General and consistent statistics for cosmological observations

Ermis Mitsou [1,*] Jaiyul Yoo [1,2,†] Ruth Durrer [3,‡] Fulvio Scaccabarozzi,[1] and Vittorio Tansella[3]

[1]*Center for Theoretical Astrophysics and Cosmology, Institute for Computational Science, University of Zurich, CH–8057 Zürich, Switzerland*
[2]*Physics Institute, University of Zürich, Winterthurerstrasse 190, CH–8057, Zürich, Switzerland*
[3]*Département de Physique Théorique and Center for Astroparticle Physics, Université de Genève,*
24 quai Ansermet, CH–1211 Genève 4, Switzerland

This paper focuses on two aspects of the statistics of cosmological observables that are important for the next stages of precision cosmology. First, we note that the theory of reduced angular N-point spectra has only been developed in detail up to the trispectrum case and in a fashion that makes it difficult to go beyond. To fill this gap, here we present a constructive approach that provides a systematic description of reduced angular N-point spectra and their covariance matrices, for arbitrary N. Second, we focus on the common practice in the literature on cosmological observables, which consists in simply discarding a part of the expression, namely, the terms containing fields evaluated at the observer position. We point out that this is not justified beyond linear order in perturbation theory, as these terms contribute to all the multipoles of the corresponding spectra and with a magnitude that is of the same order as the rest of the nonlinear corrections. We consider the possibility that the reason for neglecting these terms is a conceptual discomfort when using ensemble averages, which originates in an apparent tension between the ergodic hypothesis and the privileged position of the observer on the light-cone. We clarify this subtle issue by performing a careful derivation of the relation between the theoretical statistical predictions and the observational estimators for all N. We conclude that there is no inconsistency whatsoever in ensemble-averaging fields at and near the observer position, thus clearing the way for consistent and robust high-precision calculations.

I. INTRODUCTION

Upcoming surveys such as Refs. [1–5] will significantly enhance the quantity and quality of the data that we have at our disposal for understanding the universe. This development requires a commensurate effort on the theoretical side for the new observational input to be interpreted correctly. In particular, one needs to consider analytical expressions for cosmological observables beyond the linear order in perturbation theory, as is already clearly reflected by the extensive amount of work on the subject in the case of CMB lensing [6–17], galaxy number counts [18–32], and cosmological distances and weak lensing [33–41]. The inclusion of nonlinear corrections is not only required for increasing the precision of the basic quantities of interest, such as the power spectrum of observables, but also for obtaining the leading contributions to higher-order statistics (e.g., bispectrum, trispectrum, etc.). The latter contain crucial information about the early universe and the formation of structure, which will become more

accessible thanks to the expected advances in observations. In this paper we will elaborate on two important aspects of the statistics of cosmological observables for the next stages of precision cosmology which have not yet been considered in the literature.

The first issue is the lack of generality in our description of angular higher-order statistics. Indeed, in the literature, the rotationally invariant (or "reduced") angular N-point spectra of a given observable on the sky are defined only up to the trispectrum ($N = 4$) case [42] and built in a rather opaque way that does not generalize easily to arbitrary N. The $N = 3, 4$ cases are well studied and measured [43–47] in CMB observations, whittling down the parameter spaces of the inflationary models. It is known, however, that the hierarchy between N-point spectra is not necessarily trivial, e.g., there exist models with a large trispectrum, but negligible bispectrum [48–51]. Therefore, the study of N-point spectra beyond $N = 4$ could be relevant for models with both negligible bispectrum and trispectrum. In any case, it is always desirable to have a systematic theoretical framework for approaching a given problem, which is what we will present here. This is a transparent construction of the reduced angular N-point spectra, for arbitrary N, their covariance matrices for arbitrary N and N', as well as a flurry of useful equations for manipulating them. The central ingredient in this framework will be the generalization of the "triangular" Wigner $3 - j$ symbol to a "multilateral" one, leading in particular to a neat diagrammatic representation. Our equations will hold for the idealized case of full sky measurements, but they should not

*ermitsou@physik.uzh.ch
†jyoo@physik.uzh.ch
‡ruth.durrer@unige.ch

be difficult to generalize to the case of partial covering. Note, finally, that the number of components of the angular spectra increases rapidly with N, meaning that the new higher-order statistics we can now work with are practically impossible to store, let alone manipulate, in their entirety. Nevertheless, our construction is the indispensable prerequisite for handling these objects, as it provides the starting point for research on privileged subsets, limits, shapes and efficient estimators thereof.

The second issue we want to discuss concerns the consistency and accuracy of the theoretical predictions, especially at the nonlinear orders in perturbation theory which now need to be taken into account. The standard approach to analytical computations of cosmological observables neglects certain terms, namely, those containing fields evaluated at the observer point. As we will discuss, for linear order perturbation theory these observer-dependent contributions have no effect on multipoles $l \geqslant 2$, but we will argue that they do affect all multipoles at the nonlinear level. This is an important point, because none of the works on nonlinear effects so far [6–41] considers these observer-dependent terms. Moreover, at the linear level it has already been established that the consideration of the full expressions, i.e., including the observer-dependent terms, has some important advantages, as it leads to results that are gauge-invariant and respect the equivalence principle in k-space (absence of spurious infrared divergences) [52–63].

We identify two possible reasons behind the reticence towards considering these observer-dependent terms in the literature. The first one is rooted in a common misconception that observer-dependent terms are synonymous with "monopole," because they only affect the later in the power spectrum of the linear theory (or the s multipole for spin s observables). Since low multipoles are dominated by cosmic variance, one is then naturally led to ignore observer-dependent terms. We will show that this is an artifact of linear perturbation theory, because observer-dependent terms actually do arise in nonlinear corrections to *all* multipoles, and there is no reason to believe that these contributions are negligible. The second possible reason for ignoring these terms might be a discomfort of conceptual nature. Simply put, it is not clear whether one can consistently ensemble average them to compute correlation functions and spectra, because, contrary to the case of source points, one does not probe several "observer points" observationally for the ergodic hypothesis to apply. A first shaky aspect of this argument is that the ensemble averages that are considered in the standard approach are statistically homogeneous, so this already wipes out the observationally privileged status of the observer point, independently of whether one chooses to include the observer-dependent terms or not. More precisely, we will see that the above objection to the statistical use of observer-dependent terms arises from a misunderstanding of the way the ergodic hypothesis enters the relation between the ensemble averages of the theorist and the geometrical averages of the observer. We will perform a careful derivation of that relation under the standard assumption of statistical homogeneity and isotropy for the stochastic fluctuations, but without any assumptions about which points are included in the ensemble averages. We will then see that the error of this theoretical prediction with

respect to the data averages is nothing but cosmic variance. Thus, by carelessly ensemble averaging field products at arbitrary points on the light-cone, independently of how well we can probe them observationally, the theorist is not deviating in any new way from observation other than the fundamental limitation of cosmic variance. Let us also stress that our result is completely general, as we will work with the freshly defined N-point spectra for arbitrary N. The assumption of full sky measurement will not be problematic here, as the point we wish to make is of qualitative nature. Also, we will only consider the spectra of a single observable, but the formalism generalizes straightforwardly in the case of cross-correlations of different observables.

The paper is organized as follows. In Sec. II we discuss some technical aspects regarding cosmological observables and their theoretical statistical treatment. The reader who is familiar with cosmological observables can skip this part, although some of its definitions and equations will be used later on, so it might be useful to take a quick look anyway. In Sec. III we discuss the present situation regarding reduced angular N-point spectra in the literature and then derive our formulas for defining and handling these objects for arbitrary N. Section IV focuses on the issue of observer-dependent terms. We first discuss in more detail the inconsistencies, and potential inaccuracies, that result from ignoring them in the spectra of observables. Then, using our definition of the angular correlation functions and spectra given in the previous section, we derive the relation between the statistical ones of the theorist and the geometrical ones of the observer. This demonstrates that one is allowed to (and one should) take into account the observer-dependent terms inside ensemble averages. In Sec. V we derive the expression for the covariance matrix of two arbitrary reduced angular spectra and determine the asymptotic behavior of cosmic variance for large l values. In Sec. VI we discuss a few subtleties one should be aware of when working with the connected parts of the statistics. Finally, in Sec. VII we conclude. The Appendices contain technical derivations to avoid clogging the main text, as well as the set of Wigner $3 - j$ symbol identities we will use in this paper.

II. PRELIMINARY CONSIDERATIONS

A. Angle, redshift, and observables

A cosmological observable associated with some localized source of light (galaxy, supernovae, etc.) is a function of two space-time points $\mathcal{O}(x_o; x_s)$, the "observer" point x_o and the "source" point x_s, that are constrained to lie on a common lightlike geodesic, with x_s in the past of x_o. Thus, for a given observer at x_o, the set of possible light source points x_s detected through light forms the past light-cone of x_o. The numbers x_s^μ are not observables, as they are ambiguous due to the freedom of performing coordinate transformations. One can of course consider correlation functions of the form $\langle \mathcal{O}(x_o; x_1) \dots \mathcal{O}(x_o; x_N) \rangle$, or the associated spectra in k-space, but these cannot be related to observable information, especially in the presence of nonnegligible inhomogeneity effects. One must therefore work directly with the physically unambiguous relations, i.e., the relations between observables, so

one must parametrize the past-light cone at x_o in terms of observables. One of them is the incoming photon direction in the sky, n, leading to an angular parametrization of the sky. Thus, instead of working with the "k-space" spectra, we will work with the "l-space" ones, which are directly related to observable quantities. Finally, for the radial parametrization of the light-cone there are several choices, such as the observed redshift z or the luminosity/angular distances $D_{L,A}$. Here we will consider the former, which is also the most widely used and model independent. Thus, the observer at x_o parametrizes her light-cone in terms of the observables z and n, and the physical information lies in the relation between $\mathcal{O}$ and (z, n) that is the function $\mathcal{O}(x_o; z, n)$.

The quantities z and n are defined with respect to the observer rest-frame, i.e., a tetrad e_a at x_o,

$$g_o(e_a, e_b) \equiv \eta_{ab}, \tag{1}$$

whose time-component e_0 is the observer's 4-velocity, and the source 4-velocity u_s satisfying $g_s(u_s, u_s) \equiv -1$. More precisely, if k denotes the momentum 4-vector of the photon, then we have

$$1 + z := \frac{g(u_s, k_s)}{g(e_0, k_o)}, \quad n_i := -\frac{g(e_i, k_o)}{\sqrt{g(e_j, k_o)\, g(e_j, k_o)}}, \tag{2}$$

where $i \in \{1, 2, 3\}$ is the spatial part of the four-dimensional index $a \in \{0, 1, 2, 3\}$. Note that n_i is the observed angular direction, not the propagation direction, hence the minus sign. These quantities are invariant under space-time coordinate transformations, as any observable should, since a measurement cannot depend on how we parametrize that manifold. In the absence of a tetrad, the only available numbers are the components of k in a coordinate system k^μ and the corresponding angles in the spatial part are the ones an observer uses only if the basis ∂_μ is orthonormal, i.e., only if $g_{\mu\nu}(x_o) = \eta_{\mu\nu}$. This is not the case in most of the coordinate systems used in cosmology, which is why one requires a tetrad to obtain the correct angles n. Since this field is defined at the observer position, the corresponding correction of the angles it induces will appear as observer terms in the expressions of observables.

The observer tetrad Eq. (1) is defined only up to a Lorentz transformation $e_a \rightarrow \Lambda_a{}^b e_b$, but the boosts alter the observer 4-velocity, so the only ambiguity is the orientation of the spatial frame $e_i \rightarrow R_i{}^j e_j$, leading to an SO(3) ambiguity for n,

$$n \rightarrow R^{-1} n. \tag{3}$$

As for the observable $\mathcal{O}$, the theoretical expression must also be a scalar under coordinate transformations [58],

$$\tilde{\mathcal{O}}(\tilde{x}_o; z, n) \equiv \mathcal{O}(x_o; z, n), \tag{4}$$

but it can be a tensor with respect to the Lorentz index a of the observer tetrad. For instance, ω and n are components of the Lorentz vector $k_a := g(e_a, k_o)$ at x_o

$$(k_a) = -\omega(1, n_i), \tag{5}$$

which is why they transform nontrivially under Lorentz transformations of e_a. Since the boost part here is fixed by the definite observer 4-velocity, we only have to deal with the

SO(3) ambiguity Eq. (3). To that end, we express n in terms of observed angles (ϑ, φ),

$$n \equiv \sin \vartheta \cos \varphi\, e_1 + \sin \vartheta \sin \varphi\, e_2 + \cos \vartheta\, e_3, \tag{6}$$

so that $\mathcal{O} \equiv \mathcal{O}(x_o; z, \vartheta, \varphi)$ becomes a function on the unit sphere S_2. The latter is a two-dimensional manifold with coordinates $\vartheta^A \in \{\vartheta, \varphi\}$ and with the admissible coordinate transformations being the ones induced by the rotation Eq. (3).[1] The generic observable will therefore be a tensor field on that manifold $\mathcal{O}_{A_1 \dots A_n}$. However, in two dimensions any tensor can be locally reduced to scalars and pseudoscalars. For instance, a vector can be decomposed into a scalar v and a pseudoscalar $\tilde{v}$ via

$$V_A = \nabla_A v + \varepsilon_A{}^B \nabla_B \tilde{v}, \tag{7}$$

while a tensor can be decomposed into two scalars T, t and two pseudo-scalars $\tilde{T}, \tilde{t}$

$$T_{AB} = S_{AB} T + \varepsilon_{AB} \tilde{T} + \left[\nabla_A \nabla_B - \tfrac{1}{2} S_{AB} \nabla^2 \right] t + \varepsilon^C_{(A} \nabla_{B)} \nabla_C \tilde{t}, \tag{8}$$

where S_{AB}, ε_{AB}, and ∇_A are the metric, totally antisymmetric tensor in two-dimensional and covariant derivative on the 2-sphere S_2, respectively. Further decomposing these (pseudo-)scalars into spherical harmonics leads to the decomposition of $\mathcal{O}_{A_1 \dots A_n}$ into spin-weighted spherical harmonics. To avoid the introduction of the latter, which complicates unnecessarily the formalism, here we will assume that our observables are the (pseudo-) scalars of the above decomposition (up to possible Laplacians). For instance, in the case of the CMB polarization tensor P_{AB} we would directly work with its "electric" and "magnetic" components $\nabla_A \nabla_B P^{AB}$ and $\varepsilon_A{}^C \nabla_B \nabla_C P^{AB}$, respectively.

Finally, another important observable is the incoming photon frequency

$$\omega := -g(e_0, k_o), \tag{9}$$

so the most general parametrization of light-based observables is a spectral distribution $\mathcal{O}(x_o; z, n, \omega)$. In the case of observables associated with diffuse sources (e.g., the CMB), i.e., that do not have a particular emission moment and therefore no associated redshift, then we simply have $\mathcal{O}(x_o; n, \omega)$. For the sake of simplicity, here we will assume that ω is either fixed, or that it is integrated over with some given spectral distribution, as in the case of the CMB temperature for instance. We will therefore work with observable functions of the form $\mathcal{O}(x_o; z, n)$, but the inclusion of ω is straightforward and basically enters our equations exactly as the redshift dependence. A more detailed account of this subsection's content, and in particular of the underlying geometrical constructions, can be found in Ref. [64].

B. Ensemble averaged N-point correlation functions

The theorist works with tensor fields on the full space-time manifold which we collectively denote by $\Phi(x)$. The

---

[1] Note that this manifold is a subset of the tangent space at x_o and with a parametrization that is induced by the tetrad basis e_a of that space, so these angles are not space-time coordinates.

observable $\mathcal{O}(x_o; z, n)$ is an integrodifferential functional of these fields

$$\mathcal{O}(x_o; z, n) \equiv \mathcal{O}(x_o; z, n)[\Phi]. \tag{10}$$

Within cosmological perturbation theory, one splits Φ into a homogeneous and isotropic "background" solution $\bar{\Phi}$ and a fluctuation ϕ,

$$\Phi = \bar{\Phi} + \phi, \tag{11}$$

where the latter is subject to a gauge ambiguity. This is the manifestation of infinitesimal coordinate transformations in this framework, since the background is kept fixed under these transformation. The observable $\mathcal{O}(x_o; z, n)$ should, by definition, be independent of the choice of gauge, so one can fix the latter. Moreover, one can use any constraint equations (e.g., the Poisson equation) to reduce the number of fields down to those carrying degrees of freedom, which we denote by $\phi_{\mathfrak{a}}(x)$, i.e., the "$\mathfrak{a}$" index includes both space-time and internal indices. Thus, the $\phi_{\mathfrak{a}}(t_0, \vec{x})$ data at a given time t_0 uniquely determine the ones at any other time t,

$$\phi_{\mathfrak{a}}(t, \vec{x}) \equiv \phi_{\mathfrak{a}}(t, \vec{x})[\phi_{\mathfrak{b}}(t_0, \vec{y})]. \tag{12}$$

Here it is assumed that the $\phi_{\mathfrak{a}}$ also contain the momenta/velocities, for fields obeying second-order equations in time. Next, to compare with observations, one needs to consider a statistical ensemble of solutions. The $\phi_{\mathfrak{a}}$ are therefore promoted to stochastic fields with a corresponding probability distribution functional (pdf) associating a probability density to each field solution $\phi_{\mathfrak{a}}(x)$. Since the latter are completely determined by their configuration at some reference time $\phi_{\mathfrak{a}}(t_0, \vec{x})$, it suffices to define that pdf on these field configurations $P \equiv P[\phi(t_0)]$ (dropping the index $\mathfrak{a}$ and the position $\vec{x}$ for notational simplicity). One can then define the moments, i.e., the statistical averages of field products,

$$\langle \phi_{\mathfrak{a}_1}(t_0, \vec{x}_1) \dots \phi_{\mathfrak{a}_n}(t_0, \vec{x}_n) \rangle$$
$$:= \int D\phi(t_0) \, P[\phi(t_0)] \, \phi_{\mathfrak{a}_1}(t_0, \vec{x}_1) \dots \phi_{\mathfrak{a}_n}(t_0, \vec{x}_n), \tag{13}$$

which completely determine the functional P, and with these one can define the *field* correlation functions (FCF),

$$F_{\mathfrak{a}_1 \dots \mathfrak{a}_n}(x_1, \dots, x_n) := \langle \phi_{\mathfrak{a}_1}(t_1, \vec{x}_1) \dots \phi_{\mathfrak{a}_n}(t_n, \vec{x}_n) \rangle$$
$$\equiv \langle \phi_{\mathfrak{a}_1}(t_1, \vec{x}_1)[\phi(t_0)] \dots \phi_{\mathfrak{a}_n}(t_n, \vec{x}_n)[\phi(t_0)] \rangle, \tag{14}$$

and use the linearity of the averaging operation to express this as a functional of Eq. (13). In particular, P is chosen such that

$$\langle \phi_{\mathfrak{a}}(x) \rangle \equiv 0, \tag{15}$$

which can be alternatively stated as

$$\langle \Phi(x) \rangle \equiv \bar{\Phi}(x). \tag{16}$$

With the above definitions one can now perform statistical averages of arbitrary functionals of the $\phi_{\mathfrak{a}}(x)$. In particular, the *theoretical* correlation functions (TCF) of the observables are defined by

$$G^{\text{th}}\big(x_o; \{z_k, n_k\}_{k=1}^N\big) := \left\langle \prod_{k=1}^N \mathcal{O}(x_o; z_k, n_k)[\Phi] \right\rangle, \tag{17}$$

which are therefore ultimately a functional of $\bar{\Phi}$ and the FCFs Eq. (14). Now note that, although strict homogeneity and isotropy are lost as soon as $\phi \neq 0$, these notions can be reintroduced at the statistical level by imposing the corresponding symmetries on $P[\phi(t_0)]$. Thus, we require that, if the two configurations $\phi_{\mathfrak{a}}(t_0)$ and $\phi'_{\mathfrak{a}}(t_0)$ are related by an isometry of the background geometry, then $P[\phi(t_0)] = P[\phi'(t_0)]$. As a result, the FCFs and TCFs are invariant under isometries, meaning in particular that Gth is independent of the observer position $\vec{x}_o$,

$$G^{\text{th}}\big(x_o; \{z_k, n_k\}_{k=1}^N\big) \equiv G^{\text{th}}\big(t_o; \{z_k, n_k\}_{k=1}^N\big), \tag{18}$$

and invariant under a common rotation of its N directions,

$$G^{\text{th}}\big(t_o; \{z_k, R n_k\}_{k=1}^N\big) \equiv G^{\text{th}}\big(t_o; \{z_k, n_k\}_{k=1}^N\big). \tag{19}$$

III. THE GENERAL THEORY OF REDUCED ANGULAR N-POINT SPECTRA

A. The previous implicit approach

Consider some observable on the sky $\mathcal{O}(n)$, where for the purposes of the present discussion we can omit the redshift dependence. The quantity that is usually considered in theory is the angular N-point spectrum through ensemble averaging $\langle \mathcal{O}_{l_1 m_1} \dots \mathcal{O}_{l_N m_N} \rangle$, where $\mathcal{O}_{lm}$ are the spherical harmonic components. In the standard approach where statistical isotropy is assumed, that N-point spectrum is invariant under rotations, so that not all of its components are independent. Its information can therefore be stored in a rotationally invariant quantity with fewer indices called the "reduced" angular spectrum. Assuming for simplicity here $\langle \mathcal{O}_{lm} \rangle = 0$, the lowest-order cases are the angular power spectrum,

$$\langle \mathcal{O}_{l_1 m_1} \mathcal{O}_{l_2 m_2} \rangle = \delta_{l_1 l_2} \delta_{m_1, -m_2} (-1)^{m_2} C_{l_1}, \tag{20}$$

where C_l is the "reduced angular power spectrum," and the angular bispectrum

$$\langle \mathcal{O}_{l_1 m_1} \mathcal{O}_{l_2 m_2} \mathcal{O}_{l_3 m_3} \rangle = \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} B_{l_1 l_2 l_3}, \tag{21}$$

where $B_{l_1 l_2 l_3}$ is the "reduced angular bispectrum" and the first factor is the Wigner $3 - j$ symbol. As one could expect from isotropy, C_l and $B_{l_1 l_2 l_3}$ have no m_k dependence. The relations Eqs. (20) and (21) can be inverted and, in particular, allow one to build unbiased estimators for these quantities,

$$\hat{C}_l := \frac{1}{2l+1} \sum_m \mathcal{O}_{lm} \mathcal{O}_{lm}^*,$$

$$\hat{B}_{l_1 l_2 l_3} = \sum_{m_1, m_2, m_3} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \mathcal{O}_{l_1 m_1} \mathcal{O}_{l_2 m_2} \mathcal{O}_{l_3 m_3}, \tag{22}$$

i.e.,

$$\langle \hat{C}_l \rangle = C_l, \quad \langle \hat{B}_{l_1 l_2 l_3} \rangle = B_{l_1 l_2 l_3}. \tag{23}$$

One can therefore work exclusively with the invariant and more compact reduced spectra, both at the theoretical and data level, instead of the full ones. The theory of the reduced angular N-point spectra is well developed in the literature up to the trispectrum case [43–47,65–68], based on the pioneering work in Ref. [42] which considered the following method. First, one looks for the relation between the full spectrum and

its reduced counterpart [e.g., Eqs. (20) and (21)] by imposing on the former the invariance under rotations,

$$
\langle \mathcal{O}_{l_1 m_1} \dots \mathcal{O}_{l_N m_N} \rangle
$$
$$
= \sum_{m'_k} \langle \mathcal{O}_{l_1 m'_1} \dots \mathcal{O}_{l_N m'_N} \rangle D_{l_1, m'_1 m_1} \dots D_{l_N, m'_N m_N}, \quad (24)
$$

where $D_{l,mm'}$ is the Wigner matrix, all of them here depending on the same rotation angles. Using the identities of the Wigner matrices and $3 - j$ symbols, one must then find the most general solution to this equation, which is Eqs. (20) and (21) for the $N = 2, 3$ cases, respectively. Finally, one must also invert this relation to obtain its estimator [e.g., Eq. (22)] and, through that quantity, its explicit definition in terms of ensemble averages of the observables [e.g., Eq. (23)]. Although this procedure generalizes straightforwardly to $N = 4$ [42] and beyond,[2] it becomes quickly very difficult in practice, as one has to "guess" the relation between the full and reduced spectrum by inferring it from an increasingly complicated equation. Moreover, even after finding a solution, one must still prove that it is the most general one and then invert it, which is also harder to do for larger values of N. As a result, the generalization of Eq. (22) for arbitrary N does not exist in the literature. Finally, with this procedure the reduced spectrum is determined only up to an overall normalization, which can depend on the l_i numbers, since only the m_i numbers are summed over in the defining Eq. (24).[3]

B. A new constructive approach

Here we propose a computationally straightforward approach to the problem. Since the defining property behind the notion of reduced spectra is rotational invariance, we will first focus on this geometrical aspect, i.e., before any statistical considerations. We will therefore be constructing the generalization of the full estimators Eq. (22), which exclusively depend on observed data, with their statistical counterparts being obtainable by simply taking the ensemble average as in Eq. (23). In fact, we will refer to these estimators as the "observational" reduced spectra, as opposed to the "theoretical" ones Eq. (17) which are obtained though statistical ensembles that are observationally unavailable. This is to highlight the fact that the observational reduced spectra have a significance of their own, as they amount to all the rotationally invariant information one can extract out of the observed data. This is the physical information, as it is independent of the artificial SO(3) ambiguity of the observer's spatial reference frame e_i in Eq. (6).

---

[2]Following this procedure, in Ref. [69] the authors have found the analog of Eqs. (20) and (21) for the quadrispectrum case $N = 5$. In Ref. [70] one can find a formal expression for arbitrary N, and also the real space N-point correlation function in terms of the reduced angular spectrum, but without rigorous demonstration. In both of these references, however, the inversion to find the estimators and ensemble-averaged spectra in terms of the observables, i.e., the analogues of Eqs. (22) and (23), is not given.

[3]See, for instance, Refs. [66,70], where alternative normalizations are considered and are essentially related to the l_i-dependent factors between the Gaunt and $3 - j$ coefficients.

We start with the fact that the observer only has access to a single light-cone x_o and therefore to the observables $\mathcal{O}(x_o; z, n)$. The building blocks for rotationally invariant functionals of $\mathcal{O}(x_o; z, n)$ are the average of the products $\prod_{k=1}^{N} \mathcal{O}(x_o; z_k, n_k)$ over all possible common rotations of the directions n_k, i.e., the average over the SO(3) group,

$$
\frac{\int dR \prod_{k=1}^{N} \mathcal{O}(x_o; z_k, R^{-1} n_k)}{\int dR}, \quad (25)
$$

where dR is the Haar measure on SO(3) and is invariant under group multiplication. Thanks to this, if one rotates the n_k in Eq. (25) with the same rotation R', then this can be reabsorbed in the dummy variable R by the redefinition $R \to R'R$, thus leaving the average invariant. To turn Eq. (25) into a well-defined integral, we consider an arbitrary orthonormal reference frame e_i and parametrize R as

$$
R(\alpha, \beta, \gamma) := R_3(\alpha) R_2(\beta) R_3(\gamma), \quad (26)
$$

where $R_i(\theta)$ denotes the matrix corresponding to a rotation around e_i with angle θ, so that α, β, γ are the Euler angles,

$$
\alpha \in [0, 2\pi[, \quad \beta \in [0, \pi], \quad \gamma \in [0, 2\pi[. \quad (27)
$$

In particular, the inverse matrix is simply

$$
R^{-1}(\alpha, \beta, \gamma) \equiv R(-\gamma, -\beta, -\alpha). \quad (28)
$$

The Haar measure dR on SO(3) now reads

$$
dR(\alpha, \beta, \gamma) \equiv \sin \beta \, d\alpha \, d\beta \, d\gamma, \quad (29)
$$

so the average over SO(3) of some function $f(n_1, \dots, n_N)$ is given by

$$
\langle f(n_1, \dots, n_N) \rangle_{\text{SO(3)}}
$$
$$
:= \frac{1}{8\pi^2} \int_0^{2\pi} d\alpha \int_0^{\pi} \sin \beta \, d\beta
$$
$$
\times \int_0^{2\pi} d\gamma \, f(R^{-1}(\alpha, \beta, \gamma) n_1, \dots, R^{-1}(\alpha, \beta, \gamma) n_N),
$$
$$
(30)
$$

and thus, the N-point *observational* correlation functions (OCF) are defined by

$$
G^{\text{ob}}(x_o; \{z_k, n_k\}_{k=1}^{N}) := \left\langle \prod_{k=1}^{N} \mathcal{O}(x_o; z_k, n_k) \right\rangle_{\text{SO(3)}}. \quad (31)
$$

In particular, the $N = 1$ case reduces to the average over the sphere,

$$
G^{\text{ob}}(x_o; z) \equiv \frac{1}{4\pi} \int d\Omega \, \mathcal{O}(x_o; z, n), \quad (32)
$$

which is shown by picking $e_3 = n$. For $N > 1$, three out of the $2N$ angles in G^{ob} are redundant, since we are free to rotate at will, or equivalently, to choose the reference frame $\{e_i\}_{i=1}^{3}$ arbitrarily. For instance, one can pick (assuming that n_1 and n_2 are not parallel)

$$
e_3 = n_1, \quad e_2 = \frac{n_2 - (n_1 n_2) n_1}{\sqrt{1 - (n_1 n_2)^2}}, \quad e_1 = \frac{n_1 \times n_2}{\sqrt{1 - (n_1 n_2)^2}}, \quad (33)
$$

thus leaving us with a dependence on the $2N - 3$ angles $\vartheta_2, \ldots \vartheta_N$ and $\varphi_3, \ldots, \varphi_N$ that parametrize $\{n_k\}_{k=2}^N$ in the e_i basis Eq. (6).

Now note that the N-point OCF is a redshift-dependent function on

$$S_2^N := \underbrace{S_2 \times \cdots \times S_2}_{N \text{ times}}, \qquad (34)$$

that is symmetric under a common rotation of these N spheres, just as its theoretical counterpart Eq. (19). In particular, expressing $\mathcal{O}$ inside G^{ob} in terms of the stochastic fields $\mathcal{O} \equiv \mathcal{O}[\Phi]$ and using the linearity of averaging, we obtain the identity

$$\left\langle G^{\text{ob}}\left(x_o; \{z_k, n_k\}_{k=1}^N\right) \right\rangle \equiv G^{\text{th}}\left(t_o; \{z_k, n_k\}_{k=1}^N\right), \qquad (35)$$

showing that the G^{ob} are full unbiased estimators of the Gth. The idea now is to decompose such functions in a basis of SO(3)-invariant functions on S_2^N. For $N = 2$ the only SO(3)-invariant combination of the two vectors is $n_1 n_2 \in [-1, 1]$ and a basis of functions on that interval are the Legendre polynomials P_l, so

$$G^{\text{ob}}(x_o; z_1, z_2, n_1, n_2) \equiv G_l^{\text{ob}}(x_o; z_1, z_2)(2l + 1)P_l(n_1 \cdot n_2), \qquad (36)$$

where

$$G_l^{\text{ob}}(x_o; z_1, z_2) := \frac{\mathcal{O}_{lm}(x_o; z_1)\, \mathcal{O}_{lm}^*(x_o; z_2)}{2l + 1}, \qquad (37)$$

is the observational two-point spectrum and $\mathcal{O}_{lm}$ are the harmonic components of $\mathcal{O}$,

$$\mathcal{O}(x_o; z, n) \equiv \mathcal{O}_{lm}(x_o; z)\, Y_{lm}(n). \qquad (38)$$

Note that we work with the unit-average normalization of the spherical harmonics Eq. (A5), which is the natural one in this context and that we keep the summation over l, m indices implicit for notational simplicity. In the presence of both dummy and free l, m indices, their nature will be inferable unambiguously by looking at both sides of the equation. The m indices will always be clearly associated to some l value and therefore run from $-l$ to l, while the l indices run from s to ∞, where s is the spin of the observable under consideration.[4] For $N > 2$, the decomposition is performed in detail in Appendix A for generic functions $f(n_1, \ldots, n_N)$ and the result is the following:

$$G^{\text{ob}}\left(x_o; \{z_k, n_k\}_{k=1}^N\right) \equiv G_{l_1 \ldots l_N | L_1 \ldots L_{N-3}}^{\text{ob}}\left(x_o; \{z_k\}_{k=1}^N\right) \\ \times Y_{l_1 \ldots l_N | L_1 \ldots L_{N-3}}\left(\{n_k\}_{k=1}^N\right), \qquad (39)$$

where

$$G_{l_1 \ldots l_N | L_1 \ldots L_{N-3}}^{\text{ob}}\left(x_o; \{z_k\}_{k=1}^N\right) := W_{m_1 \ldots m_N}^{l_1 \ldots l_N | L_1 \ldots L_{N-3}} \prod_{k=1}^N \mathcal{O}_{l_k m_k}(x_o; z_k) \qquad (40)$$

and

$$Y_{l_1 \ldots l_N | L_1 \ldots L_{N-3}}\left(\{n_k\}_{k=1}^N\right) := W_{m_1 \ldots m_N}^{l_1 \ldots l_N | L_1 \ldots L_{N-3}} \prod_{k=1}^N Y_{l_k m_k}(n_k), \qquad (41)$$

and where we have defined the following coefficients:

$$W_{m_1 \ldots m_N}^{l_1 \ldots l_N | L_1 \ldots L_{N-3}} := \begin{pmatrix} l_1 & l_2 & L_1 \\ m_1 & m_2 & -M_1 \end{pmatrix} \left[\prod_{k=1}^{N-4} (-1)^{L_k + M_k} \sqrt{2L_k + 1} \begin{pmatrix} L_k & l_{k+2} & L_{k+1} \\ M_k & m_{k+2} & -M_{k+1} \end{pmatrix} \right] \\ \times (-1)^{L_{N-3} + M_{N-3}} \sqrt{2L_{N-3} + 1} \begin{pmatrix} L_{N-3} & l_{N-1} & l_N \\ M_{N-3} & m_{N-1} & m_N \end{pmatrix}. \qquad (42)$$

Although the latter appear as a sum over the M_k indices, they are actually a single product because a Wigner $3 - j$ symbol vanishes if the sum of its m_k entries is not zero. Also for that reason, on can directly check that these coefficients are nonzero only if $\sum_{k=1}^N m_k = 0$.

Equations (39) to (42) are the main result of this section. Note that they can actually also encompass the $N = 1, 2$ cases if we define the redundant notation

$$G_{l_1}^{\text{ob}}(x_o; z_1) := \delta_{l_1}^0 G^{\text{ob}}(x_o; z_1), \quad G_{l_1 l_2}^{\text{ob}}(x_o; z_1, z_2) := (-1)^{l_1} \delta_{l_1 l_2} \sqrt{2l_1 + 1} \, G_{l_1}^{\text{ob}}(x_o; z_1, z_2), \qquad (43)$$

$$Y_{l_1}(n_1) := \delta_{l_1}^0, \quad Y_{l_1 l_2}(n_1, n_2) := (-1)^{l_1} \delta_{l_1 l_2} \sqrt{2l_1 + 1} \, P_{l_1}(n_1 \cdot n_2), \qquad (44)$$

and

$$W_{m_1}^{l_1} := \begin{pmatrix} l_1 & 0 & 0 \\ m_1 & 0 & 0 \end{pmatrix} \equiv \delta_0^{l_1} \delta_{m_1}^0, \quad W_{m_1 m_2}^{l_1 l_2} := \begin{pmatrix} l_1 & l_2 & 0 \\ m_1 & m_2 & 0 \end{pmatrix} \equiv \frac{(-1)^{l_1 + m_1}}{\sqrt{2l_1 + 1}} \delta^{l_1 l_2} \delta_{m_1, -m_2}. \qquad (45)$$

---

[4]For instance, we have $s = 0$ for CMB temperature maps, while $s = 2$ for the maps of the electric and magnetic components of the polarization field.

The "harmonic" components of the N-point OCF Eq. (40) are the observational N-point spectra which generalize the structures Eq. (22) to arbitrary N. In particular, for the $N = 3, 4$ cases we recover the known results [42], up to a different normalization convention factor $(-1)^L \sqrt{2L+1}$ for $N = 4$,

$$G^{\mathrm{ob}}_{l_1 l_2 l_3}(x_o; z_1, z_2, z_3) \equiv \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \prod_{k=1}^{3} \mathcal{O}_{l_k m_k}(x_o; z_k), \tag{46}$$

$$G^{\mathrm{ob}}_{l_1 l_2 l_3 l_4 | L}(x_o; z_1, z_2, z_3, z_4) \equiv (-1)^{L+M} \sqrt{2L+1} \begin{pmatrix} l_1 & l_2 & L \\ m_1 & m_2 & -M \end{pmatrix} \begin{pmatrix} L & l_3 & l_4 \\ M & m_3 & m_4 \end{pmatrix} \prod_{k=1}^{4} \mathcal{O}_{l_k m_k}(x_o; z_k). \tag{47}$$

As for the functions Eqs. (41), they form indeed a basis for SO(3)-invariant functions on S_2^N, as shown in Appendix A. They are orthonormal Eq. (A25) and therefore we can invert Eq. (39)

$$G^{\mathrm{ob}}_{l_1 \ldots l_N | L_1 \ldots L_{N-3}}(x_o; \{z_k\}_{k=1}^N) \equiv \int \left(\prod_{k=1}^{N} \frac{d\Omega_k}{4\pi} \right) Y^*_{l_1 \ldots l_N | L_1 \ldots L_{N-3}}(\{n_k\}_{k=1}^N) G^{\mathrm{ob}}(x_o; \{z_k, n_k\}_{k=1}^N). \tag{48}$$

Observe how the quantities Eqs. (40) and (41) are explicitly SO(3)-invariant, since they only depend on total angular momentum numbers l_k and L_k, making a total of $2N - 3$ indices. As noticed earlier, this is indeed the number of independent angles present in the corresponding correlation functions. Note also that Eqs. (40) and (41) are the very same combinations of the observable components and spherical harmonics, respectively. As we will see later on, all of the nontrivial contractions of indices will always be controlled by the coefficients Eq. (42) with no extra factors, which comes from the fact that they are orthonormal in the sense of Eq. (A23). Thus, contrary to the implicit method described in the previous Sec. III A, here there exists a normalization of the reduced spectra that appears as naturally privileged for computational convenience.

For the sake of completeness, let us also present the generalization of Eq. (23) by reminding that the TCFs Eq. (17) are SO(3)-invariant functions on S_2^N [see Eq. (19)], so they can also be decomposed in the basis Eq. (41). Using Eq. (35) and the linearity of averaging we find

$$G^{\mathrm{th}}_{l_1 \ldots l_N | L_1 \ldots L_{N-3}}(t_o; \{z_k\}_{k=1}^N) := \int \left(\prod_{k=1}^{N} \frac{d\Omega_k}{4\pi} \right) Y^*_{l_1 \ldots l_N | L_1 \ldots L_{N-3}}(\{n_k\}_{k=1}^N) G^{\mathrm{th}}(t_o; \{z_k, n_k\}_{k=1}^N)$$

$$\equiv \left\langle G^{\mathrm{ob}}_{l_1 \ldots l_N | L_1 \ldots L_{N-3}}(x_o; \{z_k\}_{k=1}^N) \right\rangle \equiv W^{l_1 \ldots l_N | L_1 \ldots L_{N-3}}_{m_1 \ldots m_N} \left\langle \prod_{k=1}^{N} \mathcal{O}_{l_k m_k}(x_o; z_k) \right\rangle. \tag{49}$$

Finally, to make contact with the implicit method described in Sec. III A, let us also invert Eq. (49) to obtain the generalization of Eqs. (20) and (21),

$$\left\langle \prod_{k=1}^{N} \mathcal{O}_{l_k m_k}(z_k) \right\rangle \equiv \int \left[\prod_{k=1}^{N} \frac{d\Omega_k}{4\pi} Y^*_{l_k m_k}(n_k) \right] \left\langle \prod_{k=1}^{N} \mathcal{O}(z_k, n_k) \right\rangle \equiv \int \left[\prod_{k=1}^{N} \frac{d\Omega_k}{4\pi} Y^*_{l_k m_k}(n_k) \right] G^{\mathrm{th}}(\{z_k, n_k\}_{k=1}^N)$$

$$\equiv G^{\mathrm{th}}_{l_1 \ldots l_N | L_1 \ldots L_{N-3}}(\{z_k\}_{k=1}^N) \int \left[\prod_{k=1}^{N} \frac{d\Omega_k}{4\pi} Y^*_{l_k m_k}(n_k) \right] Y_{l_1 \ldots l_N | L_1 \ldots L_{N-3}}(\{n_k\}_{k=1}^N)$$

$$\equiv W^{l_1 \ldots l_N | L_1 \ldots L_{N-3}}_{m'_1 \ldots m'_N} G^{\mathrm{th}}_{l_1 \ldots l_N | L_1 \ldots L_{N-3}}(\{z_k\}_{k=1}^N) \int \left[\prod_{k=1}^{N} \frac{d\Omega_k}{4\pi} Y^*_{l_k m_k}(n_k) Y_{l_k m'_k}(n_k) \right]$$

$$\equiv W^{l_1 \ldots l_N | L_1 \ldots L_{N-3}}_{m_1 \ldots m_N} G^{\mathrm{th}}_{l_1 \ldots l_N | L_1 \ldots L_{N-3}}(\{z_k\}_{k=1}^N). \tag{50}$$

This is the general solution to the equation of the implicit method Eq. (24) for arbitrary N, up to a l_k, L_k-dependent normalization. This result highlights again the naturalness of the normalization used here for the reduced angular spectra, as all the complicated factors entering any of the above equations always combine to form the multilateral Wigner symbols exactly. Note also that, with our construction, the consideration of the full spectra Eq. (50) becomes superfluous, as one no longer requires them to derive the reduced ones as in the implicit method of the previous subsection.

C. Multilateral diagrammatic representation

For $N = 3$ the coefficients defined in Eq. (42) simply reduce to the Wigner $3 - j$ symbol

$$W^{l_1 l_2 l_3}_{m_1 m_2 m_3} \equiv \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix}. \tag{51}$$
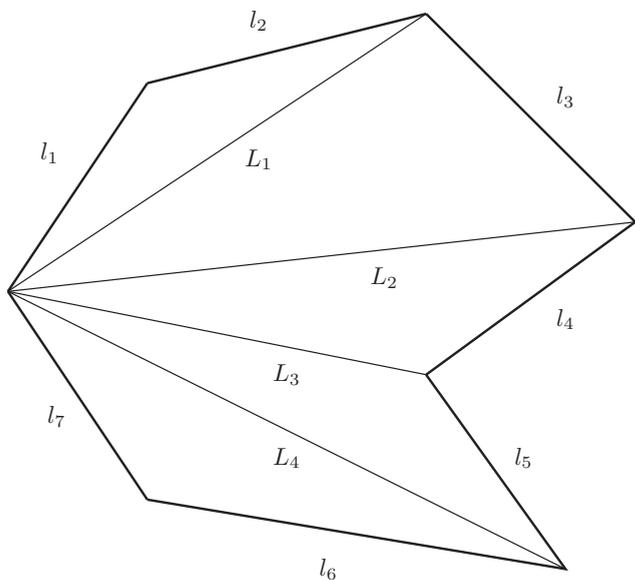
FIG. 1. A multilateral illustration for the case $N = 7$.

This quantity contains a $\{l_1 \; l_2 \; l_3\}$ factor, for which an alternative definition from Eq. (A22) is

$$\{l_1 \; l_2 \; l_3\}$$
$$:= \begin{cases} 1 & \text{if there exists a triangle with lengths } l_1, l_2, l_3 \\ 0 & \text{otherwise} \end{cases},$$

(52)

hence the common name "triangle delta." Consequently, the general coefficient Eq. (42) contains a $\{l_1 \ldots l_N | L_1 \ldots L_{N-3}\}$ factor, defined in Eq. (A24) as a product of certain $\{j_i, j_k, j_m\}$ factors. Generalizing the picture laid out in Refs. [42,69] for the $N = 4, 5$ cases,[5] this quantity Eq. (A24) can be interpreted as a "multilateral delta" in the following way. Being a product of $N - 2$ triangle deltas, it is nonzero only if all of the involved integers $l_1, \ldots, l_k$ and $L_1, \ldots, L_{N-3}$ form their respective triangles. Moreover, with respect to the triangle ordering in Eq. (A24), one of the edges of two neighboring triangles must have the same length L_k, so we can picture the triangles as being connected by common edges. The resulting shape is therefore a multilateral with N edges of lengths $l_1, \ldots, l_N$, while the $L_1, \ldots, L_{N-3}$ integers correspond to the lengths of the $N - 3$ diagonals connected to the vertex where the edges l_1 and l_N join (see Fig. 1). Thus, the multilateral delta is not nonzero for *any* set of $l_1, \ldots, l_N$ that can form a multilateral, but only for those whose diagonals also have integer length. This suggests naming the symbols $W^{\cdots}_{\cdots}$ defined in Eq. (42) the "multilateral Wigner symbols", with the $3 - j$ symbols therefore corresponding to the "triangular" case.

From this multilateral picture one can also directly infer that the decomposition Eq. (39) is not unique for $N > 3$, as was already recognized in Ref. [42] for the case $N = 4$.

---

[5]See also Sec. 2.4.1 of Ref. [70] for the same picture, and for arbitrary N, but in the case of the flat sky limit where the sphere S_2 is replaced by $\mathbb{R}^2$.

Indeed, there is an ambiguity in choosing the vertex with respect to which the diagonals are drawn. This ambiguity, modulo symmetries of the $3 - j$ symbols, corresponds to alternative orderings of the (l_k, m_k) pairs and therefore leads us to consider permutations of the multilateral Wigner symbol indices. As shown in detail [42] for the $N = 4$ case, these are achieved by taking linear combinations with Wigner $6 - j$ symbols, thus allowing one to relate all possible orderings.

IV. OBSERVER TERMS AND ENSEMBLE AVERAGING

A. The issue

An important limitation in cosmology is that, observationally, we have access only to one single realization of the statistical process under consideration, the universe, and only from one single vantage point. One therefore requires some assumptions to relate the theoretical predictions, which are of statistical nature, to observations. In the standard approach, we consider stochastic small initial fluctuations in a homogeneous and isotropic universe, which evolve under gravity. We also assume that their probability distribution functional is homogeneous and isotropic and that averaging over positions and directions is equivalent to ensemble average, i.e., an ergodic hypothesis. With these assumptions one can effectively treat different parts of the universe as different realizations and relate the theoretical ensemble averages to observational light-cone averages.

The requirement for the ergodic hypothesis to be applicable is the observational availability of a large enough number of source points. Indeed, the common motto in the literature is that the theorist is allowed to ensemble average a given product of fields evaluated at source positions, because the observer probes several of these positions. This viewpoint is perfectly sufficient as long as one is only interested in the lowest-order approximation where a cosmological observable is entirely determined by the value of fields at the source position. In the era of precision cosmology, however, the above approximation is no longer valid, because observations are able to capture several subleading effects which depend on fields evaluated on the whole line of sight from the source to the observer (e.g., weak lensing). As one approaches the observer along its past light-cone, the number of observable points decreases and therefore, one would naively infer, so does the degree of applicability of the ergodic hypothesis. The extreme case is the observer point itself, for which we can only have a single measurement. Are we then allowed to perform ensemble averages of fields at, and in the vicinity of, the observer point, as we implicitly do in Eq. (17)? The aim of this section is to answer this question and its conclusion is in the affirmative.

Before demonstrating this statement, however, let us explain in more detail why this question is relevant for precision cosmology. To that end, note first that the typical expression for a cosmological observable $\mathcal{O}$ to linear order in perturbation theory is of the form

$$\mathcal{O}^{(1)}(z, n) = X_{\mathrm{o}} + X_{\mathrm{los}}(z, n) + X_{\mathrm{s}}(z, n), \qquad (53)$$

where X_{o} and X_{s} denote field fluctuations evaluated at the observer and source positions, respectively, while X_{los} denotes an integral over fields evaluated along the background line of

sight. The quantity that is relevant for comparing with observations is then, e.g., the corresponding two-point correlation function, whose second-order contribution is

$$G_{(2)}^{\text{th}}(z, z', n, n') = \langle \mathcal{O}^{(1)}(z, n)\, \mathcal{O}^{(1)}(z', n') \rangle. \qquad (54)$$

Now note that the split Eq. (53) is not unique, as one can always transfer quantities between the three kinds of contributions by performing integrations by parts. Moreover, since we parametrize $\mathcal{O}$ with the observed redshift z and position in the sky n, the function $\mathcal{O}^{(1)}(z, n)$ is gauge-invariant, because the full observable $\mathcal{O}(z, n)$ has no dependence on coordinates whatsoever [58]. The split Eq. (53) can then be chosen such that each of the three terms are separately gauge-invariant. In practice, however, the expression Eq. (53) often presents itself with a gauge-dependent observer term X_{o}, which can therefore be set to zero with an appropriate choice of gauge, e.g., the synchronous gauge. Nevertheless, the expressions found in the literature are often in longitudinal gauge, in which case $X_{\text{o}} \neq 0$.

As already mentioned, one can consistently avoid the presence of observer terms (gauge-invariant or not) by an appropriate integration by parts, but one then needs to take into account the resulting extra terms in X_{los} and X_{s} that this manipulation introduces. Instead, the standard approach in the literature is to work in the longitudinal gauge and simply discard X_{o} as soon as it appears in the computation. The issue of whether one can ensemble average field values at the observer position therefore never arises in the standard practice. As for the line-of-sight terms X_{los}, they are ensemble averaged just like the source terms X_{s}, which implicitly assumes that ergodicity applies to all points on the light-cone, independently of how close they are to the observer point.

It is now well understood that discarding the observer terms X_{o} as described above leads to several *qualitative* problems:

(1) In most cases the resulting $\mathcal{O}^{(1)}$ is not gauge-invariant under gauge transformations at the observer because one discards a gauge-dependent term X_{o}. However, the gauge transformation of the observable only produces observer terms, by construction, which is therefore "ok" in this approach, since one is precisely neglecting such terms to begin with.

(2) Following the construction of Sec. II B, Fourier decomposing the field configurations $\phi(t_0)$ and using statistical homogeneity to eliminate one of the d^3k integrals, one obtains

$$G^{\text{th}}\big(t_o; \{z_k, n_k\}_{k=1}^{N}\big)$$
$$\equiv \int \left(\prod_{q=1}^{N-1} d^3 k_q\right) \tilde{G}^{\text{th}}\big(t_o, \{\vec{k}_q\}_{q=1}^{N-1}; \{z_k, n_k\}_{k=1}^{N}\big), \quad (55)$$

where the integrand is a combination of field k-spectra at t_0, multiplied by growth functions and $e^{i\vec{k}\cdot\vec{x}}$ factors evaluated on the light-cone. This integrand is generically divergent in the infrared, a fact which has led to confusion in the literature [71]. Moreover, the dependence on the infrared cut-off that has to be introduced to control the divergence in Eq. (55) breaks the equivalence principle, as the observable correlation function ends up depending on the uniform gravity mode. Taking into account X_{o} precisely cancels the divergent terms, thus showing that they cannot be associated with physical effects and that the equivalence principle holds as one would

expect [56,60,63].[6] As discussed at the beginning of Sec. II, the integrand in Eq. (55) is not itself an observable quantity, but it is interesting to see that the consideration of observer terms preserves the physical intuition in k-space.

(3) The sky parametrization n does not generically correspond to the one of an actual observer.[7] Indeed, the terms X_{o} contain corrections that implement the diffeomorphism relating the observer parametrization n to the one induced by the local coordinate system around the observer and therefore to the gauge that is chosen. In longitudinal gauge the coordinate-induced frame in the observer's tangent space is not orthonormal in the presence of vector and tensor modes, as the one of an observer should be, so neglecting X_{o} leads to spurious contributions to the "lensing" of observables in the sky.

Nevertheless, it turns out that there is no significant *quantitative* problem today in neglecting X_{o} *to linear order* in perturbation theory. This is because the actual observer approximates the ensemble average in Eq. (54) with the SO(3) averages of the data in the sky, denoted by $\langle \ldots \rangle_{\text{SO(3)}}$ in what follows. X_{o} is special with respect to this operation, because it has no angular dependence (if its spin vanishes) and can therefore be factored out. Consequently, the contribution $\langle X_{\text{o}}^2 \rangle_{\text{SO(3)}} \equiv X_{\text{o}}^2$ is a monopole term (for spin $s = 0$) in the corresponding power spectrum, while the cross-terms are simply zero, e.g.,

$$\langle X_{\text{o}} X_{\text{s}}(n, z) \rangle_{\text{SO(3)}} \equiv X_{\text{o}} \langle X_{\text{s}}(n, z) \rangle_{\text{SO(3)}} \approx 0. \qquad (56)$$

If $\mathcal{O}$ is an observable of nonzero spin s (e.g., for the observer velocity, where $s = 1$, or the CMB polarization, where $s = 2$), then the $\langle X_{\text{o}}^2 \rangle_{\text{SO(3)}}$ will contribute to the first nontrivial multipole which is the s multipole. But the overwhelming part of the relevant information actually lies in the rest of the spectrum, so ignoring the observer terms of cosmological observables is indeed irrelevant quantitatively.[8]

The caveat of the previous argumentation is that this is only true *to linear order in perturbation theory*. Indeed, to second order the solution of the observable will have cross-terms of the schematic form[9]

$$\mathcal{O}^{(2)}(z, n) \supset X_{\text{o}} X_{\text{s}}(z, n) + X_{\text{o}} X_{\text{los}}(z, n) + \ldots, \qquad (57)$$

---

[6]For a detailed demonstration and discussion of this cancellation in the two-point case, we refer the reader to Ref. [56] for the luminosity distance and to Ref. [63] for the galaxy number density observables.

[7]See Ref. [64] for a detailed explanation. In particular, note that while the introduction of the Sachs basis allows one to locally decompose tensors on the sky in the basis the observer uses, it does not provide the observer parametrization of the sky itself that is required to compute correlation functions and spectra as the observer does.

[8]Note that for some observables, such as galaxy number density and luminosity distance, the standard analysis is performed with correlation functions instead of spectra, in which case all multipoles contribute to the result and the lower-order ones are not subtracted. In that case, it has been shown that observer terms can be of the same order as other relativistic corrections [57,60].

[9]For a concrete example, see Eqs. (4.2) to (4.5) of Ref. [40] for the second-order part of the redshift perturbation.

implying terms of the form

$$G_{(3)}^{\text{th}}(z, z', n, n') \supset \langle X_{\text{o}} X_{\text{s}}(z, n) X_{\text{s}}(z', n') \rangle + \dots \quad (58)$$

for the next order of the two-point correlation function. Importantly, these contributions will affect *all* multipoles of the third-order power spectrum with an amplitude that is a priori comparable to the standard terms such as $\sim \langle X_{\text{s}}^2 X_{\text{los}} \rangle$, etc. Moreover, similar terms will also appear in higher-order correlation functions. Thus, in the era of precision cosmology one can no longer ignore the observer terms X_{o}, in the usual gauge choices that are considered, meaning that it is relevant to discuss their inclusion inside ensemble averages.

B. Relating theory and observation

We now wish to derive the relation between the observational and theoretical N-point functions of observables Eqs. (31) and (17). We already have Eq. (35), but this is a theoretical relation, as it involves ensemble averaging, and simply states that G^{ob} is a full unbiased estimator of Gth. Rather, we are interested in relating the two N-point functions as one does in practice, i.e., we need G^{ob} to be evaluated on a definite field configuration, and this is achieved using the ergodic hypothesis. In particular, we want to see how the latter enters the derivation and how it affects the issue of the single observer point.

We start by recalling Sec. II B and consider the set of three-dimensional field configurations $\phi_{\mathfrak{a}}(t_0)$ over which we sum when performing the ensemble average in Eq. (13). Given the invariance of $P[\phi(t_0)]$ under the isometry group, it is convenient to partition this ensemble into equivalence classes, where two field configurations are deemed equivalent if they can be related by an isometry. For the sake of simplicity, let us consider here the case of flat background space, so that the isometries form the Euclidean group $\mathbb{E} := \text{SO}(3) \ltimes \mathbb{R}^3$, and let us also use the trivial Cartesian coordinates. Any element of a given equivalence class can be described as an isometry of some fixed representative $\hat{\phi}_{\mathfrak{a}}(t_0)$,

$$\phi_{\mathfrak{a},R,\vec{c}}(t_0, \vec{x}) = M_{\mathfrak{a}}{}^{\mathfrak{b}}(R)\, \hat{\phi}_{\mathfrak{b}}(t_0, R\vec{x} + \vec{c}), \quad (59)$$

where R is a rotation matrix, $\vec{c}$ a translation vector and $M_{\mathfrak{a}}{}^{\mathfrak{b}}$ is the matrix that rotates tensor indices in $\mathfrak{a}$. We can therefore split the functional integration in Eq. (13) into an integral over the elements of a given class followed by an integral over all possible classes. By the latter we mean an integral over suitably chosen representatives $\hat{\phi}_{\mathfrak{a}}(t_0)$ such that the corresponding functional integral is well-defined.[10] Since the pdf is constant over all representatives of a given class, the integral in Eq. (13) becomes

$$\langle \phi_{\mathfrak{a}_1}(t_0, \vec{x}_1) \dots \phi_{\mathfrak{a}_n}(t_0, \vec{x}_n) \rangle$$
$$\equiv \frac{\int D\hat{\phi}(t_0)\, P[\hat{\phi}(t_0)]\, \langle \hat{\phi}_{\mathfrak{a}_1}(t_0, \vec{x}_1) \dots \hat{\phi}_{\mathfrak{a}_n}(t_0, \vec{x}_n) \rangle_{\mathbb{E}}}{\int D\hat{\phi}(t_0)\, P[\hat{\phi}(t_0)]}, \quad (60)$$

---

[10]The existence of such a splitting of the integration is a nontrivial mathematical assertion, whose proof, if possible, would go beyond the scope of this paper.

where we now integrate only over the set of representatives, and

$$\langle X[\phi(t_0)] \rangle_{\mathbb{E}} := \int \frac{d^3 c}{V} \int \frac{dR}{8\pi^2} X\big[M_{\mathfrak{a}}{}^{\mathfrak{b}}(R)\, \phi_{\mathfrak{b}}(t_0, R\vec{x} + \vec{c})\big]. \quad (61)$$

is the average over the Euclidean group action over the field configurations, V is the total volume and dR is the SO(3) Haar measure. As one may expect, the ensemble average therefore contains a purely geometric average over the symmetry group of the pdf.

Apart from subsets of measure zero, the configurations $\hat{\phi}(t_0)$ appearing in the integral Eq. (60) have a rich spatial dependence. In particular, they are nonperiodic functions which therefore probe a large variety of local field profiles for large enough V. In the $V \to \infty$ limit, which we consider here, the ergodic hypothesis states that this probing is thorough enough to make the Euclidean average in Eq. (60) independent of the configuration $\hat{\phi}(t_0)$. As a result, that average factorizes out of the integral, thus yielding the relation

$$\langle \phi_{\mathfrak{a}_1}(t_0, \vec{x}_1) \dots \phi_{\mathfrak{a}_n}(t_0, \vec{x}_n) \rangle \overset{\text{erg.}}{=} \langle \phi_{\mathfrak{a}_1}(t_0, \vec{x}_1) \dots \phi_{\mathfrak{a}_n}(t_0, \vec{x}_n) \rangle_{\mathbb{E}}. \quad (62)$$

Now on the right-hand side it is a definite, although generic field configuration that is considered. We can then generalize this manipulation straightforwardly to the case of the FCFs Eq. (14), since the Euclidean group action is independent of the time variable t, and therefore also to the case of the TCF Eq. (17),

$$G^{\text{th}}\big(t_o; \{z_k, n_k\}_{k=1}^N\big) \overset{\text{erg.}}{=} \left\langle \prod_{k=1}^N \mathcal{O}(x_o; z_k, n_k)[\Phi] \right\rangle_{\mathbb{E}}$$
$$\equiv \int \frac{d^3 c}{V} \int \frac{dR}{8\pi^2} \prod_{k=1}^N$$
$$\times \mathcal{O}(x_o; z_k, n_k)\big[M_{\mathfrak{a}}{}^{\mathfrak{b}}(R)\, \phi_{\mathfrak{b}}(t, R\vec{x} + \vec{c})\big], \quad (63)$$

where now we act with $\mathbb{E}$ directly on the four-dimensional fields $\phi_{\mathfrak{a}}(x)$ instead of the $\phi_{\mathfrak{a}}(t_0, \vec{x})$. Note also that we have not specified the $\bar{\Phi}$ dependence in Eq. (63) for simplicity. The above expression allows us to make contact with the OCFs Eq. (31). Let us first redefine the dummy variable $\vec{c} \to \vec{c} - R\vec{x}_o$ in Eq. (63) and let us also define the notation for shifted fields

$$\phi_{\mathfrak{a},\vec{c}}(t, \vec{x}) := \phi_{\mathfrak{a}}(t, \vec{x} + \vec{c}), \quad (64)$$

to get

$$G^{\text{th}}\big(t_o; \{z_k, n_k\}_{k=1}^N\big) \overset{\text{erg.}}{=} \int \frac{d^3 c}{V} \int \frac{dR}{8\pi^2} \prod_{k=1}^N$$
$$\times \mathcal{O}(x_o; z_k, n_k)\big[M_{\mathfrak{a}}{}^{\mathfrak{b}}(R)\, \phi_{\mathfrak{b},\vec{c}}(t, R(\vec{x} - \vec{x}_o))\big]. \quad (65)$$

We next observe that, for a given value of $\vec{c}$, the fields $\phi_{\mathfrak{a},\vec{c}}$ are rotated by R^{-1} around $\vec{x}_o$ in Eq. (65). But rotating the fields around $\vec{x}_o$ is tantamount to rotating the observed angles n in the opposite direction, so the SO(3) average on the fields

translates into an OCF

$$
\begin{aligned}
&G^{\mathrm{th}}\big(t_o; \{z_k, n_k\}_{k=1}^N\big) \\
&\stackrel{\mathrm{erg.}}{=} \int \frac{d^3 c}{V} \int \frac{dR}{8\pi^2} \prod_{k=1}^N \mathcal{O}(x_o; z_k, R^{-1} n_k)[\phi_{\mathfrak{a},\vec{c}}(t, \vec{x} - \vec{x}_o)] \\
&\equiv \int \frac{d^3 c}{V} G^{\mathrm{ob}}\big(x_o; \{z_k, n_k\}_{k=1}^N\big)[\phi_{\mathfrak{a},\vec{c}}(t, \vec{x} - \vec{x}_o)] \\
&\equiv \int \frac{d^3 c}{V} G^{\mathrm{ob}}\big(x_o; \{z_k, n_k\}_{k=1}^N\big)[\phi_{\mathfrak{a}}(t, \vec{x} - \vec{x}_o + \vec{c})]. \quad (66)
\end{aligned}
$$

Similarly, translating the fields by $\vec{x}_o - \vec{c}$ can be equivalently expressed as shifting the observer by $\vec{c} - \vec{x}_o$, so we finally obtain (after renaming $\vec{c} \to \vec{x}_o$),

$$
G^{\mathrm{th}}\big(t_o; \{z_k, n_k\}_{k=1}^N\big) \stackrel{\mathrm{erg.}}{=} \frac{1}{V} \int d^3 x_o\, G^{\mathrm{ob}}\big(t_o, \vec{x}_o; \{z_k, n_k\}_{k=1}^N\big), \tag{67}
$$

i.e., the average over the action of the Euclidean group on the fields $\phi_{\mathfrak{a}}$ is equivalent to averaging over all observer reference frames, as in G^{ob}, but *also* over all observer positions. This is simply because we have imposed both statistical isotropy *and* homogeneity. For instance, the CMB maps one would obtain from another vantage point of the universe $\vec{x}'_o$ would be different from the ones we observe on earth today, while the theoretical power spectrum which we calculate is the same for all vantage points. Since the relation between the correlation function and their associated spectra is linear, we have that Eq. (67) simply translates into

$$
\begin{aligned}
&G^{\mathrm{th}}_{l_1 \ldots l_N | L_1 \ldots L_{N-3}}\big(t_o; \{z_k, n_k\}_{k=1}^N\big) \\
&\stackrel{\mathrm{erg.}}{=} \frac{1}{V} \int d^3 x_o\, G^{\mathrm{ob}}_{l_1 \ldots l_N | L_1 \ldots L_{N-3}}\big(t_o, \vec{x}_o; \{z_k, n_k\}_{k=1}^N\big). \quad (68)
\end{aligned}
$$

Equations (67) and (68) are the central result of this section. The important thing to understand from them is that the application of the ergodic hypothesis singles out a definite field configuration on the right-hand side, but it does not single out the position of the actual observer, since all possible such positions are taken into account equally inside an average. Therefore, the applicability of the ergodic hypothesis depends on the representative field configuration which we choose on the right-hand side of Eq. (67), *completely independently of how this configuration is probed observationally*. In particular, as stated in the step in which ergodicity is actually used Eq. (62), what we need is that field configuration to have a rich profile, which is a decent assumption to make about the configuration of the actual universe.

What we also see, however, is that the use of ergodicity is necessary, but not sufficient to relate theory and observations yet, because the information from several $\vec{x}_o$ that is required in Eqs. (67) and (68) is not observationally available. Rather, the step which actually singles out the observer point is the approximation of the $\vec{x}_o$ averages in the right hand-sides of Eqs. (67) and (68) by the value at the actual observer position,

$$
G^{\mathrm{th}}\big(t_o; \{z_k, n_k\}_{k=1}^N\big) \stackrel{\mathrm{erg.}}{\approx} G^{\mathrm{ob}}\big(x_o; \{z_k, n_k\}_{k=1}^N\big) \tag{69}
$$

and

$$
G^{\mathrm{th}}_{l_1 \ldots l_N | L_1 \ldots L_{N-3}}\big(t_o; \{z_k, n_k\}_{k=1}^N\big) \stackrel{\mathrm{erg.}}{\approx} G^{\mathrm{ob}}_{l_1 \ldots l_N | L_1 \ldots L_{N-3}}\big(x_o; \{z_k, n_k\}_{k=1}^N\big), \tag{70}
$$

respectively. Approximating an average over some set by a single value within that set is of course the crudest possible estimator of that average. One expects this estimation to be accurate enough only if the values of the considered set exhibit a relatively small dispersion around the average value. But this dispersion in $\vec{x}_o$ space cannot be measured for the very same reason that brought us to this approximation in the first place, i.e., the observational unavailability of data at other observation points. Nevertheless, as we will see in the next subsection, the variance of G^{ob} with respect to the $\vec{x}_o$ dependence is equal to its theoretical analog, the statistical or "cosmic" variance, through the ergodic hypothesis. Thus, the error one makes in Eq. (69) or Eq. (70) is precisely cosmic variance, which is indeed negligible on small enough scales (see Sec. V B).

Equations (69) and (70) allow to relate theory and observation, as both sides can be computed in their respective domains. Finally, another way of interpreting the result Eq. (67) is by combining it with Eq. (35) to obtain

$$
\big\langle G^{\mathrm{ob}}\big(x_o; \{z_k, n_k\}_{k=1}^N\big)\big\rangle \stackrel{\mathrm{erg.}}{=} \frac{1}{V} \int d^3 x_o\, G^{\mathrm{ob}}\big(x_o; \{z_k, n_k\}_{k=1}^N\big), \tag{71}
$$

i.e., the ergodic hypothesis and statistical homogeneity equate the ensemble averaging with the translational averaging of a generic configuration. Thus, under these assumptions, while estimating G^{th} with G^{ob} first appears as a double approximation, i.e., the measurement of a single universe realization *and* from a single viewpoint, these two turn out to be the very same approximation.

C. Unambiguous covariance matrix and cosmic variance

Equation (71) shows that, assuming ergodicity and statistical homogeneity, ensemble averaging G^{ob} is tantamount to averaging it over $\vec{x}_o$ with a generic configuration of the fields. The last open question is therefore whether these two types of averaging are also equal when it comes to the variance associated with the estimation Eq. (69). The statistical covariance matrix is

$$
\begin{aligned}
&\mathrm{Cov}_{\mathrm{stat.}}(t_o; \alpha_N; \alpha'_{N'}) \\
&:= \big\langle [G^{\mathrm{ob}}(x_o; \alpha_N) - G^{\mathrm{th}}(\alpha_N)][G^{\mathrm{ob}}(x_o; \alpha'_{N'}) - G^{\mathrm{th}}(\alpha'_{N'})]\big\rangle \\
&\equiv \big\langle G^{\mathrm{ob}}(x_o; \alpha_N)\, G^{\mathrm{ob}}(x_o; \alpha'_{N'})\big\rangle - G^{\mathrm{th}}(\alpha_N)\, G^{\mathrm{th}}(\alpha'_{N'}), \quad (72)
\end{aligned}
$$

where α_N collectively denotes the set $\{z_k, n_k\}_{k=1}^N$ for notational simplicity, while the spatial one is

$$
\begin{aligned}
\mathrm{Cov}_{\mathrm{spat.}}(t_o; \alpha_N; \alpha'_{N'}) &:= \frac{1}{V} \int d^3 x_o [G^{\mathrm{ob}}(x_o; \alpha_N) - G^{\mathrm{th}}(\alpha_N)] \\
&\quad \times [G^{\mathrm{ob}}(x_o; \alpha'_{N'}) - G^{\mathrm{th}}(\alpha'_{N'})] \\
&\stackrel{\mathrm{erg.}}{=} \frac{1}{V} \int d^3 x_o\, G^{\mathrm{ob}}(x_o; \alpha_N)\, G^{\mathrm{ob}}(x_o; \alpha'_{N'}) \\
&\quad - G^{\mathrm{th}}(\alpha_N)\, G^{\mathrm{th}}(\alpha'_{N'}). \quad (73)
\end{aligned}
$$

The former leads to the notion of cosmic variance, so, if both of them are the same, we would have shown that our indiscriminate use of ensemble averaging leads to no new uncertainties between theory and observation. This is simply achieved by noting that a product of two OCFs can be expressed as a partial SO(3) average of a single OCF,

$$
\begin{aligned}
&G^{\mathrm{ob}}(x_o; \alpha_N)\, G^{\mathrm{ob}}(x_o; \alpha'_{N'}) \\
&\equiv \frac{1}{(8\pi^2)^2} \int dR \int dR' \prod_{k=1}^{N} \mathcal{O}(x_o; z_k, R^{-1} n_k) \\
&\quad \times \prod_{k'=1}^{N'} \mathcal{O}(x_o; z'_{k'}, R'^{-1} n'_{k'}) \\
&\stackrel{R \to RR'}{\equiv} \frac{1}{(8\pi^2)^2} \int dR \int dR' \prod_{k=1}^{N} \mathcal{O}(x_o; z_k, R'^{-1} R^{-1} n_k) \\
&\quad \times \prod_{k'=1}^{N'} \mathcal{O}(x_o; z'_{k'}, R'^{-1} n'_{k'}) \\
&\equiv \frac{1}{8\pi^2} \int dR\, G^{\mathrm{ob}}\big(x_o; \{z_k, R^{-1} n_k\}_{k=1}^{N}, \{z'_{k'}, n'_{k'}\}_{l=1}^{N'}\big). \quad (74)
\end{aligned}
$$

Inserting this in Eqs. (72) and (73) and using Eqs. (67) and (35) we find that both covariances are equal,

$$
\mathrm{Cov}_{\mathrm{spat.}} \stackrel{\mathrm{erg.}}{=} \mathrm{Cov}_{\mathrm{stat.}} =: \mathrm{Cov}, \quad (75)
$$

and that

$$
\begin{aligned}
\mathrm{Cov}(t_o; \alpha_N; \alpha'_M) &\equiv \frac{1}{8\pi^2} \int dR\, G^{\mathrm{th}}\big(\{z_k, R^{-1} n_k\}_{k=1}^{N}, \{z'_l, n'_l\}_{l=1}^{M}\big) \\
&\quad - G^{\mathrm{th}}(\alpha_N)\, G^{\mathrm{th}}(\alpha'_M). \quad (76)
\end{aligned}
$$

Thus, the absolute "1-sigma" error of the estimator Eq. (69),

$$
\Sigma(t_o; \alpha_N) := \sqrt{\mathrm{Cov}(t_o; \alpha_N; \alpha_N)}, \quad (77)
$$

is unambiguously cosmic variance. We see that our derivation provides a refined understanding of that notion, which is not usually expressed in cosmology textbooks. Given our statistical assumptions, cosmic variance is the error due to the fact that we observe a single realization of the universe *and* from a single vantage point $\bar{x}_o$. In accordance with the discussion of the previous subsection, if either of these two conditions were dropped, then there would be no cosmic variance. On one hand, if we had simultaneous access to the data of a single realization from all possible $\bar{x}_o$, then the ergodic hypothesis Eq. (67) would allow us to match the theoretical predictions exactly.[11] On the other hand, if we could observe all possible universe realizations, even from a single viewpoint $\bar{x}_o$, then we would be able to compute directly the theoretical N-point functions, which are independent of $\bar{x}_o$. In the first case we are technically setting $\Sigma_{\mathrm{spat.}} \to 0$, whereas in the second one it is rather $\Sigma_{\mathrm{stat.}} \to 0$.

---

[11]Here we neglect the fact that this information would also require a time $\sim V^{1/3}$ to be collected by a main observer and therefore analyzed.
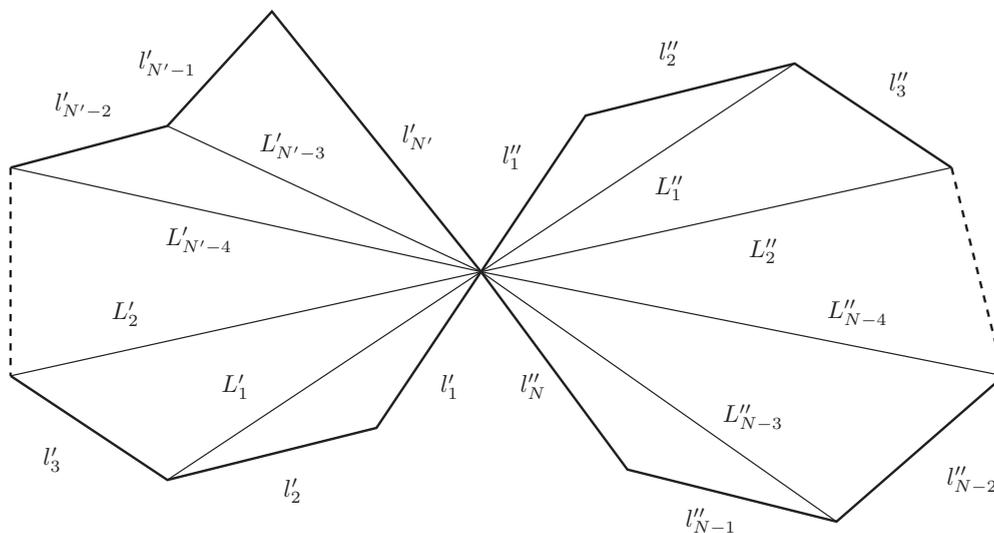
In the next section we will see that, although cosmic variance is always nonzero in practice, the associated relative uncertainty on the spectra must tend to zero when $l_k \to \infty$ as a consequence of statistical isotropy and the central limit theorem. Technically, this corresponds to $\Sigma_{\mathrm{stat.}}$ becoming negligible compared to G^{th} for large enough l_k values, with statistical homogeneity and the ergodic hypothesis then implying that so does $\Sigma_{\mathrm{spat.}}$, as we just saw in Eq. (75). This is the reason why, in practice, one can still obtain statistical cosmological information even from a single observational point.

V. COVARIANCE IN l-SPACE

A. The covariance matrix of angular N-point spectra

From now on we drop the dependence on t_o for notational simplicity. Here we wish to derive the expression for the covariance matrix of the reduced N-point and N'-point spectra, i.e., the harmonic analog of Eq. (76). To that end, we first note that $\mathrm{Cov}(t_o; \alpha_N; \alpha'_{N'})$ is a function on $S_2^N \times S_2^{N'}$ that is invariant under independent rotations on each of these two components, so we can decompose each of the corresponding arguments in the basis Eq. (41). The harmonic components are then found by projecting twice

$$
\begin{aligned}
&\mathrm{Cov}_{l_1 \ldots l_N | L_1 \ldots L_{N-3}; l'_1 \ldots l'_{N'} | L'_1 \ldots L'_{N'-3}}\big(\{z_k\}_{k=1}^{N}; \{z'_{k'}\}_{k'=1}^{N'}\big) \\
&:= \int \left(\prod_{k=1}^{N} \frac{d\Omega_k}{4\pi}\right)\left(\prod_{k'=1}^{N'} \frac{d\Omega'_{k'}}{4\pi}\right) \\
&\quad \times Y^*_{l_1 \ldots l_N | L_1 \ldots L_{N-3}}\big(\{n_k\}_{k=1}^{N}\big) Y^*_{l'_1 \ldots l'_{N'} | L'_1 \ldots L'_{N'-3}}\big(\{n'_{k'}\}_{k'=1}^{N'}\big) \\
&\quad \times \mathrm{Cov}\big(\{z_k, n_k\}_{k=1}^{N}; \{z'_{k'}, n'_{k'}\}_{k'=1}^{N'}\big). \quad (78)
\end{aligned}
$$

To express this in terms of the reduced theoretical spectra, we use Eq. (76) and decompose the TCFs in the basis Eq. (41). By proceeding exactly as in Appendix A to eliminate the SO(3) average in Eq. (76), and using the orthonormality relations Eqs. (A5) and (A23), we find

$$
\begin{aligned}
&\mathrm{Cov}_{l_1 \ldots l_N | L_1 \ldots L_{N-3}; l'_1 \ldots l'_{N'} | L'_1 \ldots L'_{N'-3}}\big(\{z_k\}_{k=1}^{N}; \{z'_{k'}\}_{k'=1}^{N'}\big) \\
&\equiv W^{l_1 \ldots l_N | L_1 \ldots L_{N-3}}_{m''_1 \ldots m''_N} W^{l''_1 \ldots l''_N l'_1 \ldots l'_{N'} | L''_1 \ldots L''_{N+N'-3}}_{m''_1 \ldots m''_N m'_1 \ldots m'_{N'}} W^{l'_1 \ldots l'_{N'} | L'_1 \ldots L'_{N'-3}}_{m'_1 \ldots m'_{N'}} \\
&\quad \times G^{\mathrm{th}}_{l''_1 \ldots l''_N l'_1 \ldots l'_{N'} | L''_1 \ldots L''_{N+N'-3}}\big(\{z_k\}_{k=1}^{N}, \{z'_{k'}\}_{k'=1}^{N'}\big) \\
&\quad - G^{\mathrm{th}}_{l_1 \ldots l_N | L_1 \ldots L_{N-3}}\big(\{z_k\}_{k=1}^{N}\big) G^{\mathrm{th}}_{l'_1 \ldots l'_{N'} | L'_1 \ldots L'_{N'-3}}\big(\{z'_{k'}\}_{k'=1}^{N'}\big). \quad (79)
\end{aligned}
$$

Now the second product of Wigner symbols can be simplified using Eq. (B5) iteratively, then Eq. (B6) and finally Eq. (B1) to find

$$
\begin{aligned}
&\sum_{m'_k} W^{l''_1 \ldots l''_N l'_1 \ldots l'_{N'} | L''_1 \ldots L''_{N+N'-3}}_{m''_1 \ldots m''_N m'_1 \ldots m'_{N'}} W^{l'_1 \ldots l'_{N'} | L'_1 \ldots L'_{N'-3}}_{m'_1 \ldots m'_{N'}} \\
&\equiv \{l'_1 \ldots l'_{N'} | L'_1 \ldots L'_{N'-3}\} \delta_{l''_N L''_{N-2}} \delta_{0 L''_{N-1}} \delta_{l'_1 L''_N} \\
&\quad \times \left(\prod_{k=1}^{N'-3} \delta_{L'_k L''_{N+k}}\right) W^{l''_1 \ldots l''_N | L''_1 \ldots L''_{N-3}}_{m''_1 \ldots m''_N}. \quad (80)
\end{aligned}
$$

FIG. 2. The type of diagram composing the covariance matrix of the N-point and N'-point spectra.

We therefore obtain

$$
\begin{aligned}
&\mathrm{Cov}_{l_1\ldots l_N|L_1\ldots L_{N-3};l_1'\ldots l_{N'}'|L_1'\ldots L_{N'-3}'}\big(\{z_k\}_{k=1}^N;\{z_{k'}'\}_{k'=1}^{N'}\big) \\
&\equiv A_{l_1\ldots l_N|L_1\ldots L_{N-3}}^{l_1''\ldots l_N''|L_1''\ldots L_{N-3}''} \\
&\quad \times G_{l_1''\ldots l_N'' l_1'\ldots l_{N'}'|L_1''\ldots L_{N-3}'' l_N'' 0 l_1' L_1'\ldots L_{N'-3}'}^{\mathrm{th}}\big(\{z_k\}_{k=1}^N,\{z_{k'}'\}_{k'=1}^{N'}\big) \\
&\quad - G_{l_1\ldots l_N|L_1\ldots L_{N-3}}^{\mathrm{th}}\big(\{z_k\}_{k=1}^N\big) G_{l_1'\ldots l_{N'}'|L_1'\ldots L_{N'-3}'}^{\mathrm{th}}\big(\{z_{k'}'\}_{k'=1}^{N'}\big), \quad (81)
\end{aligned}
$$

where we have defined the following contraction of multilateral Wigner symbols of equal order

$$
A_{l_1'\ldots l_N'|L_1'\ldots L_{N-3}'}^{l_1\ldots l_N|L_1\ldots L_{N-3}} := W_{m_1\ldots m_N}^{l_1\ldots l_N|L_1\ldots L_{N-3}} W_{m_1\ldots m_N}^{l_1'\ldots l_N'|L_1'\ldots L_{N-3}'}, \qquad (82)
$$

which is rotationally invariant, as it does not depend on m_k indices. From Eq. (81) we see that the covariance matrix of the N-point and N'-point spectra is controlled by the squeezed $(N+N')$-point spectrum, as depicted in Fig. 2, which arises from gluing together an N-point and an N'-point diagram. Therefore, the squeezed $(N+N')$-point spectrum here is already proportional to $\{l_1'\ldots l_{N'}'|L_1'\ldots L_{N'-3}'\}$, which is why we were able to discard that factor coming from Eq. (80).

The coefficients defined in Eq. (82) form a symmetric square matrix in the space of reduced N-point spectra and in Eq. (81) we see that this matrix acts on the N-point component of the squeezed $(N+N')$-point spectrum. Comparing with the analogous expression for the correlation functions Eq. (76), we thus see that the A matrix Eq. (82) implements the operation of partial SO(3) average on invariant functions in harmonic space. Finally, note that one can extend Eq. (81) to the cases $N, N' = 1, 2$ by using a redundant notation analogous to Eqs. (43), (44), and (45).

B. Cosmic variance in the large l_k limit

Let us consider the case $N = N'$, where the covariance matrix quantifies the typical error of the $G^{\mathrm{th}} \approx G^{\mathrm{ob}}(\vec{x}_o)$ approximation that is cosmic variance, so that the l-space analog

of Eq. (77) reads

$$
\begin{aligned}
&\Sigma_{l_1\ldots l_N|L_1\ldots L_{N-3}}\big(\{z_k\}_{k=1}^N\big) \\
&:= \sqrt{\mathrm{Cov}_{l_1\ldots l_N|L_1\ldots L_{N-3};l_1\ldots l_N|L_1\ldots L_{N-3}}\big(\{z_k\}_{k=1}^N\big)}. \quad (83)
\end{aligned}
$$

We now note that this quantity has a simple asymptotic behavior at large l_k, which is completely determined by the central limit theorem [72]. Indeed, thanks to the assumption of statistical isotropy, the $\mathcal{O}_{lm}(\vec{x}_o)$ for a given value of l are $2l + 1$ independent and equally distributed random variables. Thus, the product $\prod_{k=1}^N \mathcal{O}_{l_k m_k}(\vec{x}_o)$ for given l_k values consists of $\mathcal{O}(\prod_{k=1}^N l_k)$ independent and equally distributed random variables. The observational spectrum component $G_{l_1\cdots l_N|L_1\cdots L_{N-3}}^{\mathrm{ob}}(\vec{x}_o)$ then corresponds to a weighted sum of these products over their m_k indices Eq. (40). However, since the multilateral Wigner symbol Eq. (42) is zero unless $\sum_{k=1}^N m_k = 0$, the sum actually only contains $\mathcal{O}(\prod_{k=1}^{N-1} l_k)$ independent and equally distributed random variables. Finally, noting that $G^{\mathrm{ob}}(\vec{x}_o)$ is technically a sample average and G^{th} is its ensemble average, the central limit theorem [72] states that

$$
\Sigma_{l_1\cdots l_N|L_1\cdots L_{N-3}}\big|_{l_k \gg 1} \propto \frac{1}{\sqrt{\prod_{k=1}^{N-1} l_k}} G_{l_1\cdots l_N|L_1\cdots L_{N-3}}^{\mathrm{th}}. \qquad (84)
$$

Consequently, the relative cosmic variance tends to zero as $l_k \to \infty$. In particular, this means that $G_{l_1\cdots l_N|L_1\cdots L_{N-3}}^{\mathrm{ob}}(\vec{x}_o)$ tends to the $\vec{x}_o$-independent result $G_{l_1\cdots l_N|L_1\cdots L_{N-3}}^{\mathrm{th}}$, without the need to perform the average over $\vec{x}_o$ as in Eq. (68), thus justifying the approximation Eq. (69) for small enough scales. This is therefore how one can obtain accurate statistical cosmological information from a single vantage point for sufficiently large l_k. Following the discussion at the end of Sec. IV, this loss of the $\vec{x}_o$ information in the $l_k \to \infty$ limit is a consequence of statistical homogeneity and the ergodic hypothesis.

VI. CONNECTED ANGULAR *N*-POINT SPECTRA

A. Definitions, modified ergodic relation and covariance matrices

Here we focus on the part that is usually of most interest: the *connected* component of a given correlation function or associated spectrum. This is a nonlinear combination of the full statistics G, so one should expect the linear ergodic relations Eqs. (67) and (68) to be modified. Here we will illustrate this modification by considering the case of the connected two-, three-, and four-point functions. We start by defining the observable fluctuations, in the observational and theoretical cases, as the deviation of the observable from the respective one-point functions

$$\Delta^{\text{ob}}\mathcal{O}(\vec{x}_o; z, n) := \mathcal{O}(\vec{x}_o; z, n) - G^{\text{ob}}(\vec{x}_o; z), \quad (85)$$

$$\Delta^{\text{th}}\mathcal{O}(\vec{x}_o; z, n) := \mathcal{O}(\vec{x}_o; z, n) - G^{\text{th}}(z), \quad (86)$$

or, in terms of the harmonic components,

$$\Delta^{\text{ob}}\mathcal{O}_{lm}(\vec{x}_o; z) := \mathcal{O}_{lm}(\vec{x}_o; z) - \delta_l^0 \delta_m^0 G^{\text{ob}}(\vec{x}_o; z), \quad (87)$$

$$\Delta^{\text{th}}\mathcal{O}_{lm}(\vec{x}_o; z) := \mathcal{O}_{lm}(\vec{x}_o; z) - \delta_l^0 \delta_m^0 G^{\text{th}}(z). \quad (88)$$

We note in particular that, by construction, the observational monopole is identically zero, while the theoretical one captures precisely the difference between the two one-point functions

$$\Delta^{\text{ob}}\mathcal{O}_{00}(\vec{x}_o; z) \equiv G^{\text{ob}}(\vec{x}_o, z) - G^{\text{ob}}(\vec{x}_o, z) \equiv 0, \quad (89)$$

$$\Delta^{\text{th}}\mathcal{O}_{00}(\vec{x}_o; z) \equiv G^{\text{ob}}(\vec{x}_o, z) - G^{\text{th}}(z) \neq 0. \quad (90)$$

The connected two-, three-, and four-point functions can then be defined by

$$C^\star(z_1, z_2, n_1, n_2) := \langle \Delta^\star\mathcal{O}(z_1, n_1) \, \Delta^\star\mathcal{O}(z_2, n_2) \rangle_\star, \quad (91)$$

$$B^\star(z_1, z_2, z_3, n_1, n_2, n_3) := \langle \Delta^\star\mathcal{O}(z_1, n_1) \, \Delta^\star\mathcal{O}(z_2, n_2) \, \Delta^\star\mathcal{O}(z_3, n_3) \rangle_\star, \quad (92)$$

$$
\begin{aligned}
T^\star(z_1, z_2, z_3, z_4, n_1, n_2, n_3, n_4) := &\langle \Delta^\star\mathcal{O}(z_1, n_1) \, \Delta^\star\mathcal{O}(z_2, n_2) \, \Delta^\star\mathcal{O}(z_3, n_3) \, \Delta^\star\mathcal{O}(z_4, n_4) \rangle_\star \\
&- \langle \Delta^\star\mathcal{O}(z_1, n_1) \, \Delta^\star\mathcal{O}(z_2, n_2) \rangle_\star \langle \Delta^\star\mathcal{O}(z_3, n_3) \, \Delta^\star\mathcal{O}(z_4, n_4) \rangle_\star \\
&- \langle \Delta^\star\mathcal{O}(z_1, n_1) \, \Delta^\star\mathcal{O}(z_3, n_3) \rangle_\star \langle \Delta^\star\mathcal{O}(z_2, n_2) \, \Delta^\star\mathcal{O}(z_4, n_4) \rangle_\star \\
&- \langle \Delta^\star\mathcal{O}(z_1, n_1) \, \Delta^\star\mathcal{O}(z_4, n_4) \rangle_\star \langle \Delta^\star\mathcal{O}(z_2, n_2) \, \Delta^\star\mathcal{O}(z_3, n_3) \rangle_\star,
\end{aligned}
\quad (93)
$$

where here the star is a placeholder for "ob" and "th," the corresponding averages are respectively $\langle \ldots \rangle_{\text{SO}(3)}$ and $\langle \ldots \rangle$ and we have omitted the $\vec{x}_o$ dependencies in the "ob" case for notational simplicity. Clearly, the these functions are invariant under a common rotation of their arguments, so we can compute their harmonic components by projecting on the basis Eq. (41). Using the connected analog of Eq. (36) and also Eqs. (A15), (B5), (B6), and the $3 - j$ symbol symmetries, we recover the known expressions for the reduced power spectrum, bispectrum, and trispectrum estimators [42][12]

$$C_l^{\text{ob}}(\vec{x}_o; z_1, z_2) \equiv \frac{\Delta^{\text{ob}}\mathcal{O}_{lm}(\vec{x}_o; z_1) \, \Delta^{\text{ob}}\mathcal{O}_{lm}^*(\vec{x}_o; z_2)}{2l + 1}, \quad (94)$$

$$B_{l_1 l_2 l_3}^{\text{ob}}(\vec{x}_o; z_1, z_2, z_3) \equiv \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \prod_{k=1}^3 \Delta^{\text{ob}}\mathcal{O}_{l_k m_k}(\vec{x}_o; z_k), \quad (95)$$

$$
\begin{aligned}
T_{l_1 l_2 l_3 l_4 | L}^{\text{ob}}(\vec{x}_o; z_1, z_2, z_3, z_4) \equiv &(-1)^{L+M} \sqrt{2L+1} \begin{pmatrix} l_1 & l_2 & L \\ m_1 & m_2 & -M \end{pmatrix} \begin{pmatrix} L & l_3 & l_4 \\ M & m_3 & m_4 \end{pmatrix} \prod_{k=1}^4 \Delta^{\text{ob}}\mathcal{O}_{l_k m_k}(\vec{x}_o; z_k) \\
&- (-1)^{l_1+l_3} \sqrt{(2l_1+1)(2l_3+1)} \, \delta_{l_1 l_2} \delta_{l_3 l_4} \delta_{L0} C_{l_1}^{\text{ob}}(\vec{x}_o; z_1, z_2) \, C_{l_3}^{\text{ob}}(\vec{x}_o; z_3, z_4) \\
&- (-1)^{l_1+l_2} \sqrt{2L+1} \{ l_1 \ l_2 \ L \} \, \delta_{l_1 l_3} \delta_{l_2 l_4} C_{l_1}^{\text{ob}}(\vec{x}_o; z_1, z_3) \, C_{l_2}^{\text{ob}}(\vec{x}_o; z_2, z_4) \\
&- (-1)^L \sqrt{2L+1} \{ l_1 \ l_2 \ L \} \, \delta_{l_1 l_4} \delta_{l_2 l_3} C_{l_1}^{\text{ob}}(\vec{x}_o; z_1, z_4) \, C_{l_2}^{\text{ob}}(\vec{x}_o; z_2, z_3).
\end{aligned}
\quad (96)
$$

The theoretical counterparts are then obtained by replacing "ob" $\to$ "th" and ensemble averaging, again in accordance with the $N = 2, 3, 4$ results of Ref. [42].

Let us now see how the linear ergodic relation Eq. (68) looks like in terms of the connected parts. First, because of Eq. (89), the observational spectra vanish identically whenever at least one of their l_k entries is zero, i.e., they have vanishing "monopoles" by construction. In contrast, this is not the case for the theoretical spectra Eq. (90). As a concrete example, let us consider the relation between the connected two-point correlation functions, by expressing them in terms of the full functions and using

---

[12]Note that Eq. (96) agrees with Eqs. (26) and (20) of Ref. [42], as the normalization convention is different by an overall factor of $(-1)^L \sqrt{2L+1}$.

Eq. (67)

$$C^{\text{th}}(z_1, z_2, n_1, n_2) \equiv G^{\text{th}}(z_1, z_2, n_1, n_2) - G^{\text{th}}(z_1) G^{\text{th}}(z_2) \overset{\text{erg.}}{=} \frac{1}{V} \int d^3 x_o\, G^{\text{ob}}(\vec{x}_o; z_1, z_2, n_1, n_2) - G^{\text{th}}(z_1) G^{\text{th}}(z_2)$$

$$\equiv \frac{1}{V} \int d^3 x_o [C^{\text{ob}}(\vec{x}_o; z_1, z_2, n_1, n_2) + G^{\text{ob}}(\vec{x}_o; z_1) G^{\text{ob}}(\vec{x}_o; z_2)] - G^{\text{th}}(z_1) G^{\text{th}}(z_2)$$

$$\overset{\text{erg.}}{=} \frac{1}{V} \int d^3 x_o\, C^{\text{ob}}(\vec{x}_o; z_1, z_2, n_1, n_2) + \text{Cov}(z_1; z_2), \tag{97}$$

where $\text{Cov}(z_1; z_2)$ is the covariance matrix of the one-point function with itself [see Eq. (73)]. Alternatively, in harmonic space,

$$C_l^{\text{th}}(z_1, z_2) \overset{\text{erg.}}{=} \frac{1}{V} \int d^3 x_o\, C_l^{\text{ob}}(\vec{x}_o; z_1, z_2) + \delta_{l0} \text{Cov}(z_1; z_2). \tag{98}$$

The difference with Eq. (68) in the connected case is a monopole term, which is also exactly the multipole for which the observational spectrum vanishes identically $C_0^{\text{ob}}(z_1, z_2) \equiv 0$. Therefore, a better display of Eq. (98) could be

$$C_{l>0}^{\text{th}}(z_1, z_2) \overset{\text{erg.}}{=} \frac{1}{V} \int d^3 x_o\, C_{l>0}^{\text{ob}}(\vec{x}_o; z_1, z_2), \quad C_0^{\text{th}}(z_1, z_2) \overset{\text{erg.}}{=} \text{Cov}(z_1; z_2). \tag{99}$$

Thus, the ergodic relation Eq. (68) still holds for the components containing physical information $C_l^{\text{ob}} \neq 0$, while the leftover equation relates the covariance matrix of the one-point function to the monopole of the theoretical power spectrum. As a result, the absolute one-sigma error Eq. (77) associated with the one-point function approximation,

$$G^{\text{ob}}(x_o; z) \approx G^{\text{th}}(z) \pm \Sigma(z), \tag{100}$$

is $\hspace{12cm}$ simply

$$\Sigma(z) := \sqrt{\text{Cov}(z; z)} \overset{\text{erg.}}{=} \sqrt{C_0^{\text{th}}(z, z)}. \tag{101}$$

This picture generalizes to the case of higher N, as one can check by expressing $\Delta^{\text{th}} \mathcal{O}_{lm}$ in terms of $\Delta^{\text{ob}} \mathcal{O}_{lm}$ in the theoretical spectra B^{th} and T^{th}. One finds that the resulting equation splits into two sets of equations: the nonmonopole ones, i.e., the ones where none of the l_k is zero and for which Eq. (68) still holds, and the ones relating the theoretical monopoles to the covariance matrix and higher order analogues (skewness, kurtosis, etc.) of lower-N spectra. For this reason, from now on all of our equations will hold up to monopole terms, so that we do not have to deal with this subtlety.

The next nontrivial case is the trispectrum, where one now subtracts products of two-point functions. For the purposes of our point it suffices to describe it schematically as

$$T^{\text{th}}\big(\{n_k, z_k\}_{k=1}^4\big) \equiv \left\langle \prod_{k=1}^{4} \Delta^{\text{th}} \mathcal{O}(\vec{x}_o; z_k, n_k) \right\rangle - C^{\text{th}}(z_1, z_2, n_1, n_2) C^{\text{th}}(z_3, z_4, n_3, n_4) - \dots, \tag{102}$$

where the ellipses here denote the other two possible orderings of the two-point function arguments. In complete analogy with the derivation in the $N = 2$ case Eq. (97), using Eq. (67) we find

$$T^{\text{th}}\big(\{n_k, z_k\}_{k=1}^4\big) \equiv G^{\text{th}}\big(\{n_k, z_k\}_{k=1}^4\big) - G^{\text{th}}(z_1, z_2, n_1, n_2) G^{\text{th}}(z_3, z_4, n_3, n_4) - \dots + \text{mon.}$$

$$\overset{\text{erg.}}{=} \frac{1}{V} \int d^3 x_o\, G^{\text{ob}}\big(\vec{x}_o; \{n_k, z_k\}_{k=1}^4\big) - G^{\text{th}}(z_1, z_2, n_1, n_2) G^{\text{th}}(z_3, z_4, n_3, n_4) - \dots + \text{mon.}$$

$$\equiv \frac{1}{V} \int d^3 x_o \Big[C^{\text{ob}}\big(\vec{x}_o; \{n_k, z_k\}_{k=1}^4\big) + G^{\text{ob}}(\vec{x}_o; z_1, z_2, n_1, n_2) G^{\text{ob}}(\vec{x}_o; z_3, z_4, n_3, n_4) + \dots \Big]$$

$$- G^{\text{th}}(z_1, z_2, n_1, n_2) G^{\text{th}}(z_3, z_4, n_3, n_4) - \dots + \text{mon.}$$

$$\overset{\text{erg.}}{=} \frac{1}{V} \int d^3 x_o\, C^{\text{ob}}\big(\vec{x}_o; \{n_k, z_k\}_{k=1}^4\big) + \text{Cov}(z_1, z_2, n_1, n_2; z_3, z_4, n_3, n_4) + \dots + \text{mon.}, \tag{103}$$

so the difference is again a covariance matrix, but now the one of the two-point function with itself. In harmonic space we then find

$$T_{l_1 l_2 l_3 l_4 | L}^{\text{th}}(z_1, z_2, z_3, z_4) \overset{\text{erg.}}{=} \frac{1}{V} \int d^3 x_o\, T_{l_1 l_2 l_3 l_4 | L}^{\text{ob}}(\vec{x}_o; z_1, z_2, z_3, z_4) + (-1)^{l_1 + l_3} \sqrt{(2l_1 + 1)(2l_3 + 1)}\, \delta_{l_1 l_2} \delta_{l_3 l_4} \delta_{L0} \text{Cov}_{l_1; l_3}(z_1, z_2; z_3, z_4)$$

$$+ (-1)^{l_1 + l_2} \sqrt{2L + 1}\, \{l_1 \ l_2 \ L\}\, \delta_{l_1 l_3} \delta_{l_2 l_4} \text{Cov}_{l_1; l_2}(z_1, z_3; z_2, z_4)$$

$$+ (-1)^{L} \sqrt{2L + 1}\, \{l_1 \ l_2 \ L\}\, \delta_{l_1 l_4} \delta_{l_2 l_3} \text{Cov}_{l_1; l_2}(z_1, z_4; z_2, z_3) + \text{mon.}, \tag{104}$$

where we recognize the same structure as in Eq. (96) by construction and we have the covariance matrix of the two-point spectra

$$\text{Cov}_{l;l'}(z_1, z_2; z_3, z_4) \stackrel{\text{erg.}}{=} \frac{1}{V} \int d^3 x_o \, G_l^{\text{ob}}(\vec{x}_o; z_1, z_2) \, G_{l'}^{\text{ob}}(\vec{x}_o; z_3, z_4) - G_l^{\text{th}}(z_1, z_2) \, G_{l'}^{\text{th}}(z_3, z_4). \tag{105}$$

A first important difference with the two-point case Eq. (98) is that the extra terms here affect multipoles of arbitrary magnitude, although only a small subset, the one of pairwise equal l_k's. A second important difference is that the components of T^{ob} with pairwise equal l_k's are not identically zero in general, i.e., the equations no longer split as in the monopole case. One must therefore take Eq. (104) as it is and infer that the analog of the single-observer approximation Eq. (70) in this case is actually

$$T_{l_1 l_2 l_3 l_4 | L}^{\text{ob}}(\vec{x}_o; z_1, z_2, z_3, z_4) \stackrel{\text{erg.}}{\approx} T_{l_1 l_2 l_3 l_4 | L}^{\text{th}}(z_1, z_2, z_3, z_4) - (-1)^{l_1 + l_3} \sqrt{(2l_1 + 1)(2l_3 + 1)} \, \delta_{l_1 l_2} \delta_{l_3 l_4} \delta_{L0} \text{Cov}_{l_1; l_3}(z_1, z_2; z_3, z_4)$$

$$- (-1)^{l_1 + l_2} \sqrt{2L + 1} \, \{l_1 \;\; l_2 \;\; L\} \, \delta_{l_1 l_3} \delta_{l_2 l_4} \text{Cov}_{l_1; l_2}(z_1, z_3; z_2, z_4)$$

$$- (-1)^{L} \sqrt{2L + 1} \, \{l_1 \;\; l_2 \;\; L\} \, \delta_{l_1 l_4} \delta_{l_2 l_3} \text{Cov}_{l_1; l_2}(z_1, z_4; z_2, z_3) + \text{mon.} \tag{106}$$

Finally, just as the two-point case Eq. (98) with $l = 0$ provides the cosmic variance of the 1-point function for free, the four-point case Eq. (104) with $l_1 = l_2 =: l$, $l_3 = l_4 =: l'$ and $L = 0$ provides the cosmic variance for the power spectrum. However, here $T_{ll'l'l'|0}^{\text{ob}} \neq 0$, but rather [use Eq. (96) and properties of the $3 - j$ symbols],

$$T_{ll'l'l'|0}^{\text{ob}}(\vec{x}_o; z_1, z_2, z_3, z_4) \equiv -\delta_{ll'} \big[C_l^{\text{ob}}(\vec{x}_o; z_1, z_3) C_l^{\text{ob}}(\vec{x}_o; z_2, z_4) + (z_3 \leftrightarrow z_4) \big], \tag{107}$$

so, inserting this in Eq. (104), using Eq. (106) and isolating the covariance matrix, we find

$$\text{Cov}_{l;l'}(z_1, z_2, z_3, z_4) \equiv \frac{(-1)^{l+l'}}{\sqrt{(2l + 1)(2l' + 1)}} \Big[\delta_{ll'} \big[C_l^{\text{th}}(z_1, z_3) C_l^{\text{th}}(z_2, z_4) + (z_3 \leftrightarrow z_4) \big] + T_{ll'l'l'|0}^{\text{th}}(z_1, z_2, z_3, z_4) \Big] + \text{mon.} \tag{108}$$

This generalizes the result of Ref. [42], worked out for the CMB where there is no redshift dependence, to the case of generic redshifts along the light-cone. In particular, the cosmic variance of the power spectrum, i.e., the absolute 1-sigma error Eq. (77) in the approximation

$$C_l^{\text{ob}}(\vec{x}_o; z, z') \approx C_l^{\text{th}}(z, z') \pm \Sigma_l(z, z'), \tag{109}$$

is

$$\Sigma_l^2(z, z') := \text{Cov}_{l;l}(z, z', z, z') \equiv \frac{1}{2l + 1} \Big[\big[C_l^{\text{th}}(z, z') \big]^2 + C_l^{\text{th}}(z, z) C_l^{\text{th}}(z', z') + T_{llll|0}^{\text{th}}(z, z', z, z') \Big] + \text{mon.} \tag{110}$$

Finally, for Gaussian statistics and equal redshifts, we recover the well-known result[13]

$$\Sigma_l(z, z)|_{\text{Gauss.}} = \sqrt{\frac{2}{2l + 1}} \, C_l^{\text{th}}(z, z). \tag{111}$$

Given the discussion of Sec. V B and comparing Eq. (110) with Eq. (84), we find that here the central limit theorem essentially states that the trispectrum decays faster with growing l than the C_l^2 terms.

B. Mind the monopole when using relative fluctuations

Until now we have considered the "absolute" fluctuations $\Delta^{\text{ob}} \mathcal{O}$ and $\Delta^{\text{th}} \mathcal{O}$, but in practice the most convenient ones to work with are the relative ones (if $G^{\text{th}}(z) \neq 0$),

$$\delta^{\text{ob}} \mathcal{O}(\vec{x}_o; z, n) := \frac{\Delta^{\text{ob}} \mathcal{O}(\vec{x}_o; z, n)}{G^{\text{ob}}(\vec{x}_o; z)}, \qquad \delta^{\text{th}} \mathcal{O}(\vec{x}_o; z, n) := \frac{\Delta^{\text{th}} \mathcal{O}(\vec{x}_o; z, n)}{G^{\text{th}}(z)}. \tag{112}$$

This introduces a nonlinear difference between the observational and theoretical definitions and will affect the covariance matrices of the corresponding spectra. Indeed, consider for instance the case of the relative power spectra

$$\tilde{C}_l^{\text{ob}}(\vec{x}_o; z_1, z_2) := \frac{\delta^{\text{ob}} \mathcal{O}_{lm}(\vec{x}_o; z_1) \, \delta^{\text{ob}} \mathcal{O}_{lm}^*(\vec{x}_o; z_2)}{2l + 1} \equiv \frac{C_l^{\text{ob}}(\vec{x}_o; z_1, z_2)}{G^{\text{ob}}(\vec{x}_o; z_1) \, G^{\text{ob}}(\vec{x}_o; z_2)}, \tag{113}$$

$$\tilde{C}_l^{\text{th}}(z_1, z_2) \stackrel{\text{erg.}}{=} \frac{1}{V} \int d^3 x_o \, \frac{\delta^{\text{th}} \mathcal{O}_{lm}(\vec{x}_o; z_1) \, \delta^{\text{th}} \mathcal{O}_{lm}^*(\vec{x}_o; z_2)}{2l + 1} \equiv \frac{C_l^{\text{th}}(z_1, z_2)}{G^{\text{th}}(z_1) \, G^{\text{th}}(z_2)}, \tag{114}$$

which are denoted using tilded letters to distinguish them from the absolute ones C_l^{ob} and C_l^{th}. The corresponding covariance matrix is given by

$$\widetilde{\text{Cov}}_{l;l'}(z_1, z_2, z_3, z_4) \stackrel{\text{erg.}}{=} \frac{1}{V} \int d^3 x_o \big[\tilde{C}_l^{\text{ob}}(\vec{x}_o; z_1, z_2) - \tilde{C}_l^{\text{th}}(z_1, z_2) \big] \big[\tilde{C}_{l'}^{\text{ob}}(\vec{x}_o; z_3, z_4) - \tilde{C}_{l'}^{\text{th}}(z_3, z_4) \big] \tag{115}$$

---

[13]See e.g., [73] for the case of the CMB power spectrum.

and one would naively expect this to be simply

$$\widetilde{\text{Cov}}_{l;l'}(z_1, z_2, z_3, z_4) = \frac{\text{Cov}_{l;l'}(z_1, z_2, z_3, z_4)}{G^{\text{th}}(z_1)\,G^{\text{th}}(z_2)\,G^{\text{th}}(z_3)\,G^{\text{th}}(z_4)} + \text{mon.}, \tag{116}$$

where here $\text{Cov}_{l;l'}$ is the covariance matrix of the absolute two-point spectrum Eq. (108). In particular, one would then infer the analog of Eq. (111) for the corresponding cosmic variance in the equal redshift Gaussian case,

$$\tilde{\Sigma}_l(z, z)|_{\text{Gauss.}} = \sqrt{\frac{2}{2l+1}}\,\tilde{C}_l^{\text{th}}(z, z). \tag{117}$$

However, as we will now show, Eqs. (116) and (117) are actually only approximate, because to obtain them one must wrongly assume that $G^{\text{ob}}(\vec{x}_o; z) = G^{\text{th}}(z)$ or, according to Eq. (89), that the monopole of the observable is zero at $\vec{x}_o$ for all z,

$$\delta^{\text{th}}\mathcal{O}_{00}(\vec{x}_o, z) = 0. \tag{118}$$

Indeed,

$$\text{Cov}_{l;l'}(z_1, z_2, z_3, z_4) \stackrel{\text{erg.}}{=} \frac{1}{V}\int d^3x_o \big[C_l^{\text{ob}}(\vec{x}_o; z_1, z_2)\,C_{l'}^{\text{ob}}(\vec{x}_o; z_3, z_4) \big] - C_l^{\text{th}}(z_1, z_2)\,C_{l'}^{\text{th}}(z_3, z_4) + \text{mon.} \tag{119}$$

$$\equiv \frac{1}{V}\int d^3x_o\, G^{\text{ob}}(\vec{x}_o; z_1)\,G^{\text{ob}}(\vec{x}_o; z_2)\,G^{\text{ob}}(\vec{x}_o; z_3)\,G^{\text{ob}}(\vec{x}_o; z_4)\,\tilde{C}_l^{\text{ob}}(\vec{x}_o; z_1, z_2)\,\tilde{C}_{l'}^{\text{ob}}(\vec{x}_o; z_3, z_4)$$

$$- G^{\text{th}}(z_1)\,G^{\text{th}}(z_2)\,G^{\text{th}}(z_3)\,G^{\text{th}}(z_4)\,\tilde{C}_l^{\text{th}}(z_1, z_2)\,\tilde{C}_{l'}^{\text{th}}(z_3, z_4) + \text{mon.}$$

$$\equiv G^{\text{th}}(z_1)\,G^{\text{th}}(z_2)\,G^{\text{th}}(z_3)\,G^{\text{th}}(z_4)[\widetilde{\text{Cov}}_{ll'}(z_1, z_2, z_3, z_4) + \tilde{R}_{ll'}(z_1, z_2, z_3, z_4)] + \text{mon.}, \tag{120}$$

where the remainder term $\tilde{R}_{ll'}$ depends on the monopole $\delta^{\text{th}}\mathcal{O}_{00}(\vec{x}_o; z_k)$ that comes from converting the $G^{\text{ob}}(\vec{x}_o, z_k)$ into $G^{\text{ob}}(z_k)$ and which cannot go through the integral over $\vec{x}_o$, because of its dependence on that variable. The remainder term is therefore controlled by a combination of monopoles of the theoretical power spectrum, bispectrum and trispectrum which, unlike the extra "mon." terms described in the previous section, contribute here to *all* l values, not just to $l = 0$. Thus, it is important to know whether one compares absolute quantities or relative ones, because the corresponding covariance matrices are different and therefore so will be the results of the corresponding Fisher forecasts.

VII. CONCLUSION

In this paper we have addressed two potentially important issues for the next stages of precision cosmology. First, we have developed the general theory of the rotationally invariant ("reduced") angular N-point spectra. The method employed so far [42] is not suited for obtaining expressions for arbitrary N, because it requires solving some increasingly complicated algebraic equation as N grows. As a result, one must work out each N case separately and the present literature contains a detailed description only up to the $N = 4$ case [42]. We have presented an alternative construction which provides all the relevant definitions and relations associated with reduced angular N-point spectra straightforwardly and for all N. This includes the covariance matrix of these spectra for arbitrary N and N', which is controlled by a squeezed $(N + N')$-point spectrum and leads to a cosmic variance of the N-point spectrum that decays as $l^{(N-1)/2}$ for large l_k numbers. Our construction is based on the introduction of the "multilateral" Wigner symbols, which generalize the Wigner $3 - j$ symbols of triangles to polygons. With these we have

built an orthonormal harmonic basis for N-point correlation functions on the sphere, such that the reduced spectra appear as the coefficients in this basis. We have also discussed a corresponding diagrammatic representation, generalizing the one of Refs. [42,69] for the cases $N = 4, 5$, which was also considered in Ref. [70] in the flat sky limit. Even though the determination of a generic N-point spectrum might be prohibitive numerically, our general framework allows easily for specialization, e.g., to squeezed spectra with only one l varying while the others are fixed to some low value.

The second important part of the paper consists in motivating and justifying the consideration of observer terms of cosmological observables inside the ensemble averages of correlation functions and spectra. Through a careful derivation of the relation between theoretical predictions ("G^{th}") and observations ("G^{ob}"), under the assumptions of statistical homogeneity and isotropy and ergodicity, we have shown that no special treatment of the observer point is required whatsoever on the theoretical side. Ensemble averaging field products at this point does not introduce any inconsistencies, nor does it imply any new assumptions or uncertainties. This motivates us to consider the *full* analytical expressions of cosmological observables in calculations, especially at nonlinear order in perturbations where observer terms are intertwined with source and line-of-sight contributions in a nontrivial way, as we argued in Sec. IV A. Thus, with this conceptual ambiguity being lifted, we can now safely state that a rigorous treatment of cosmological observables should include the observer terms, or at least a check of the magnitude of their contribution in the final result if they are neglected.

We have also taken advantage of the present framework to discuss some subtle aspects of working with the connected part of spectra, as one usually does, when relating theory to observation. When working directly with fluctuations of

observables around their average (one-point function) value, as is often the case in cosmology, the difference between full and connected spectra is simply a sum of products of spectra with $N > 1$. However, here we pointed out that the notion of fluctuation is different in the theoretical and observational cases, since the one-point functions are different (ensemble versus sky average). As a result, the relation between theoretical and observational *connected* spectra involves extra "monopole" terms. Moreover, if one works with relative, instead of absolute, fluctuations of observables, then one must take into account additional corrections which now affect all multipoles, not just the monopoles. A direct consequence of this fact is that the covariance matrices computed with the relative and absolute fluctuations are not proportional to each other, but have an additional remainder term. Thus, this can lead to different results in the corresponding Fisher analyses, depending on whether one uses absolute or relative observational fluctuations.

Finally, and remarkably, the fact that the observational and theoretical one-point functions are not equal in general might also be relevant in the context of parameter estimation involving CMB data. Indeed, the standard treatment of the CMB temperature is to simply neglect cosmic variance and therefore consider these two quantities as equal $\bar{T} := \langle T \rangle = \langle T \rangle_{\mathrm{SO(3)}}(\vec{x}_o)$. This is justified by the high precision of the average CMB temperature measurement [74] and, most importantly, by the smallness of the associated relative cosmic variance $C_0 \sim 10^{-5}$, which sets the fundamental uncertainty floor. In principle, however, one should consider $\bar{T}$ (or ω_γ) as an extra cosmological parameter to be varied in the likelihood analysis. Through a careful Fisher analysis that incorporates the above-mentioned subtleties [75], the impact of an independent $\bar{T}$ on the other cosmological parameters and their errors is estimated in Ref. [76].

APPENDIX A: HARMONIC DECOMPOSITION OF SO(3)-INVARIANT FUNCTIONS ON S_2^N

Consider a function f on S_2^N that satisfies

$$f(n_1, \ldots, n_N) \equiv f(Rn_1, \ldots, Rn_N), \quad \text{(A1)}$$

for arbitrary rotations R. We can then average over the SO(3) group on the right-hand side [see Eq. (30)],

$$f(n_1, \ldots, n_N) \equiv \langle f(n_1, \ldots, n_N) \rangle_{\mathrm{SO(3)}}. \quad \text{(A2)}$$

We then decompose each one of the n_k dependencies into spherical harmonics

$$f(n_1, \ldots, n_N) \equiv f_{m_1 \ldots m_N}^{l_1 \ldots l_N} Y_{l_1 m_1}(n_1) \ldots Y_{l_N m_N}(n_N), \quad \text{(A3)}$$

where

$$f_{m_1 \ldots m_N}^{l_1 \ldots l_N} := \int \left[\prod_{k=1}^{N} \frac{d\Omega_k}{4\pi} Y_{l_k m_k}^*(n_k) \right] f(n_1, \ldots, n_N). \quad \text{(A4)}$$

The summation over l, m indices will be kept implicit for notational simplicity. In what follows, we will encounter both dummy and free l, m indices, so their nature will be inferable by looking at both sides of the equation. The m indices will always be clearly associated to some l value and therefore run from $-l$ to l, while the l indices run from 0 to ∞. Also, note that we use the less conventional normalization of spherical harmonics,

$$\frac{1}{4\pi} \int d\Omega \, Y_{lm}(n) \, Y_{l'm'}^*(n) \equiv \delta_{ll'} \delta_{mm'}. \quad \text{(A5)}$$

Equation (A2) now reads

$$f(n_1, \ldots, n_N) \equiv f_{m_1 \ldots m_N}^{l_1 \ldots l_N} \times \frac{1}{8\pi^2} \int dR \prod_{k=1}^{N} Y_{l_k m_k}(R^{-1} n_k), \quad \text{(A6)}$$

where the Haar measure dR is given in Eq. (29). We can then extract the R-dependence out of the spherical harmonics by using their transformation property under rotations,

$$Y_{lm}(R^{-1}(\alpha, \beta, \gamma) n) \equiv Y_{lm'}(n) D_{l, m'm}(\alpha, \beta, \gamma), \quad \text{(A7)}$$

where the D_l are the Wigner matrices forming the $(2l + 1)$-dimensional irreducible representation of SO(3),

$$D_l(R) D_l(R') \equiv D_l(RR'), \quad D_l(R) D_l^\dagger(R) \equiv \mathbb{I}. \quad \text{(A8)}$$

We thus have

$$f(n_1, \ldots, n_N) \equiv I_{m_1' \ldots m_N', m_1 \ldots m_N}^{l_1 \ldots l_N} f_{m_1 \ldots m_N}^{l_1 \ldots l_N} \prod_{k=1}^{N} Y_{l_k m_k'}(n_k), \quad \text{(A9)}$$

where

$$I_{m_1' \ldots m_N', m_1 \ldots m_N}^{l_1 \ldots l_N} := \frac{1}{8\pi^2} \int_0^{2\pi} d\alpha \int_0^{\pi} \sin \beta \, d\beta$$
$$\times \int_0^{2\pi} d\gamma \prod_{k=1}^{N} D_{l_k, m_k' m_k}(\alpha, \beta, \gamma). \quad \text{(A10)}$$

In the $N = 2$ case, we can use the identity

$$D_{l, mm'}^* \equiv (-1)^{m+m'} D_{l, -m-m'}, \quad \text{(A11)}$$

and the orthonormality relation

$$\frac{1}{8\pi^2} \int_0^{2\pi} d\alpha \int_0^{\pi} \sin \beta \, d\beta \int_0^{2\pi} d\gamma \, D_{l_1, m_1 m_1'}(\alpha, \beta, \gamma)$$
$$\times D_{l_2, m_2 m_2'}^*(\alpha, \beta, \gamma)$$
$$\equiv \frac{1}{2l_1 + 1} \delta_{l_1 l_2} \delta_{m_1 m_2} \delta_{m_1' m_2'}, \quad \text{(A12)}$$

to obtain

$$I_{m_1' m_2', m_1 m_2}^{l_1 l_2} \equiv \frac{(-1)^{m_1 + m_1'}}{2l_1 + 1} \delta_{l_1 l_2} \delta_{-m_1 m_2} \delta_{-m_1' m_2'}, \quad \text{(A13)}$$

and thus

$$f(n_1, n_2) \equiv f_l (2l + 1) P_l(n_1 n_2), \quad \text{(A14)}$$

where the P_l are the Legendre polynomials, arising from the identity

$$P_l(n_1 n_2) \equiv \frac{1}{2l+1} Y_{lm}(n_1) Y^*_{lm}(n_2), \qquad (A15)$$

and

$$f_l := \frac{(-1)^m f^{ll}_{m,-m}}{2l+1}. \qquad (A16)$$

In the case of observable products, we obtain the well-known results Eqs. (36) and (37). For the case of most interest $N > 2$, we need to use iteratively the "Clebsch-Gordan" composition rule,

$$D_{l_1,m_1 m'_1} D_{l_2,m_2 m'_2} = \begin{pmatrix} l_1 & l_2 & L \\ m_1 & m_2 & -M \end{pmatrix} \begin{pmatrix} l_1 & l_2 & L \\ m'_1 & m'_2 & -M' \end{pmatrix}$$
$$\times (2L+1)(-1)^{M+M'} D_{L,MM'}, \qquad (A17)$$

to reduce the product in Eq. (A10) down to a single pair, in which case we can use again Eqs. (A11) and (A12). The result is

$$I^{l_1 \dots l_N}_{m'_1 \dots m'_N, m_1 \dots m_N} \equiv W^{l_1 \dots l_N | L_1 \dots L_{N-3}}_{m'_1 \dots m'_N} W^{l_1 \dots l_N | L_1 \dots L_{N-3}}_{m_1 \dots m_N}, \qquad (A18)$$

where the coefficients on the right-hand side are given in Eq. (42). Going back to Eq. (A9), this finally implies the decomposition

$$f(n_1, \dots, n_N) \equiv f_{l_1 \dots l_N | L_1 \dots L_{N-3}} Y_{l_1 \dots l_N | L_1 \dots L_{N-3}}(n_1, \dots, n_N), \qquad (A19)$$

where

$$f_{l_1 \dots l_N | L_1 \dots L_{N-3}} := W^{l_1 \dots l_N | L_1 \dots L_{N-3}}_{m_1 \dots m_N} f^{l_1 \dots l_N}_{m_1 \dots m_N} \qquad (A20)$$

and

$$Y_{l_1 \dots l_N | L_1 \dots L_{N-3}}(n_1, \dots, n_N) := W^{l_1 \dots l_N | L_1 \dots L_{N-3}}_{m_1 \dots m_N} \prod_{k=1}^N Y_{l_k m_k}(n_k). \qquad (A21)$$

The latter form a basis for SO(3)-invariant functions on S_2^N. Indeed, we have just shown that they generate that space, so we must still show that they are orthogonal under the natural scalar product. Using the identity Eq. (B5), where

$$\{l_1 \ l_2 \ l_3\} := \begin{cases} 1 & \text{if } l_1 \in \{|l_2 - l_3|, \dots, l_2 + l_3\} \\ 0 & \text{otherwise} \end{cases}, \qquad (A22)$$

one obtains

$$W^{l_1 \dots l_N | L_1 \dots L_{N-3}}_{m_1 \dots m_N} W^{l_1 \dots l_N | L'_1 \dots L'_{N-3}}_{m_1 \dots m_N}$$
$$\equiv \{l_1 \dots l_N | L_1 \dots L_{N-3}\} \prod_{k=1}^{N-3} \delta^{L_k L'_k}, \qquad (A23)$$

where we have defined

$$\{l_1 \dots l_N | L_1 \dots L_{N-3}\} := \{l_1 \ l_2 \ L_1\} \left(\prod_{k=1}^{N-4} \{L_k \ l_{k+2} \ L_{k+1}\} \right)$$
$$\times \{L_{N-3} \ l_{N-1} \ l_N\}, \qquad (A24)$$

and therefore the orthonormality relation

$$\int \left(\prod_{k=1}^N \frac{d\Omega_k}{4\pi} \right) Y_{l_1 \dots l_N | L_1 \dots L_{N-3}}(n_1, \dots, n_N)$$
$$\times Y^*_{l'_1 \dots l'_N | L'_1 \dots L'_{N-3}}(n_1, \dots, n_N)$$
$$\equiv \{l_1 \dots l_N | L_1 \dots L_{N-3}\} \prod_{k=1}^N \delta^{l_k l'_k} \prod_{k=1}^{N-3} \delta^{L_k L'_k}. \quad (A25)$$

Finally, one can use this to invert Eq. (A19),

$$f_{l_1 \dots l_N | L_1 \dots L_{N-3}} \equiv \int \left(\prod_{k=1}^N \frac{d\Omega_k}{4\pi} \right) Y^*_{l_1 \dots l_N | L_1 \dots L_{N-3}}(n_1, \dots, n_N)$$
$$\times f(n_1, \dots, n_N). \qquad (A26)$$

APPENDIX B: SOME IDENTITIES FOR WIGNER $3 - j$ SYMBOLS

Here we display all the identities satisfied by the Wigner $3 - j$ symbols that are used in this paper.

(1) Special cases:

$$\begin{pmatrix} l_1 & 0 & 0 \\ m_1 & 0 & 0 \end{pmatrix} \equiv \delta^{l_1}_0 \delta^0_{m_1},$$

$$\begin{pmatrix} l_1 & l_2 & 0 \\ m_1 & m_2 & 0 \end{pmatrix} \equiv \frac{(-1)^{l_1+m_1}}{\sqrt{2l_1+1}} \delta^{l_1 l_2} \delta_{m_1,-m_2}. \qquad (B1)$$

(2) Column exchange symmetry:
even permutations

$$\begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \equiv \begin{pmatrix} l_2 & l_3 & l_1 \\ m_2 & m_3 & m_1 \end{pmatrix} \equiv \begin{pmatrix} l_3 & l_1 & l_2 \\ m_3 & m_1 & m_2 \end{pmatrix}, \qquad (B2)$$

odd permutations

$$\begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \equiv (-1)^{l_1+l_2+l_3} \begin{pmatrix} l_2 & l_1 & l_3 \\ m_2 & m_1 & m_3 \end{pmatrix}$$
$$\equiv (-1)^{l_1+l_2+l_3} \begin{pmatrix} l_1 & l_3 & l_2 \\ m_1 & m_3 & m_2 \end{pmatrix}, \qquad (B3)$$

(3) m-sign flip symmetry

$$\begin{pmatrix} l_1 & l_2 & l_3 \\ -m_1 & -m_2 & -m_3 \end{pmatrix} \equiv (-1)^{l_1+l_2+l_3} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix}, \qquad (B4)$$

(4) Sum identities

$$\begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \begin{pmatrix} l_1 & l_2 & l'_3 \\ m_1 & m_2 & m'_3 \end{pmatrix}$$
$$\equiv \frac{1}{2l_3+1} \{l_1 \ l_2 \ l_3\} \delta_{l_3 l'_3} \delta_{m_3 m'_3}, \qquad (B5)$$

$$(-1)^m \begin{pmatrix} L & l & l \\ M & m & -m \end{pmatrix} \equiv \sqrt{2l+1} (-1)^l \delta^L_0 \delta^0_M. \qquad (B6)$$

[1] M. Levi, C. Bebek, T. Beers, R. Blum, R. Cahn, D. Eisenstein *et al.*, The DESI experiment, a whitepaper for snowmass, arXiv:1308.0847.

[2] P. A. Abell *et al.* (LSST Science Collaborations), LSST Science Book, Version 2.0, arXiv:0912.0201 (2009).

[3] J. Green, P. Schechter, C. Baltay, R. Bean, D. Bennett, R. Brown *et al.*, Wide-field infrared survey telescope (WFIRST) final report, arXiv:1208.4012.

[4] P. E. Dewdney, P. J. Hall, R. T. Schilizzi, and T. J. L. W. Lazio, The square kilometre array, IEEE Proc. **97**, 1482 (2009).

[5] R. Laureijs, J. Amiaux, S. Arduini, J. Auguères, J. Brinchmann, R. Cole *et al.*, Euclid definition study report, arXiv:1110.3193.

[6] A. Challinor and A. Lewis, Lensed CMB power spectra from all-sky correlation functions, Phys. Rev. D **71**, 103010 (2005).

[7] A. Lewis and A. Challinor, Weak gravitational lensing of the CMB, Phys. Rep. **429**, 1 (2006).

[8] D. Hanson, A. Challinor, and A. Lewis, Weak lensing of the CMB, Gen. Relativ. Gravit. **42**, 2197 (2010).

[9] A. Lewis, A. Challinor, and D. Hanson, The shape of the CMB lensing bispectrum, J. Cosmol. Astropart. Phys. 03 (2011) 018.

[10] G. W. Pettinari, C. Fidler, R. Crittenden, K. Koyama, A. Lewis, and D. Wands, Impact of polarization on the intrinsic cosmic microwave background bispectrum, Phys. Rev. D **90**, 103010 (2014).

[11] C. Bonvin, C. Clarkson, R. Durrer, R. Maartens, and O. Umeh, Do we care about the distance to the CMB? Clarifying the impact of second-order lensing, J. Cosmol. Astropart. Phys. 06 (2015) 050.

[12] G. Marozzi, G. Fanizza, E. Di Dio, and R. Durrer, CMB-lensing beyond the Born approximation, J. Cosmol. Astropart. Phys. 09 (2016) 028.

[13] G. Marozzi, G. Fanizza, E. Di Dio, and R. Durrer, Impact of Next-to-Leading Order Contributions to Cosmic Microwave Background Lensing, Phys. Rev. Lett. **118**, 211301 (2017).

[14] G. Marozzi, G. Fanizza, E. Di Dio, and R. Durrer, CMB-lensing beyond the leading order: Temperature and polarization anisotropies, Phys. Rev. D **98**, 023535 (2018).

[15] G. Pratten and A. Lewis, Impact of post-Born lensing on the CMB, J. Cosmol. Astropart. Phys. 08 (2016) 047.

[16] A. Lewis, A. Hall, and A. Challinor, Emission-angle and polarization-rotation effects in the lensed CMB, J. Cosmol. Astropart. Phys. 08 (2017) 023.

[17] E. Di Dio, R. Durrer, G. Fanizza, and G. Marozzi, Rotation of the CMB polarization by foreground lensing, Phys. Rev. D **100**, 043508 (2019).

[18] E. Di Dio, R. Durrer, G. Marozzi, and F. Montanari, Galaxy number counts to second order and their bispectrum, J. Cosmol. Astropart. Phys. 12 (2014) 017.

[19] D. Bertacca, R. Maartens, and C. Clarkson, Observed galaxy number counts on the lightcone up to second order: I. Main result, J. Cosmol. Astropart. Phys. 09 (2014) 037.

[20] D. Bertacca, R. Maartens, and C. Clarkson, Observed galaxy number counts on the lightcone up to second order: II. Derivation, J. Cosmol. Astropart. Phys. 11 (2014) 013.

[21] E. Di Dio, R. Durrer, G. Marozzi, and F. Montanari, The bispectrum of relativistic galaxy number counts, J. Cosmol. Astropart. Phys. 01 (2016) 016.

[22] J. T. Nielsen, and R. Durrer, Higher order relativistic galaxy number counts: Dominating terms, J. Cosmol. Astropart. Phys. 03 (2017) 010.

[23] E. Di Dio, H. Perrier, R. Durrer, G. Marozzi, A. Moradinezhad Dizgah, J. Norea *et al.*, Non-Gaussianities due to relativistic corrections to the observed galaxy bispectrum, J. Cosmol. Astropart. Phys. 03 (2017) 006.

[24] O. Umeh, S. Jolicoeur, R. Maartens, and C. Clarkson, A general relativistic signature in the galaxy bispectrum: The local effects of observing on the lightcone, J. Cosmol. Astropart. Phys. 03 (2017) 034.

[25] S. Jolicoeur, O. Umeh, R. Maartens,, and C. Clarkson, Imprints of local lightcone projection effects on the galaxy bispectrum. Part II, J. Cosmol. Astropart. Phys. 09 (2017) 040.

[26] S. Jolicoeur, O. Umeh, R. Maartens,, and C. Clarkson, Imprints of local lightcone projection effects on the galaxy bispectrum. Part III. Relativistic corrections from nonlinear dynamical evolution on large-scales, J. Cosmol. Astropart. Phys. 03 (2018) 036.

[27] S. Jolicoeur, A. Allahyari, C. Clarkson, J. Larena, O. Umeh, and R. Maartens, Imprints of local lightcone projection effects on the galaxy bispectrum IV: Second-order vector and tensor contributions, J. Cosmol. Astropart. Phys. 03 (2019) 004.

[28] K. Koyama, O. Umeh, R. Maartens, and D. Bertacca, The observed galaxy bispectrum from single-field inflation in the squeezed limit, J. Cosmol. Astropart. Phys. 07 (2018) 050.

[29] E. Di Dio, R. Durrer, R. Maartens, F. Montanari, and O. Umeh, The full-sky angular bispectrum in redshift space, J. Cosmol. Astropart. Phys. 04 (2019) 053.

[30] C. Clarkson, E. M. de Weerd, S. Jolicoeur, R. Maartens, and O. Umeh, The dipole of the galaxy bispectrum, Mon. Not. R. Astron. Soc. **486**, L101 (2019).

[31] M. Jalilvand, B. Ghosh, E. Majerotto, B. Bose, R. Durrer, and M. Kunz, Non-linear contributions to angular power spectra, Phys. Rev. D **101**, 043530 (2020).

[32] J. L. Fuentes, J. C. Hidalgo, and K. A. Malik, Galaxy number counts at second order: An independent approach, arXiv:1908.08400.

[33] F. Bernardeau, C. Bonvin, and F. Vernizzi, Full-sky lensing shear at second order, Phys. Rev. D **81**, 083002 (2010).

[34] F. Bernardeau, C. Bonvin, N. Van de Rijt, and F. Vernizzi, Cosmic shear bispectrum from second-order perturbations in general relativity, Phys. Rev. D **86**, 023001 (2012).

[35] O. Umeh, C. Clarkson, and R. Maartens, Nonlinear relativistic corrections to cosmological distances, redshift, and gravitational lensing magnification: I. Key results, Classical Quantum Grav. **31**, 202001 (2014).

[36] O. Umeh, C. Clarkson, and R. Maartens, Nonlinear relativistic corrections to cosmological distances, redshift and gravitational lensing magnification. II. Derivation, Classical Quantum Gravity **31**, 205001 (2014).

[37] S. Andrianomena, C. Clarkson, P. Patel, O. Umeh, and J.-P. Uzan, Non-linear relativistic contributions to the cosmological weak-lensing convergence, J. Cosmol. Astropart. Phys. 06 (2014) 023.

[38] G. Marozzi, The luminosity distance–redshift relation up to second order in the Poisson gauge with anisotropic stress, Classical Quantum Gravity **32**, 045004 (2015).

[39] C. Bonvin, C. Clarkson, R. Durrer, R. Maartens, and O. Umeh, Cosmological ensemble and directional averages of observables, J. Cosmol. Astropart. Phys. 07 (2015) 040.

[40] G. Fanizza, J. Yoo, and S. G. Biern, Non-linear general relativistic effects in the observed redshift, J. Cosmol. Astropart. Phys. 09 (2018) 037.

[41] H. A. Gressel, C. Bonvin, M. Bruni, and D. Bacon, Full-sky weak lensing: A nonlinear post-Friedmann treatment, J. Cosmol. Astropart. Phys. 05 (2019) 045.

[42] W. Hu, Angular trispectrum of the cosmic microwave background, Phys. Rev. D **64**, 083005 (2001).

[43] J. R. Fergusson and E. P. S. Shellard, Primordial non-Gaussianity and the CMB bispectrum, Phys. Rev. D **76**, 083523 (2007).

[44] K. M. Smith, L. Senatore, and M. Zaldarriaga, Optimal limits on f_{NL}^{local} from WMAP 5-year data, J. Cosmol. Astropart. Phys. 09 (2009) 006.

[45] J. Smidt, A. Amblard, P. Serra, and A. Cooray, Measurement of primordial non-Gaussianity using the WMAP 5-year temperature skewness power spectrum, Phys. Rev. D **80**, 123005 (2009).

[46] L. Senatore, K. M. Smith, and M. Zaldarriaga, Non-Gaussianities in single field inflation and their optimal limits from the WMAP 5-year data, J. Cosmol. Astropart. Phys. 01 (2010) 028.

[47] C. Feng, A. Cooray, J. Smidt, J. O'Bryan, B. Keating, and D. Regan, Planck trispectrum constraints on primordial non-Gaussianity at cubic order, Phys. Rev. D **92**, 043509 (2015).

[48] X. Chen, M.-X. Huang, S. Kachru, and G. Shiu, Observational signatures and non-Gaussianities of general single field inflation, J. Cosmol. Astropart. Phys. 01 (2007) 002.

[49] T. Suyama and M. Yamaguchi, Non-Gaussianity in the modulated reheating scenario, Phys. Rev. D **77**, 023505 (2008).

[50] X. Chen, B. Hu, M.-X. Huang, G. Shiu, and Y. Wang, Large primordial trispectra in general single field inflation, J. Cosmol. Astropart. Phys. 08 (2009) 008.

[51] N. S. Sugiyama, E. Komatsu, and T. Futamase, Non-Gaussianity Consistency Relation for Multifield Inflation, Phys. Rev. Lett. **106**, 251301 (2011).

[52] J. Yoo, A. L. Fitzpatrick, and M. Zaldarriaga, A new perspective on galaxy clustering as a cosmological probe: General relativistic effects, Phys. Rev. D **80**, 083514 (2009).

[53] J. Yoo, General relativistic description of the observed galaxy power spectrum: Do we understand what we measure? Phys. Rev. D **82**, 083508 (2010).

[54] J. Yoo, Relativistic effect in galaxy clustering, Classical Quantum Gravity **31**, 234001 (2014).

[55] J. Yoo and M. Zaldarriaga, Beyond the linear-order relativistic effect in galaxy clustering: Second-order gauge-invariant formalism, Phys. Rev. D **90**, 023513 (2014).

[56] S. G. Biern and J. Yoo, Gauge-invariance and infrared divergences in the luminosity distance, J. Cosmol. Astropart. Phys. 04 (2017) 045.

[57] S. G. Biern and J. Yoo, Correlation function of the luminosity distances, J. Cosmol. Astropart. Phys. 09 (2017) 026.

[58] J. Yoo and R. Durrer, Gauge-transformation properties of cosmological observables and its application to the light-cone average, J. Cosmol. Astropart. Phys. 09 (2017) 016.

[59] F. Scaccabarozzi and J. Yoo, Light-cone observables and gauge-invariance in the geodesic light-cone formalism, J. Cosmol. Astropart. Phys. 06 (2017) 007.

[60] F. Scaccabarozzi, J. Yoo, and S. G. Biern, Galaxy two-point correlation function in general relativity, J. Cosmol. Astropart. Phys. 10 (2018) 024.

[61] J. Yoo, N. Grimm, E. Mitsou, A. Amara, and A. Refregier, Gauge-invariant formalism of cosmological weak lensing, J. Cosmol. Astropart. Phys. 04 (2018) 029.

[62] N. Grimm and J. Yoo, Jacobi Mapping approach for a precise cosmological weak lensing formalism, J. Cosmol. Astropart. Phys. 07 (2018) 067.

[63] N. Grimm, F. Scaccabarozzi, J. Yoo, S. G. Biern, and J.-O. Gong, Galaxy power spectrum in general relativity, arXiv:2005.06484.

[64] E. Mitsou and J. Yoo, Tetrad formalism for exact cosmological observables, in Springer Briefs in Physics (in press), arXiv:1908.10757.

[65] T. Okamoto and W. Hu, Angular trispectra of CMB temperature and polarization, Phys. Rev. D **66**, 063008 (2002).

[66] D. M. Regan, E. P. S. Shellard, and J. R. Fergusson, General CMB and primordial trispectrum estimation, Phys. Rev. D **82**, 023520 (2010).

[67] M. Kamionkowski, T. L. Smith, and A. Heavens, CMB bispectrum, trispectrum, non-Gaussianity, and the Cramer-Rao bound, Phys. Rev. D **83**, 023007 (2011).

[68] T. L. Smith and M. Kamionkowski, Probability distribution for non-Gaussianity estimators constructed from the CMB trispectrum, Phys. Rev. D **86**, 063009 (2012).

[69] L. R. Abramo and T. S. Pereira, Testing Gaussianity, homogeneity, and isotropy with the cosmic microwave background, Adv. Astron. **2010**, 378203 (2010).

[70] F. Lacasa, Non-Gaussianity and extragalactic foregrounds to the cosmic microwave background, arXiv:1406.0441.

[71] E. Barausse, S. Matarrese, and A. Riotto, The Effect of inhomogeneities on the luminosity distance-redshift relation: Is dark energy necessary in a perturbed Universe? Phys. Rev. D **71**, 063537 (2005).

[72] W. Feller, *Introduction to Probability Theory and Its Application* (John Wiley & Sons, New York, 1968), Vol. 1.

[73] R. Durrer, *The Cosmic Microwave Background* (Cambridge University Press, Cambridge, 2008).

[74] J. C. Mather *et al.*, Measurement of the cosmic microwave background spectrum by the COBE FIRAS instrument, Astrophys. J. **420**, 439 (1994).

[75] J. Yoo, E. Mitsou, N. Grimm, R. Durrer, and A. Refregier, Cosmological information contents on the light-cone, J. Cosmol. Astropart. Phys. 12 (2019) 015.

[76] J. Yoo, E. Mitsou, Y. Dirian, and R. Durrer, Background photon temperature $\bar{T}$: A new cosmological Parameter? Phys. Rev. D **100**, 063510 (2019).