do not enter the constraint. We will show below that these terms only increase the variance and as a consequence one can take them equal to zero. In particular, note that this is the case for a coefficient d with $l'_1 = l'_2$ and $l'_3 = \text{odd}$. This property will turn out to be useful in what follows.

We are now in a position to calculate the variance of the estimator. Looking at Eq. (44) we see that this requires the computation of the sixth moment of the a_l^m 's; see Eq. (9). After having made use of the properties of the coefficients d and rearranging the resulting 15 terms into two groups, straightforward algebra yields

$$\langle [\mathcal{E}(C_{l_1 l_2 l_3})]^2 \rangle = \sum_{l'_1 m'_1} \sum_{l'_2 m'_2} \sum_{l'_3 m'_3} C_{l'_1} C_{l'_2} C_{l'_3} \left[6d \begin{Bmatrix} l_1 & l_2 & l_3^* \\ l'_1 & l'_2 & l'_3 \\ m'_1 & m'_2 & m'_3 \end{Bmatrix} d \begin{Bmatrix} l_1 & l_2 & l_3 \\ l'_1 & l'_2 & l'_3 \\ m'_1 & m'_2 & m'_3 \end{Bmatrix} + 9(-1)^{m'_1+m'_2} d \begin{Bmatrix} l_1 & l_2 & l_3^* \\ l'_1 & l'_1 & l'_3 \\ m'_1 & -m'_1 & m'_3 \end{Bmatrix} d \begin{Bmatrix} l_1 & l_2 & l_3 \\ l'_2 & l'_2 & l'_3 \\ m'_2 & -m'_2 & m'_3 \end{Bmatrix} \right]. \quad (47)$$

The square of the variance of $\mathcal{E}(C_{l_1 l_2 l_3})$ is given by

$$\sigma_{\mathcal{E}(C_{l_1 l_2 l_3})}^2 = \langle [\mathcal{E}(C_{l_1 l_2 l_3})]^2 \rangle - \langle \mathcal{E}(C_{l_1 l_2 l_3}) \rangle^2. \quad (48)$$

Following the discussion in Sec. II, the term $\langle [\mathcal{E}(C_{l_1 l_2 l_3})]^2 \rangle$ is of order ϵ^0 whereas the lowest nonvanishing order of $\langle \mathcal{E}(C_{l_1 l_2 l_3}) \rangle^2$ is ϵ^2 . Therefore, the latter one will not enter the minimization procedure and the variance squared will be written as $\sigma_{\mathcal{E}(C_{l_1 l_2 l_3})}^2 \approx \langle [\mathcal{E}(C_{l_1 l_2 l_3})]^2 \rangle$. However, let us notice that this does not occur in the case of the two-point correlator. Indeed, as we have seen, in that case both terms contributing to the square of the variance are of the same order in ϵ . Then, it follows that Eq. (47) corresponds to Eq. (16) in the last section.

Let us now examine the structure of the variance in more detail. The term $6d^*d \propto \Re^2(d) + \Im^2(d)$ in Eq. (47) is analogous to the one in Eq. (16). However, here there is another contribution, the $9d^*d$ term, which will play a crucial role in what follows. The imaginary part of this term of course vanishes, as the variance must be real. However, the real part of it contains a contribution like

$$9 \sum_{l'_3} C_{l'_3} \sum_{m'_3} \left[\sum_{l'_1} C_{l'_1} \sum_{m'_1} (-1)^{m'_1} \Im \left(d \begin{Bmatrix} l_1 & l_2 & l_3 \\ l'_1 & l'_1 & l'_3 \\ m'_1 & -m'_1 & m'_3 \end{Bmatrix} \right) \right] \left[\sum_{l'_2} C_{l'_2} \sum_{m'_2} (-1)^{m'_2} \Im \left(d \begin{Bmatrix} l_1 & l_2 & l_3 \\ l'_2 & l'_2 & l'_3 \\ m'_2 & -m'_2 & m'_3 \end{Bmatrix} \right) \right] = 9 \sum_{l'_3} C_{l'_3} \sum_{m'_3} [\dots]^2, \quad (49)$$

where the quantity $[\dots]^2$ depends on the indices l_1, l_2, l_3 and l'_3, m'_3 and is strictly positive or zero. Of course, there is a similar term coming from the real part of d . Thus, we see that the various contributions of the imaginary part of the coefficients d to the two terms, $6d^*d$ and $9d^*d$, only increase the variance. Since we know that a vanishing imaginary part does satisfy the constraint Eq. (45), it can be disregarded in the sequel. Therefore, Eq. (47) can then be written solely in terms of *real* coefficients d as follows:

$$\sigma_{\mathcal{E}(C_{l_1 l_2 l_3})}^2 = \sum_{l'_1 m'_1} \sum_{l'_2 m'_2} \sum_{l'_3 m'_3} C_{l'_1} C_{l'_2} C_{l'_3} \left[6 \left(d \begin{Bmatrix} l_1 & l_2 & l_3 \\ l'_1 & l'_2 & l'_3 \\ m'_1 & m'_2 & m'_3 \end{Bmatrix} \right)^2 + 9(-1)^{m'_1+m'_2} d \begin{Bmatrix} l_1 & l_2 & l_3 \\ l'_1 & l'_1 & l'_3 \\ m'_1 & -m'_1 & m'_3 \end{Bmatrix} d \begin{Bmatrix} l_1 & l_2 & l_3 \\ l'_2 & l'_2 & l'_3 \\ m'_2 & -m'_2 & m'_3 \end{Bmatrix} \right]. \quad (50)$$

Our next move now is to minimize this variance with respect to the coefficients d , taking into account the constraint of Eq. (45):

$$\delta \left\{ \sigma_{\mathcal{E}(C_{l_1 l_2 l_3})}^2 + \sum_{l'_1 l'_2 l'_3} \lambda_{l'_1 l'_2 l'_3} \left[\sum_{m'_1 m'_2 m'_3} d \begin{Bmatrix} l_1 & l_2 & l_3 \\ l'_1 & l'_2 & l'_3 \\ m'_1 & m'_2 & m'_3 \end{Bmatrix} \begin{Bmatrix} l'_1 & l'_2 & l'_3 \\ m'_1 & m'_2 & m'_3 \end{Bmatrix} - \delta_S^{l'_1 l'_2} \right] \right\} = 0. \quad (51)$$

This equation is analogous to Eq. (17). As before, one needs to give a concrete meaning to the symbol δ of the variation. Its definition must respect the symmetries of the coefficients d and hence we take

$$\frac{\delta d \begin{Bmatrix} l_1 & l_2 & l_3 \\ l'_1 & l'_2 & l'_3 \\ m'_1 & m'_2 & m'_3 \end{Bmatrix}}{d \begin{Bmatrix} l_1 & l_2 & l_3 \\ l'_1 & l'_2 & l'_3 \\ m'_1 & m'_2 & m'_3 \end{Bmatrix}} = \frac{1}{6} \left\{ \frac{1}{2} \delta_{l'_1 l'_1} \delta_{l'_2 l'_2} \delta_{l'_3 l'_3} [\delta_{m'_1 m'_1} \delta_{m'_2 m'_2} \delta_{m'_3 m'_3} + (-1)^{m'_1+m'_2+m'_3} \delta_{m'_1 - m'_1} \delta_{m'_2 - m'_2} \delta_{m'_3 - m'_3}] + 5 \text{ terms} \right\}. \quad (52)$$

Then, Eq. (51) leads to

$$12 C_{l'_1} C_{l'_2} C_{l'_3} d \begin{Bmatrix} l_1 & l_2 & l_3 \\ l'_1 & l'_2 & l'_3 \\ m'_1 & m'_2 & m'_3 \end{Bmatrix} + \lambda_{l'_1 l'_2 l'_3} \begin{Bmatrix} l'_1 & l'_2 & l'_3 \\ m'_1 & m'_2 & m'_3 \end{Bmatrix} + 6(-1)^{m'_2} C_{l'_2} C_{l'_3} \delta_{l'_1 l'_1} \delta_{m'_1 - m'_2} \sum_{lm} C_l (-1)^m d \begin{Bmatrix} l_1 & l_2 & l_3 \\ l & l & l'_3 \\ m & -m & m'_3 \end{Bmatrix} \\ + 6(-1)^{m'_3} C_{l'_3} C_{l'_1} \delta_{l'_2 l'_2} \delta_{m'_2 - m'_3} \sum_{lm} C_l (-1)^m d \begin{Bmatrix} l_1 & l_2 & l_3 \\ l & l & l'_1 \\ m & -m & m'_1 \end{Bmatrix} + 6(-1)^{m'_1} C_{l'_1} C_{l'_2} \delta_{l'_3 l'_3} \delta_{m'_3 - m'_1} \sum_{lm} C_l (-1)^m d \begin{Bmatrix} l_1 & l_2 & l_3 \\ l & l & l'_2 \\ m & -m & m'_2 \end{Bmatrix} = 0. \quad (53)$$

This formula and Eq. (45) form a set of equations which completely determines the best unbiased estimator. We see the complicated structure of these equations. The last three terms come from the $9dd$ term in the variance and are not present in the case of the angular spectrum.

From this last equation and using the constraint, Eq. (45), we can get the general expression for the Lagrange multipliers. Thus, we multiply Eq. (53) by the appropriate $3j$ symbol and we sum over the three indices m'_i . The first term is exactly the constraint and produces a $\delta_S^{l'_i l'_j}$. Using the fact that a triple sum over the m_i 's of the squared of a $3j$ symbol gives unity, the second term yields the Lagrange multipliers themselves. Finally, the last three terms vanish: indeed, after straightforward manipulations one generates a term like

$$d_{\left\{ \begin{smallmatrix} l_1 & l_2 & l_3 \\ m & -m & 0 \end{smallmatrix} \right\}}^{l'_1 l'_2 l'_3} \sum_{m'_1} (-1)^{m'_1} \begin{pmatrix} l'_1 & l'_1 & l'_3 \\ m'_1 & -m'_1 & 0 \end{pmatrix} \quad (54)$$

for the third term of Eq. (53) and analogously for the last two ones. As we mentioned previously, the coefficient d in Eq. (54) vanishes unless l'_3 is even. In this case, recalling the identity [18],

$$\sum_{m'_1} (-1)^{m'_1} \begin{pmatrix} l'_1 & l'_1 & l'_3 \\ m'_1 & -m'_1 & 0 \end{pmatrix} = (-1)^{l'_1} \sqrt{2l'_1 + 1} \delta_{l'_3 0}, \quad l'_3 = \text{even}, \quad (55)$$

we see that there will only be a nonvanishing term if $l'_3 = 0$. But, the corresponding d is zero because $l'_3 < 2$ and therefore terms of this kind do not contribute to the Lagrange multipliers. Then, these are given by

$$\lambda_{l'_1 l'_2 l'_3}^{l_1 l_2 l_3} = -12 C_{l'_1} C_{l'_2} C_{l'_3} \delta_S^{l'_i l'_j}. \quad (56)$$

Plugging this into Eq. (53), one has

$$\begin{aligned} & 12 C_{l'_1} C_{l'_2} C_{l'_3} \left[d_{\left\{ \begin{smallmatrix} l_1 l_2 l_3 \\ m_1 m_2 m_3 \end{smallmatrix} \right\}}^{l'_1 l'_2 l'_3} - \delta_S^{l'_i l'_j} \begin{pmatrix} l'_1 & l'_2 & l'_3 \\ m'_1 & m'_2 & m'_3 \end{pmatrix} \right] + 6(-1)^{m'_2} C_{l'_2} C_{l'_3} \delta_{l'_1 l'_2} \delta_{m'_1 - m'_2} \sum_{lm} C_l (-1)^m d_{\left\{ \begin{smallmatrix} l_1 l_2 l_3 \\ m - m m'_3 \end{smallmatrix} \right\}}^{l'_1 l'_2 l'_3} \\ & + 6(-1)^{m'_3} C_{l'_3} C_{l'_1} \delta_{l'_2 l'_3} \delta_{m'_2 - m'_3} \sum_{lm} C_l (-1)^m d_{\left\{ \begin{smallmatrix} l_1 l_2 l_3 \\ m - m m'_1 \end{smallmatrix} \right\}}^{l'_1 l'_2 l'_3} + 6(-1)^{m'_1} C_{l'_1} C_{l'_2} \delta_{l'_3 l'_1} \delta_{m'_3 - m'_1} \sum_{lm} C_l (-1)^m d_{\left\{ \begin{smallmatrix} l_1 l_2 l_3 \\ m - m m'_2 \end{smallmatrix} \right\}}^{l'_1 l'_2 l'_3} = 0. \end{aligned} \quad (57)$$

This is the final equation to be solved in order to determine the best unbiased estimator. A solution is

$$d_{\left\{ \begin{smallmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{smallmatrix} \right\}}^{l'_1 l'_2 l'_3} = \begin{pmatrix} l'_1 & l'_2 & l'_3 \\ m'_1 & m'_2 & m'_3 \end{pmatrix} \delta_S^{l'_i l'_j}. \quad (58)$$

This leads to

$$\mathcal{E}_{\text{best}}(C_{l_1 l_2 l_3}) = \sum_{m'_1 m'_2 m'_3} \begin{pmatrix} l_1 & l_2 & l_3 \\ m'_1 & m'_2 & m'_3 \end{pmatrix} a_{l_1}^{m'_1} a_{l_2}^{m'_2} a_{l_3}^{m'_3}. \quad (59)$$

This is the main result of this subsection.

Given that we now know the best unbiased estimator for $C_{l_1 l_2 l_3}$, one can compute its variance, the smallest one among all possible estimator variances. In [15] we have already calculated it, and it reads

$$\sigma_{\mathcal{E}_{\text{best}}(C_{l_1 l_2 l_3})}^2 = C_{l_1} C_{l_2} C_{l_3} (1 + \delta_{l_1 l_2} + \delta_{l_2 l_3} + \delta_{l_3 l_1} + 2 \delta_{l_1 l_2} \delta_{l_2 l_3}). \quad (60)$$

In the same reference a plot of this variance for low order multipoles can also be found. This is what one could dub

(the square of) the “bispectrum cosmic variance” in perfect analogy with $\sigma_{\mathcal{E}_{\text{best}}(C_l)}^2 = 2C_l^2/(2l+1)$, which is (the square of) the variance of the best unbiased estimator for the angular spectrum, commonly known as the “cosmic variance.”

Let us conclude this subsection by comparing our results with those that recently appeared in the literature. An estimator restricted to the diagonal case $l_1 = l_2 = l_3$ has been proposed in [4] (see also [64] for an extension of their analysis) for $\mathcal{B}_l \equiv C_{lll}$ and reads

$$\begin{aligned} \mathcal{E}(\mathcal{B}_l) &= \frac{1}{2l+1} \begin{pmatrix} l & l & l \\ 0 & 0 & 0 \end{pmatrix}^{-3/2} \\ &\times \sum_{m_1 m_2 m_3} \begin{pmatrix} l & l & l \\ m_1 & m_2 & m_3 \end{pmatrix} a_l^{m_1} a_l^{m_2} a_l^{m_3}. \end{aligned} \quad (61)$$

In that work, the aim of the authors was not to seek the best estimator, but to use Eq. (61) to analyze the non-Gaussian features of the 4-yr COBE-DMR data. It is easy to see that their estimator does *not* satisfy the constraint (45); i.e., the estimator is biased. This is due to the presence of the overall prefactor in front of the triple sum in Eq. (61). However, as

we have proved above, getting rid of it produces the best unbiased estimator $\mathcal{E}_{\text{best}}(\mathcal{C}_{l_1 l_2 l_3})$, Eq. (59).

B. Best estimator for the third moment $\mathcal{B}^\alpha$

We now seek an estimator for $\mathcal{B}^\alpha \equiv \langle a_{l_1}^{m_1} a_{l_2}^{m_2} a_{l_3}^{m_3} \rangle$ where, as in the last section, it is convenient to define a collective index

$$\alpha \equiv \begin{Bmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{Bmatrix}.$$

This question has already been addressed in [45]. As we did with the second moment, our starting expression for an unbiased cubic estimator $\mathcal{E}(\mathcal{B}^\alpha)$ of $\mathcal{B}^\alpha$ will be in the form

$$\mathcal{E}(\mathcal{B}^\alpha) = \int \int \int d\Omega_1 d\Omega_2 d\Omega_3 E^\alpha(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) \times \Delta(\mathbf{e}_1) \Delta(\mathbf{e}_2) \Delta(\mathbf{e}_3). \quad (62)$$

The goal is to find the weight function $E^\alpha(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ that minimizes the variance of the estimator. As above, the quantity $\mathcal{B}^\alpha$ is unchanged if we permute arbitrary columns of indices in α : $\mathcal{B}^\alpha = \mathcal{B}^{\bar{\alpha}}$, where, for instance,

$$\bar{\alpha} \equiv \begin{Bmatrix} l_2 & l_1 & l_3 \\ m_2 & m_1 & m_3 \end{Bmatrix}.$$

$\mathcal{E}(\mathcal{B}^\alpha)$ has the same properties as $\mathcal{B}^\alpha$ and then it follows that $\mathcal{E}(\mathcal{B}^\alpha) = \mathcal{E}(\mathcal{B}^{\bar{\alpha}})$ for any column-permuted $\bar{\alpha}$. This implies that the $E^\alpha(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ satisfies $E^\alpha(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) = E^{\bar{\alpha}}(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$.

Now, is $E^\alpha(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ also symmetric under permutations in the directions $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$? From its definition we cannot know, for these directions are integrated over in the above defining equation. Unlike the case for the second moment discussed before, here E^α cannot be decomposed into symmetric and antisymmetric parts. However, we can always write $E^\alpha = E_S^\alpha + \text{something}$, and show that this last contribution to Eq. (62) vanishes. Therefore, there is no loss of generality in working with $E_S^\alpha(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) \equiv \frac{1}{6} [E^\alpha(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) + 5\text{terms}]$ which is symmetric under arbitrary permutations of directions $\mathbf{e}_i$.

Demanding the estimator $\mathcal{E}(\mathcal{B}^\alpha)$ to be unbiased, $\langle \mathcal{E}(\mathcal{B}^\alpha) \rangle = \mathcal{B}^\alpha$, yields the first constraint equation that the weight function $E_S^\alpha(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ must satisfy:

$$\int \int \int d\Omega_1 d\Omega_2 d\Omega_3 E_S^\alpha(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) R_R^{\alpha'}(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) = \delta_S^{\alpha\alpha'}, \quad (63)$$

where $R^{\alpha'}(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) \equiv Y_{l_1}^{m_1'}(\mathbf{e}_1) Y_{l_2}^{m_2'}(\mathbf{e}_2) Y_{l_3}^{m_3'}(\mathbf{e}_3)$ and

$R_R^{\alpha'}(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ is its real part. The form of R^α comes from the expression of ξ_3 , viz., $\langle \Delta(\mathbf{e}_1) \Delta(\mathbf{e}_2) \Delta(\mathbf{e}_3) \rangle = \sum_\alpha \mathcal{B}^\alpha R^\alpha(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$. The symmetrized Kröneckers symbol can be written as $\delta_S^{\alpha\alpha'} \equiv \frac{1}{6} (\delta_{l_1 l_1'} \delta_{m_1 m_1'} \delta_{l_2 l_2'} \delta_{m_2 m_2'} \delta_{l_3 l_3'} \delta_{m_3 m_3'} + 5\text{terms})$ as required to comply with the symmetry under permutations in the columns of α in $E_S^\alpha(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$. In the above equation, $R_R^{\alpha'}(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ is clearly nonsymmetric under a permutation of directions $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$. However, as above, we can define a symmetrized combination $R_S^\alpha(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) \equiv \frac{1}{6} [R_R^\alpha(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) + 5\text{terms}]$ (12 terms). Symmetrizing either in directions or in the columns of α in $R_R^\alpha(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ yields exactly the same $R_S^\alpha(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$. In the last equation and in what follows the weight function E_S^α is real for reasons similar to the ones exposed around Eq. (28) in the last section.

Proceeding as in Sec. III and using Eq. (9), one gets

$$\begin{aligned} \sigma_{\mathcal{E}(\mathcal{B}^\alpha)}^2 &= \int d\Omega_1 \cdots \int d\Omega_6 E_S^\alpha(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) E_S^\alpha(\mathbf{e}_4, \mathbf{e}_5, \mathbf{e}_6) \\ &\quad \times \xi(\mathbf{e}_1 \cdot \mathbf{e}_4) [6 \xi(\mathbf{e}_2 \cdot \mathbf{e}_5) \xi(\mathbf{e}_3 \cdot \mathbf{e}_6) \\ &\quad + 9 \xi(\mathbf{e}_2 \cdot \mathbf{e}_3) \xi(\mathbf{e}_5 \cdot \mathbf{e}_6)], \end{aligned} \quad (64)$$

where, utilizing the symmetry of the coefficients $E_S^\alpha(\mathbf{e}_i, \mathbf{e}_j, \mathbf{e}_k)$ under $\mathbf{e}$ -direction permutations, only two types of ξ products remain: first type, six terms where all the three ξ 's mix directions of the first and second E_S^α 's and, second type, nine terms where only one ξ [in the above equation, $\xi(\mathbf{e}_1 \cdot \mathbf{e}_4)$] does it.

To minimize the remaining variance under the constraint (63) we introduce a set of Lagrange multipliers λ^α and write

$$\begin{aligned} \delta \left[\sigma_{\mathcal{E}(\mathcal{B}^\alpha)}^2 - \sum_{\alpha'} \lambda_{\alpha'}^\alpha \left(\int \int \int d\Omega_1 d\Omega_2 d\Omega_3 E_S^\alpha(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) \right. \right. \\ \left. \left. \times R_R^{\alpha'}(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) - \delta_S^{\alpha\alpha'} \right) \right] = 0. \end{aligned} \quad (65)$$

As already noted for the second moment, the symmetries of the weight function must be respected in the variation; hence we have

$$\frac{\delta E_S^\alpha(\mathbf{e}_i, \mathbf{e}_j, \mathbf{e}_k)}{\delta E_S^\beta(\mathbf{e}_i, \mathbf{e}_j, \mathbf{e}_k)} = \frac{1}{6} \left[\frac{\delta(\mathbf{e}_i \cdot \mathbf{e}_i - 1)}{2\pi} \frac{\delta(\mathbf{e}_j \cdot \mathbf{e}_j - 1)}{2\pi} \frac{\delta(\mathbf{e}_k \cdot \mathbf{e}_k - 1)}{2\pi} + 5 \text{ terms} \right] \delta_S^{\alpha\beta}. \quad (66)$$

Now, we vary Eq. (65) and, after some algebra, we come up with

$$\begin{aligned}
& 2 \int \int \int d\Omega_4 d\Omega_5 d\Omega_6 \frac{1}{6} \{ \xi(\mathbf{e}_1 \cdot \mathbf{e}_4) [6\xi(\mathbf{e}_j \cdot \mathbf{e}_5) \xi(\mathbf{e}_k \cdot \mathbf{e}_6) + 9\xi(\mathbf{e}_j \cdot \mathbf{e}_k) \xi(\mathbf{e}_5 \cdot \mathbf{e}_6)] + 5 \text{ terms} \} E_S^\alpha(\mathbf{e}_4, \mathbf{e}_5, \mathbf{e}_6) \\
& = \sum_{\alpha'} \lambda_{\alpha'}^\alpha R_S^{\alpha'}(\mathbf{e}_i, \mathbf{e}_j, \mathbf{e}_k),
\end{aligned} \tag{67}$$

an expression which is symmetric in the arbitrary directions e_i, e_j, e_k as it should be (note the presence of $R_S^{\alpha'}$).

We aim at getting an explicit expression for $E_S^\alpha(\mathbf{e}_4, \mathbf{e}_5, \mathbf{e}_6)$. A glance at the previous equation shows that we need to multiply both sides of it by the inverse of the correlation function ξ^{-1} defined in the Appendix. Concretely, we multiply Eq. (67) by $\xi^{-1}(\mathbf{e}_i \cdot \mathbf{e}_{i'}) \xi^{-1}(\mathbf{e}_j \cdot \mathbf{e}_{j'}) \xi^{-1}(\mathbf{e}_k \cdot \mathbf{e}_{k'})$ and integrate over directions e_i, e_j, e_k ; we get

$$\begin{aligned}
& 6E_S^\alpha(\mathbf{e}_{i'}, \mathbf{e}_{j'}, \mathbf{e}_{k'}) + 9 \frac{(2\pi)^2}{3} \int \int d\Omega_5 d\Omega_6 \xi(\mathbf{e}_5 \cdot \mathbf{e}_6) [\xi^{-1}(\mathbf{e}_{i'} \cdot \mathbf{e}_{j'}) E_S^\alpha(\mathbf{e}_{k'}, \mathbf{e}_5, \mathbf{e}_6) \\
& + \xi^{-1}(\mathbf{e}_{j'}, \mathbf{e}_{k'}) E_S^\alpha(\mathbf{e}_{i'}, \mathbf{e}_5, \mathbf{e}_6) + \xi^{-1}(\mathbf{e}_{k'}, \mathbf{e}_{i'}) E_S^\alpha(\mathbf{e}_{j'}, \mathbf{e}_5, \mathbf{e}_6)] \\
& = \frac{1}{2} \int \int \int d\Omega_i d\Omega_j d\Omega_k \xi^{-1}(\mathbf{e}_i \cdot \mathbf{e}_{i'}) \xi^{-1}(\mathbf{e}_j \cdot \mathbf{e}_{j'}) \xi^{-1}(\mathbf{e}_k \cdot \mathbf{e}_{k'}) \sum_{\alpha'} \lambda_{\alpha'}^\alpha R_S^{\alpha'}(\mathbf{e}_i, \mathbf{e}_j, \mathbf{e}_k).
\end{aligned} \tag{68}$$

As before, the 2π factors on the left hand side come from operations like $2\pi f(\mathbf{e}_i) = \int d\Omega_k \delta(\mathbf{e}_i \cdot \mathbf{e}_k - 1) f(\mathbf{e}_k)$, for an arbitrary function $f(\mathbf{e}_k)$. However, we do not have an explicit expression for E_S^α yet. We see that expressions of the type $\xi(\mathbf{e}_5 \cdot \mathbf{e}_6) \xi^{-1}(\mathbf{e}_{i'} \cdot \mathbf{e}_{j'}) E_S^\alpha(\mathbf{e}_{k'}, \mathbf{e}_5, \mathbf{e}_6)$ are the ones that prevent us from isolating the weight function. To deal with this, it is convenient to construct the combination $\int \int d\Omega_5 d\Omega_6 E_S^\alpha(\mathbf{e}_{i'}, \mathbf{e}_5, \mathbf{e}_6) \xi(\mathbf{e}_5 \cdot \mathbf{e}_6)$. To reach this goal, we multiply both sides of Eq. (68) by $\xi(\mathbf{e}_{j'} \cdot \mathbf{e}_{k'})$ and integrate over directions $e_{j'}$ and $e_{k'}$. This operation produces a divergence on the left hand side of this equation in a form of a Dirac function “ $\delta(0)$.” In the continuous case, all methods lead to this unavoidable problem and although it has already appeared in the literature [45], it has never been treated so far. That this divergence is a mathematical artifact we can see from the fact that, in practice, we never deal with an ideal experiment: the problem is solved when we take into account the fact that each different experimental setting is limited by a finite angular resolution. This is usually quantified in terms of an l -dependent window function $\mathcal{W}_l$ (the circularly symmetric pattern of the observation beam in l space), although more involved scanning techniques are also employed [65,66]. Then, only a finite number of multipoles will effectively contribute to the correlation function and, as a result, the above-mentioned divergence is regularized; indeed, we have

$$\xi(\mathbf{e}_i \cdot \mathbf{e}_j) = \sum_l \frac{2l+1}{4\pi} C_l \mathcal{W}_l^2 P_l(\mathbf{e}_i \cdot \mathbf{e}_j), \tag{69}$$

which, upon using the expression for ξ^{-1} given in the Appendix, leads to the quantity

$$\begin{aligned}
\mathcal{A}_{ik} & \equiv \int d\Omega_j \xi(\mathbf{e}_i \cdot \mathbf{e}_j) \xi^{-1}(\mathbf{e}_j \cdot \mathbf{e}_k) \\
& = 2\pi \sum_{lm} \mathcal{W}_l^2 Y_l^m(\mathbf{e}_i) Y_l^{m*}(\mathbf{e}_k).
\end{aligned} \tag{70}$$

In particular, the previous divergence “ $\delta(0)$ ” now becomes

$$\mathcal{A} \equiv \mathcal{A}_{ii} = \sum_l \frac{2l+1}{2} \mathcal{W}_l^2. \tag{71}$$

This is a more realistic and finite object to work with in the case where two directions on the microwave sky coincide for a given experience. Notice that for an ideal experimental setting in which the window function $\mathcal{W}_l \rightarrow 1$ or, equivalently, the beam $\sigma \rightarrow 0$ in the case of a Gaussian profile, $\mathcal{A}$ blows up. Since the terms of the type $9\xi\xi$ in Eq. (64) are not present in the case of the two-point correlators, this problem did not appear there. From now on, strictly speaking, all expressions should incorporate the window function. However, in what follows and for computational convenience, we will keep $\mathcal{A}_{ik} \approx \delta(\mathbf{e}_i \cdot \mathbf{e}_k - 1)$ for $e_i \neq e_k$, a good approximation as we can see from Fig. 2.

Endowed now with the above regularization method, we present the term $\int \int d\Omega_5 d\Omega_6 E_S^\alpha(\mathbf{e}_{i'}, \mathbf{e}_5, \mathbf{e}_6) \xi(\mathbf{e}_5 \cdot \mathbf{e}_6)$ in the following form:

$$\begin{aligned}
& \int \int d\Omega_5 d\Omega_6 E_S^\alpha(\mathbf{e}_4, \mathbf{e}_5, \mathbf{e}_6) \xi(\mathbf{e}_5 \cdot \mathbf{e}_6) \\
& = \frac{\mathcal{D}}{24\pi^2} \sum_{\alpha'} \lambda_{\alpha'}^\alpha \int \int \int d\Omega_{i'} d\Omega_{j'} d\Omega_{k'} \\
& \quad \times \xi^{-1}(\mathbf{e}_4 \cdot \mathbf{e}_{i'}) \xi^{-1}(\mathbf{e}_{j'} \cdot \mathbf{e}_{k'}) R_S^{\alpha'}(\mathbf{e}_{i'}, \mathbf{e}_{j'}, \mathbf{e}_{k'}),
\end{aligned} \tag{72}$$

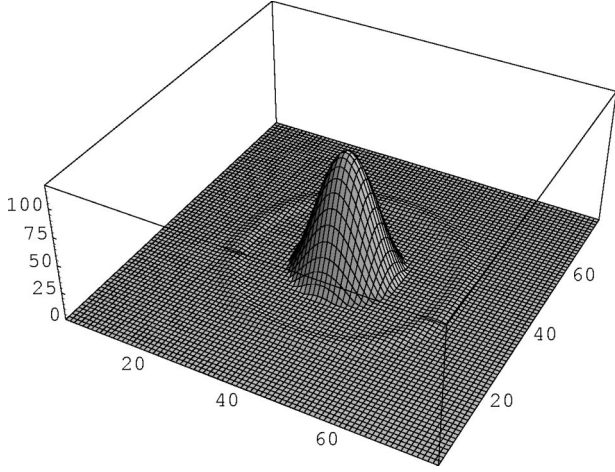


FIG. 2. Values of $\mathcal{A}_{ik}$ (as defined in the text) with fixed direction e_i at the center of the plot and direction e_k scanning a square of side 90° . We show the result for the COBE-DMR window function. The full width at half maximum of the plot is roughly 10° , of the same order as the resolution of the COBE-DMR detector. We can then expect the plotted “spike” to select just one experimental pixel on the sky map, making the relation $\mathcal{A}_{ik} \approx \delta(\mathbf{e}_i \cdot \mathbf{e}_k - 1)$ a good approximation. Note that higher resolution experiments will yield a more peaked curve, although one would expect the goodness of the approximation to remain roughly the same.

with $\mathcal{D} \equiv 24\pi^3 \{6 + 9[(2\pi)^2/3](2 \times 2\pi + 4\pi\mathcal{A})\}^{-1}$ where, as expected, the factor $\mathcal{A}$ appears explicitly. Now, we just replace the six terms with prefactor $9\xi\xi$ on the left hand side of Eq. (67) with Eq. (72), resulting in

$$\begin{aligned}
 & \int \int \int d\Omega_4 d\Omega_5 d\Omega_6 \xi(\mathbf{e}_i \cdot \mathbf{e}_4) \xi(\mathbf{e}_j \cdot \mathbf{e}_5) \xi(\mathbf{e}_k \cdot \mathbf{e}_6) \\
 & \times E_S^\alpha(\mathbf{e}_4, \mathbf{e}_5, \mathbf{e}_6) \\
 & = \frac{1}{12} \sum_{\alpha'} \lambda_{\alpha'}^\alpha \left\{ R_S^{\alpha'}(\mathbf{e}_i, \mathbf{e}_j, \mathbf{e}_k) \right. \\
 & \quad - \frac{\mathcal{D}}{2\pi} \int \int d\Omega_{j'} d\Omega_{k'} \xi^{-1}(\mathbf{e}_{j'} \cdot \mathbf{e}_{k'}) \\
 & \quad \times [R_S^{\alpha'}(\mathbf{e}_i, \mathbf{e}_{j'}, \mathbf{e}_{k'}) \xi(\mathbf{e}_j \cdot \mathbf{e}_k) + R_S^{\alpha'}(\mathbf{e}_j, \mathbf{e}_{j'}, \mathbf{e}_{k'}) \xi(\mathbf{e}_k \cdot \mathbf{e}_i) \\
 & \quad \left. + R_S^{\alpha'}(\mathbf{e}_k, \mathbf{e}_{j'}, \mathbf{e}_{k'}) \xi(\mathbf{e}_i \cdot \mathbf{e}_j) \right\}. \quad (73)
 \end{aligned}$$

This contains just one appearance of E^α ; to get the weight function explicitly, we only need to multiply the equation by $\xi^{-1}(\mathbf{e}_i \cdot \mathbf{e}_1) \xi^{-1}(\mathbf{e}_j \cdot \mathbf{e}_2) \xi^{-1}(\mathbf{e}_k \cdot \mathbf{e}_3)$ and integrate over directions e_i, e_j, e_k . Finally, we get the expression for E_S^α in terms of the Lagrange multipliers λ^α :

$$\begin{aligned}
 E_S^\alpha(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) & = \frac{1}{12 \times (2\pi)^3} \\
 & \times \sum_{\alpha'} \lambda_{\alpha'}^\alpha \left\{ \int \int \int d\Omega_i d\Omega_j d\Omega_k \xi^{-1}(\mathbf{e}_i \cdot \mathbf{e}_1) \right. \\
 & \times \xi^{-1}(\mathbf{e}_j \cdot \mathbf{e}_2) \xi^{-1}(\mathbf{e}_k \cdot \mathbf{e}_3) R_S^{\alpha'}(\mathbf{e}_i, \mathbf{e}_j, \mathbf{e}_k) \\
 & - \mathcal{D} \int \int \int d\Omega_i d\Omega_j d\Omega_k R_S^{\alpha'}(\mathbf{e}_i, \mathbf{e}_j, \mathbf{e}_k) \\
 & \times \xi^{-1}(\mathbf{e}_j \cdot \mathbf{e}_k) [\xi^{-1}(\mathbf{e}_i \cdot \mathbf{e}_1) \xi^{-1}(\mathbf{e}_2 \cdot \mathbf{e}_3) \\
 & + \xi^{-1}(\mathbf{e}_i \cdot \mathbf{e}_2) \xi^{-1}(\mathbf{e}_3 \cdot \mathbf{e}_1) \\
 & \left. + \xi^{-1}(\mathbf{e}_i \cdot \mathbf{e}_3) \xi^{-1}(\mathbf{e}_1 \cdot \mathbf{e}_2)] \right\}. \quad (74)
 \end{aligned}$$

Having reached this point, we have an explicit expression for E_S^α , but still dependent on the Lagrange multipliers. This equation is well defined as it contains the renormalized quantity $\mathcal{A}$. Within a particular experiment with a given resolution, the value $\mathcal{A}$ takes depends on what one means by two coincident directions. For example, for the COBE-DMR window-function specification (a Gaussian beam with dispersion $\sigma = 3.2^\circ$) this yields roughly $\mathcal{A} \approx 158.5$, including the quadrupole. It is not difficult to extend this to other scanning techniques. The previous equation is the analogue of our Eq. (34) corresponding to the second moment and also to Eq. (21) of [45]. In that article, a similar analysis is done but for the discretized CMB sky. We remark that no divergence appears in his case. Indeed, all relevant quantities are finite when evaluated for two directions pointing towards the same pixel. Our prefactor $\mathcal{D}$ corresponds to $3/(2+3N)$ in that paper, where N represents the number of pixels in the map. Clearly, $N \rightarrow \infty$ when the pixel size goes to zero, as well as $\mathcal{A} \rightarrow \infty$ when the window function $\mathcal{W}_l \rightarrow 1$. At this intermediate step our corresponding expressions need not coincide because both depend on the particular regularization scheme used (be it discretization or usage of a window function). Despite appearances, we will show below that the final expression for the best unbiased estimator does not depend on these schemes. This cannot be inferred from Eq. (74) because we still need to remove the Lagrange multiplier. Unlike what was done in [45], we now proceed further and express the weight function explicitly. Hence, we multiply both sides of Eq. (74) by $R_R^{\alpha''}(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$, where

$$\alpha'' \equiv \begin{Bmatrix} l_1'' & l_2'' & l_3'' \\ m_1'' & m_2'' & m_3'' \end{Bmatrix}$$

and then integrate over directions e_1, e_2 , and e_3 , and, upon using the constraint equation (63), we get

$$\begin{aligned}
12\delta_S^{\alpha\alpha''} = & \frac{1}{6} \frac{1}{C_{l_1''} C_{l_2''} C_{l_3''}} \left\{ \left[(-1)^{m_1''+m_2''+m_3''} \lambda \begin{Bmatrix} l_1'' & l_2'' & l_3'' \\ -m_1'' & -m_2'' & -m_3'' \end{Bmatrix} + \lambda \begin{Bmatrix} l_1'' & l_2'' & l_3'' \\ m_1'' & m_2'' & m_3'' \end{Bmatrix} \right] + 5 \text{ additional permutations in index } \alpha'' \right\} \\
& - \frac{\mathcal{D}}{3} \left\{ \left[\frac{\delta_{l_1'' l_2''} \delta_{m_1'' - m_2''} (-1)^{m_2''}}{C_{l_2''} C_{l_3''}} \sum_{lm} \frac{(-1)^m}{C_l} \left((-1)^{m_3''} \lambda \begin{Bmatrix} l & l & l_3'' \\ m & -m & -m_3'' \end{Bmatrix} + \lambda \begin{Bmatrix} l & l & l_3'' \\ -m & m & m_3'' \end{Bmatrix} + (-1)^{m_3''} \lambda \begin{Bmatrix} l_3'' & l & l \\ -m_3'' & m & -m \end{Bmatrix} \right. \right. \right. \\
& \left. \left. + \lambda \begin{Bmatrix} l_3'' & l & l \\ m_3'' & -m & m \end{Bmatrix} + (-1)^{m_3''} \lambda \begin{Bmatrix} l & l_3'' & l \\ m & -m_3'' & -m \end{Bmatrix} + \lambda \begin{Bmatrix} l & l_3'' & l \\ -m & m_3'' & m \end{Bmatrix} \right) \right] + \left[\begin{matrix} 1 \rightarrow 2 \\ 2 \rightarrow 3 \\ 3 \rightarrow 1 \end{matrix} \right] + \left[\begin{matrix} 1 \rightarrow 3 \\ 2 \rightarrow 1 \\ 3 \rightarrow 2 \end{matrix} \right] \right\}. \quad (75)
\end{aligned}$$

Equation (75) is the final algebraic equation that the multipliers must satisfy. For fixed, α , a natural way to proceed would be to get an explicit expression for $\lambda_{\alpha'}^{\alpha}$. Another way to solve the problem goes on a line analogous to the case of the two-point correlators: we just need to identify the complicated combination of Lagrange multipliers on the right hand side of Eq. (75) with the one on the right hand side of Eq. (73) [or, equivalently, Eq. (74)]. In order to do that, we now multiply both sides of Eq. (75) by $R_R^{\alpha''}(\mathbf{e}_i, \mathbf{e}_j, \mathbf{e}_k)$, perform the six sums over the indices in α'' , and we end up with

$$\begin{aligned}
R_S^{\alpha}(\mathbf{e}_i, \mathbf{e}_j, \mathbf{e}_k) C_{l_1} C_{l_2} C_{l_3} &= \sum_{\alpha''} R_R^{\alpha''}(\mathbf{e}_i, \mathbf{e}_j, \mathbf{e}_k) C_{l_1''} C_{l_2''} C_{l_3''} \delta_S^{\alpha\alpha''} \\
&= \frac{1}{12} \sum_{\alpha'} \lambda_{\alpha'}^{\alpha} R_S^{\alpha'}(\mathbf{e}_i, \mathbf{e}_j, \mathbf{e}_k) - \frac{\mathcal{D}}{36} \sum_{LM} \sum_{lm} \frac{(-1)^M}{C_L} [Y_L^m(\mathbf{e}_i) \xi(\mathbf{e}_j \cdot \mathbf{e}_k) + Y_L^m(\mathbf{e}_j) \xi(\mathbf{e}_k \cdot \mathbf{e}_i) + Y_L^m(\mathbf{e}_k) \xi(\mathbf{e}_i \cdot \mathbf{e}_j)] \\
&\quad \times \left[\lambda \begin{Bmatrix} l & L & L \\ m & M & -M \end{Bmatrix} + (-1)^m \lambda \begin{Bmatrix} l & L & L \\ -m & -M & M \end{Bmatrix} + \lambda \begin{Bmatrix} L & l & L \\ -M & m & M \end{Bmatrix} + (-1)^m \lambda \begin{Bmatrix} L & l & L \\ M & -m & -M \end{Bmatrix} \right. \\
&\quad \left. + \lambda \begin{Bmatrix} L & L & l \\ M & -M & m \end{Bmatrix} + (-1)^m \lambda \begin{Bmatrix} L & L & l \\ -M & M & -m \end{Bmatrix} \right]. \quad (76)
\end{aligned}$$

This is the equivalent of Eq. (35). The aim now is to show that the combination of Lagrange multipliers on the right hand side of the previous equation is precisely the one which appears on the right hand side of Eq. (73). In the latter, let us express both $R_S^{\alpha'}$ and ξ^{-1} in the second term on the right hand side in terms of spherical harmonics. After some algebra we get

$$\begin{aligned}
&\int \int \int d\Omega_4 d\Omega_5 d\Omega_6 \xi(\mathbf{e}_i \cdot \mathbf{e}_4) \xi(\mathbf{e}_j \cdot \mathbf{e}_5) \xi(\mathbf{e}_k \cdot \mathbf{e}_6) E_S^{\alpha}(\mathbf{e}_4, \mathbf{e}_5, \mathbf{e}_6) \\
&= \frac{1}{12} \sum_{\alpha'} \lambda_{\alpha'}^{\alpha} R_S^{\alpha'}(\mathbf{e}_i, \mathbf{e}_j, \mathbf{e}_k) - \frac{\mathcal{D}}{36} \sum_{LM} \sum_{lm} \frac{(-1)^M}{C_L} [Y_L^m(\mathbf{e}_i) \xi(\mathbf{e}_j \cdot \mathbf{e}_k) + Y_L^m(\mathbf{e}_j) \xi(\mathbf{e}_k \cdot \mathbf{e}_i) + Y_L^m(\mathbf{e}_k) \xi(\mathbf{e}_i \cdot \mathbf{e}_j)] \\
&\quad \times \left[\lambda \begin{Bmatrix} l & L & L \\ m & M & -M \end{Bmatrix} + (-1)^m \lambda \begin{Bmatrix} l & L & L \\ -m & -M & M \end{Bmatrix} + \lambda \begin{Bmatrix} L & l & L \\ -M & m & M \end{Bmatrix} + (-1)^m \lambda \begin{Bmatrix} L & l & L \\ M & -m & -M \end{Bmatrix} + \lambda \begin{Bmatrix} L & L & l \\ M & -M & m \end{Bmatrix} + (-1)^m \lambda \begin{Bmatrix} L & L & l \\ -M & M & -m \end{Bmatrix} \right]. \quad (77)
\end{aligned}$$

which, as advertised, yields the same combination of Lagrange multipliers of Eq. (76). Then, putting the last two equations together, multiplying by ξ^{-1} , and integrating 3 times, we finally get the weight function associated with the best unbiased estimator:

$$E_{S, \text{best}}^{\alpha}(\mathbf{e}_i, \mathbf{e}_j, \mathbf{e}_k) = R_S^{\alpha}(\mathbf{e}_i, \mathbf{e}_j, \mathbf{e}_k), \quad (78)$$

which implies that the best unbiased estimator itself is given by

$$\mathcal{E}_{\text{best}}(\mathcal{B}^{\alpha}) = \frac{1}{2} (a_{l_1}^{m_1} a_{l_2}^{m_2} a_{l_3}^{m_3} + a_{l_1}^{m_1*} a_{l_2}^{m_2*} a_{l_3}^{m_3*}). \quad (79)$$

This is the final answer and it is a new result. Let us make a few remarks. First, this does not depend on $\mathcal{A}$ which shows that Eq. (79) is independent of the regularization scheme used. Second, [63] used the following complex unbiased estimator $\mathcal{E}(\mathcal{B}^{\alpha}) = a_{l_1}^{m_1} a_{l_2}^{m_2} a_{l_3}^{m_3}$, although he did not claim it to be the best one. Third, as for the two-point correlators, one cannot use Eq. (79) in order to infer $\mathcal{E}_{\text{Best}}(C_{l_1 l_2 l_3})$. Indeed, promoting the equation

$$\langle a_{l_1}^{m_1} a_{l_2}^{m_2} a_{l_3}^{m_3} \rangle = \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} C_{l_1 l_2 l_3}, \quad (80)$$

valid for the mean values of the estimators, to the equation

$$\mathcal{E}_{\text{best}}(\mathcal{B}^\alpha) = \frac{1}{2} (a_{l_1}^{m_1} a_{l_2}^{m_2} a_{l_3}^{m_3} + a_{l_1}^{m_1*} a_{l_2}^{m_2*} a_{l_3}^{m_3*}) = \mathcal{E}_{\text{best}}(C_{l_1 l_2 l_3}) \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \quad (\text{false}), \quad (81)$$

valid for the estimators themselves, is an unjustified step. If, nevertheless, we used this false relation, we would get

$$\mathcal{E}_{\text{best}}(\mathcal{B}^\alpha) = \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \sum_{m'_1 m'_2 m'_3} \begin{pmatrix} l_1 & l_2 & l_3 \\ m'_1 & m'_2 & m'_3 \end{pmatrix} a_{l_1}^{m'_1} a_{l_2}^{m'_2} a_{l_3}^{m'_3} \quad (\text{false}), \quad (82)$$

which cannot be cast into $\frac{1}{2} (a_{l_1}^{m_1} a_{l_2}^{m_2} a_{l_3}^{m_3} + a_{l_1}^{m_1*} a_{l_2}^{m_2*} a_{l_3}^{m_3*})$. So, like for the two-point correlators, we see that one cannot infer the $\mathcal{E}_{\text{best}}(\mathcal{B}^\alpha)$ from $\mathcal{E}_{\text{best}}(C_{l_1 l_2 l_3})$ and vice versa.

Endowed now with the best unbiased estimator, one can compute its variance squared, the “third-moment cosmic variance,” which reads

$$\begin{aligned} \sigma_{\mathcal{E}_{\text{best}}(\mathcal{B}^\alpha)}^2 = & \frac{1}{2} \{ C_{l_1} C_{l_2} C_{l_3} [1 + \delta_{m_1 0} \delta_{m_2 0} \delta_{m_3 0}] + C_{l_1}^3 \delta_{l_1 l_2 l_3} [8 \delta_{m_1 0} \delta_{m_2 0} \delta_{m_3 0} + 2(\delta_{m_1 m_3} + \delta_{m_1 - m_3})(\delta_{m_1 m_2} + \delta_{m_1 - m_2})] \\ & + C_{l_1} C_{l_2}^2 \delta_{l_2 l_3} [2 \delta_{m_1 0} \delta_{m_2 - m_3} + \delta_{m_2 m_3} + \delta_{m_2 - m_3}] + C_{l_2} C_{l_3}^2 \delta_{l_3 l_1} [2 \delta_{m_2 0} \delta_{m_3 - m_1} + \delta_{m_3 m_1} + \delta_{m_3 - m_1}] \\ & + C_{l_3} C_{l_1}^2 \delta_{l_1 l_2} [2 \delta_{m_3 0} \delta_{m_1 - m_2} + \delta_{m_1 m_2} + \delta_{m_1 - m_2}] \}. \end{aligned} \quad (83)$$

For example, from this we can now compute the cosmic variance for the third-moment estimator $\mathcal{E}_{\text{best}}(\mathcal{B}^\beta)$ with

$$\beta \equiv \begin{pmatrix} l & l & l \\ m & m & -2m \end{pmatrix},$$

where $l = \text{even}$ and $m \neq 0$. This particular case is often treated in the literature: see, e.g., [63,45]. We find $\sigma_{\mathcal{E}_{\text{best}}(\mathcal{B}^\beta)}^2 = C_l^3$ whereas the variance of the estimator used in [63] yields $\sigma_{\mathcal{E}(\mathcal{B}^\beta)}^2 = 2C_l^3$. Another example comes from taking

$$\gamma \equiv \begin{pmatrix} l & l & l \\ m_1 & m_2 & m_3 \end{pmatrix},$$

where $|m_i| \neq |m_j|$ for any i, j . With this choice we get $\sigma_{\mathcal{E}_{\text{best}}(\mathcal{B}^\gamma)}^2 = C_l^3/2$ whereas [63] obtains $\sigma_{\mathcal{E}(\mathcal{B}^\gamma)}^2 = C_l^3$. Note that in both examples the results differ by a factor of 1/2. This can be traced back to the form of the best estimator in Eq. (79). The variance computed in [63] is consistent with his choice of the estimator. However, unlike what is stated in that paper, this variance does not deserve the name “cosmic” because, as we saw above, this estimator is not the best one. As expected, the variance of the best estimator is smaller than the variance computed in his article. Since we have now the correct expression for the cosmic variance, its numerical value should be reestimated. To be specific, let us take

$$\beta \equiv \begin{pmatrix} 2 & 2 & 2 \\ 1 & 1 & -2 \end{pmatrix}.$$

The cosmic variance is then $\sigma_{\mathcal{E}_{\text{best}}(\mathcal{B}^\beta)}^2 = C_2^{3/2} = (4\pi/5)^{3/2} Q_{\text{rms-PS}}^3 / T_0^3 \approx 1.3 \times 10^{-15}$ where we used $T_0 = 2.7 \text{ K}$ and $Q_{\text{rms-PS}} = 18.7 \mu\text{K}$ [67]. Although this figure is close to Luo’s result (1.4×10^{-15}) cited after Eq. (32) in [45], this does not imply that the two variances are not different by a factor of 1/2, as we have just seen. This might probably be due to a difference in the quadrupole normalizations.

V. CONCLUSIONS

Optimized analyses of CMB data sets involve the use of appropriate methods in order to reduce the various uncertainties. In particular, the theoretical error bars due to the cosmic variance can be minimized by working with the method of the best unbiased estimators. In this article, we have applied this technique for the study of CMB non-Gaussian features. These are often characterized by means of the third moment for the a_l^m ’s or by the angular bispectrum $C_{l_1 l_2 l_3}$. We have found the best unbiased estimators in both cases. These are the quantities that should be used in future data analyses and would be important for upcoming megapixel experiments [68,69] such as the microwave anisotropy probe (MAP)¹ and Planck Surveyor.² In addition to this, we have displayed both the angular bispectrum and the third moment cosmic vari-

¹<http://map.gsfc.nasa.gov/>

²<http://astro.estec.esa.nl/Planck/>

ances, the smallest possible uncertainties attached to the bispectrum and the third moment, which would be present in any ideal experiment when all other sources of noise have been removed.

ACKNOWLEDGMENTS

A.G. is a member of CONICET, Argentina.

APPENDIX: THE INVERSE TWO-POINT CORRELATION FUNCTION

In this appendix, we derive the exact expression of the inverse of the two-point correlation function ξ^{-1} , defined according to

$$\int d\Omega_j \xi(\mathbf{e}_i \cdot \mathbf{e}_j) \xi^{-1}(\mathbf{e}_j \cdot \mathbf{e}_k) \equiv \delta(\mathbf{e}_i \cdot \mathbf{e}_k - 1). \quad (\text{A1})$$

As usual, one expands the two-point correlation function on the basis of the Legendre polynomials

$$\xi(\mathbf{e}_i \cdot \mathbf{e}_j) = \sum_l \frac{2l+1}{4\pi} C_l P_l(\mathbf{e}_i \cdot \mathbf{e}_j). \quad (\text{A2})$$

In general, ξ^{-1} can be expanded on the spherical harmonics basis as follows:

$$\xi^{-1}(\mathbf{e}_i \cdot \mathbf{e}_j) = \sum_{lm} \sum_{l'm'} b_{ll'mm'} Y_l^m(\mathbf{e}_i) Y_{l'}^{m'*}(\mathbf{e}_j). \quad (\text{A3})$$

Our aim now is to determine the coefficients $b_{ll'mm'}$. Using Eqs. (A2) and (A3), the completeness relation for the Legendre polynomials together with the addition theorem of spherical harmonics in Eq. (A1), one gets

$$\sum_{lm} \sum_{l'm'} C_l b_{ll'mm'} Y_l^m(\mathbf{e}_i) Y_{l'}^{m'*}(\mathbf{e}_k) = \sum_{l''} \left(l'' + \frac{1}{2} \right) P_{l''}(\mathbf{e}_i \cdot \mathbf{e}_k). \quad (\text{A4})$$

The coefficients $b_{ll'mm'}$ are easy to read off and one obtains

$$b_{ll'mm'} = \frac{2\pi}{C_l} \delta_{ll'} \delta_{mm'}, \quad (\text{A5})$$

from which one deduces

$$\xi^{-1}(\mathbf{e}_i \cdot \mathbf{e}_j) = \sum_l \frac{l+1/2}{C_l} P_l(\mathbf{e}_i \cdot \mathbf{e}_j). \quad (\text{A6})$$

This is the expression used in the main text.

-
- [1] D. Scott, astro-ph/9912038.
 - [2] R. Scaramella and N. Vittorio, *Astrophys. J.* **375**, 439 (1991).
 - [3] M. Srednicki, *Astrophys. J. Lett.* **416**, L1 (1993).
 - [4] P. G. Ferreira, J. Magueijo, and K. M. Górski, *Astrophys. J. Lett.* **503**, L1 (1998).
 - [5] J. Pando, D. Vallas-Gabaud, and L. Fang, *Phys. Rev. Lett.* **79**, 1611 (1998).
 - [6] B. Bromley and M. Tegmark, *Astrophys. J. Lett.* **524**, L79 (2000).
 - [7] A. J. Banday, S. Zaroubi, and K. M. Górski, astro-ph/9908070.
 - [8] F. Bouchet and R. Gispert, *New Astron.* **4**, 443 (1999).
 - [9] M. Tegmark *et al.*, *Astrophys. J.* **530**, 133 (2000).
 - [10] J. Martin, A. Riazuelo, and M. Sakellariadou, *Phys. Rev. D* **61**, 083518 (2000).
 - [11] C. R. Contaldi, R. Bean, and J. Magueijo, *Phys. Lett. B* **468**, 189 (1999).
 - [12] A. Gangui, F. Lucchin, S. Matarrese, and S. Mollerach, *Astrophys. J.* **430**, 447 (1994).
 - [13] A. Gangui, *Phys. Rev. D* **50**, 3684 (1994).
 - [14] A. Linde and V. Mukhanov, *Phys. Rev. D* **56**, 535 (1997).
 - [15] A. Gangui and J. Martin, *Mon. Not. R. Astron. Soc.* **313**, 323 (2000).
 - [16] L. Wang and M. Kamionkowski, *Phys. Rev. D* **61**, 063504 (2000).
 - [17] X. Luo and D. N. Schramm, *Phys. Rev. Lett.* **71**, 1124 (1993).
 - [18] S. Mollerach, A. Gangui, F. Lucchin, and S. Matarrese, *Astrophys. J.* **453**, 1 (1995).
 - [19] D. Munshi, T. Souradeep, and A. A. Starobinsky, *Astrophys. J.* **454**, 552 (1995).
 - [20] F. R. Bouchet, D. P. Bennett, and A. Stebbins, *Nature (London)* **335**, 410 (1998).
 - [21] P. Ferreira and J. Magueijo, *Phys. Rev. D* **55**, 3358 (1997).
 - [22] P. P. Avelino *et al.*, *Astrophys. J. Lett.* **507**, L101 (1998).
 - [23] A. Gangui and L. Perivolaropoulos, *Astrophys. J.* **447**, 1 (1995).
 - [24] A. Gangui and S. Mollerach, *Phys. Rev. D* **54**, 4750 (1996).
 - [25] D. M. Goldberg and D. N. Spergel, *Phys. Rev. D* **59**, 103002 (1999).
 - [26] D. N. Spergel and D. M. Goldberg, *Phys. Rev. D* **59**, 103001 (1999).
 - [27] A. Cooray and W. Hu, *Astrophys. J.* **534**, 533 (2000).
 - [28] E. Bunn and D. Scott, astro-ph/9906044.
 - [29] T. Falk, R. Rangarajan, and M. Srednicki, *Astrophys. J. Lett.* **403**, L1 (1993).
 - [30] G. Hinshaw *et al.*, *Astrophys. J.* **431**, 1 (1994).
 - [31] G. Hinshaw *et al.*, *Astrophys. J. Lett.* **446**, L67 (1995).
 - [32] F. Bernardeau, *Astron. Astrophys.* **324**, 15 (1997).
 - [33] S. Winitzki, astro-ph/9806105.
 - [34] M. Zaldarriaga, *Phys. Rev. D* **62**, 063510 (2000).
 - [35] P. Coles, *Mon. Not. R. Astron. Soc.* **234**, 509 (1989).
 - [36] J. R. Gott *et al.*, *Astrophys. J.* **352**, 1 (1990).
 - [37] G. Smoot *et al.*, *Astrophys. J.* **437**, 1 (1994).
 - [38] J. R. Bond and G. Efstathiou, *Mon. Not. R. Astron. Soc.* **226**, 655 (1987).
 - [39] A. Kogut *et al.*, *Astrophys. J. Lett.* **439**, L29 (1995).
 - [40] A. Kogut *et al.*, *Astrophys. J. Lett.* **464**, L29 (1996).
 - [41] A. Heavens and R. Sheth, astro-ph/9904307.
 - [42] S. Winitzki and A. Kosowsky, *New Astron.* **3**, 75 (1998).

- [43] L. Popa, astro-ph/9806086.
- [44] M. Hobson *et al.*, astro-ph/9810200.
- [45] A. Heavens, Mon. Not. R. Astron. Soc. **299**, 805 (1998).
- [46] A. Guth, Phys. Rev. D **23**, 347 (1981).
- [47] A. Linde, Phys. Lett. **108B**, 389 (1983).
- [48] A. Albrecht and P. J. Steinhardt, Phys. Rev. Lett. **48**, 1220 (1982).
- [49] A. Linde, Phys. Lett. **129B**, 177 (1983).
- [50] V. F. Mukhanov and G. Chibisov, Pis'ma Zh. Eksp. Teor. Fiz. **33**, 5489 (1981) [JETP Lett. **33**, 532 (1981)].
- [51] S. W. Hawking, Phys. Lett. **115B**, 295 (1982).
- [52] A. A. Starobinsky, Phys. Lett. **117B**, 175 (1982).
- [53] J. M. Bardeen, P. J. Steinhardt, and M. S. Turner, Phys. Rev. D **28**, 679 (1983).
- [54] V. F. Mukhanov, H. A. Feldman, and R. H. Brandenberger, Phys. Rep. **215**, 203 (1992).
- [55] L. P. Grishchuk, Phys. Rev. D **48**, 3513 (1993).
- [56] J. Martin and D. J. Schwarz, Phys. Rev. D **57**, 3302 (1998).
- [57] J. Martin and D. J. Schwarz, Phys. Rev. D (to be published), astro-ph/9911225.
- [58] L. P. Grishchuk and Y. Sidorov, Phys. Rev. D **42**, 3413 (1990).
- [59] J. J. Halliwell, in *Quantum Cosmology and Baby Universes*, Proceedings of the 1989 Jerusalem Winter School for Theoretical Physics, edited by S. Coleman *et al.* (World Scientific, Singapore, 1989).
- [60] R. H. Brandenberger and C. T. Hill, Phys. Lett. B **179**, 30 (1986).
- [61] L. P. Grishchuk and J. Martin, Phys. Rev. D **56**, 1924 (1997).
- [62] K. M. Górski, Astrophys. J. Lett. **430**, L85 (1994).
- [63] X. Luo, Astrophys. J. Lett. **427**, L71 (1994).
- [64] J. Magueijo, Astrophys. J. Lett. **528**, L57 (2000).
- [65] M. White and M. Srednicki, Astrophys. J. **443**, 6 (1995).
- [66] L. Knox, Phys. Rev. D **60**, 103516 (1999).
- [67] E. F. Bunn and M. White, Astrophys. J. **480**, 6 (1997).
- [68] M. Tegmark, Astrophys. J. Lett. **480**, L87 (1997).
- [69] J. Borrill, astro-ph/9911389.