

Optimal estimation of the binned mask-free power spectrum, bispectrum, and trispectrum on the full sky: Tensor edition

Oliver H. E. Philcox^{*}

*Center for Theoretical Physics, Department of Physics, Columbia University,
New York, New York 10027, USA
and Simons Society of Fellows, Simons Foundation, New York, New York 10010, USA*

 (Received 8 June 2023; accepted 9 August 2023; published 8 September 2023)

We derive optimal estimators for the binned two-, three-, and four-point correlators of statistically isotropic tensor fields defined on the sphere, in the presence of arbitrary beams, inpainting, and masking. This is a conceptually straightforward extension of the associated scalar-field estimators [O. H. E. Philcox, *Phys. Rev. D* **107**, 123516 (2023).], but upgraded to include spin-2 fields such as cosmic microwave background polarization and galaxy shear, and parity-violating physics in all correlators. All estimators can be realized using spin-weighted spherical harmonic transforms and Monte Carlo summation and are implemented in the public code PolyBin, with computation scaling, at most, with the total number of bins. We perform a suite of validation tests verifying that the estimators are unbiased and, in limiting regimes, minimum variance. These facilitate general binned analyses of higher-point functions, and allow constraints to be placed on various phenomena, such as nonseparable inflationary physics (novelly including polarized trispectra), nonlinear evolution in the late Universe, and cosmic parity violation.

DOI: [10.1103/PhysRevD.108.063506](https://doi.org/10.1103/PhysRevD.108.063506)

I. INTRODUCTION

Cosmology abounds with scalar fields: the density of galaxies, the temperature of the cosmic microwave background (CMB), and the ionization fraction of the Universe, are a few such examples. Although these fields exist in three-dimensions, geometry forces us to observe them only on the surface of a two-sphere, whose radius corresponds to their distance, or, equivalently, age. Whilst the analysis of such observables has yielded significant information about the Universe's structure, composition and evolution, there is more information to be wrought if one additionally considers tensor fields. Such quantities (which are again observed on the two-sphere) contain also directional information, with canonical examples being the polarization of the CMB and the elliptical shapes of galaxies [1–5]. Many tensor fields exhibit symmetry under 90° rotations; these are known as spin-2 fields [5], and their study has yielded important constraints on the dynamics of inflation (through E and B modes in the CMB) and the growth of structure (through gravitational lensing) e.g., [6–12]. Moreover, there exist cosmological fields with other spins, including the spin-1 orientation field of galaxies e.g., [13], which shows 180° rotational symmetries, and a plurality of examples beyond cosmology, arising in disciplines such as oceanography and atmospheric physics.

Information can be extracted from stochastic scalar and tensor fields such as the above through careful measurement and modeling of their statistical properties. This is principally performed using N -point functions, which encapsulate the correlations between values of the field at N points on the sphere [3,14]. On large scales, the standard cosmological model asserts that the asymptotic distribution of these fields is Gaussian, and thus fully described by its two-point function or power spectrum [8]. This provides a direct window to study the physics of the early Universe, with higher-point functions, such as the bispectrum and trispectrum allowing for tests of our assumptions, as well as parametrization of gravitationally induced nonlinearity. For this reason, it has been commonplace for CMB experiments to study three- and four-point functions of temperature fluctuations, allowing for a direct probe of the physics of inflation, and, on small-scales, gravitational lensing e.g., [6,7,15–21]. Many of these include polarization data, though this has not yet been considered for large-scale trispectra. Similarly, the distribution of matter in the late Universe has begun to be probed using non-Gaussian statistics, both in photometric and spectroscopic surveys e.g., [4,22–37], though the latter usually works in deprojected three-dimensional space.

To perform accurate analyses of cosmological fields on the two-sphere, we require robust estimators for the various statistics. In general, this is nontrivial, since the fields in question are rarely observed uniformly or even across the entire sky. Furthermore, experimental noise and telescope

^{*}ohp2@cantab.ac.uk

beams impact different scales by different amounts, and there are often gaps in the survey observations due to the observer effects such as Milky Way extinction and bright stars. This imprints a complex “window function” (or “mask”) onto the observations, leading to leakage between different scales and polarization components. Whilst robust methods exist for accounting for such effects in the power spectrum [38], there has been comparably little discussion for higher-order statistics, particularly when tensors are involved (see [17–19,39–49]). Much of this is pragmatic; many of the current analyses of CMB bispectra and trispectra look only for the amplitude of a single inflationary component (parametrized by some “nonlinear” amplitude, such as f_{NL} or g_{NL}), whose biases can be well established using Monte Carlo simulations [7,43]. If we wish to study more complex inflationary models (which are not separable in harmonic space e.g., [50]) or to include gravitational contributions to the model, it is advantageous to instead measure the full statistic in some set of bins (rather than just a characteristic nonlinear amplitude, such as f_{NL}) cf., [41,51]. In this case, it is important to account for the various observational effects listed above, such that the output spectra can be reliably compared to models.

In [52] we introduced a formalism for efficient and accurate computation of N -point functions sourced by scalar fields on the two-sphere. Starting from the theoretical likelihood for the observed field, this derived optimal estimators for scalar polyspectra, which, by definition, had minimum variance in the Gaussian limit (following previous work including [47,48,53–59]). A slight generalization of these allowed for polyspectra to be efficiently computed using arbitrary weighting schemes (allowing for any linear masking, inpainting, deprojection, etc., utilizing iterated spherical harmonic transforms [52], which become highly efficient on large scales. An important feature of such estimates is that they are unbiased, i.e., they deconvolve all mask- and beam-induced effects such that the output spectra can be directly compared to theoretical models. The physical utility of these has already been demonstrated; in [60], we performed the first measurement of the parity-odd component of the *Planck* CMB trispectrum, placing strong constraints on parity-breaking models of the early Universe, and refuting (scalar) inflationary interpretation of the recent claims of parity-violation in galaxy survey data [29,34,35].

This work considers the analogous problem applied to tensor fields. In particular, we construct estimators for the power spectrum, bispectrum, and trispectrum of correlated spin-0 and spin-2 fields on the two-sphere, in the presence of masking, inpainting, and arbitrary linear weighting schemes. These are significantly more nuanced than their scalar equivalents, due to the introduction of additional variables (E and B fields), and the corresponding proliferation of output spectra. However, these can still be computed efficiently on large-scales, with computation time scaling at most linearly with the number of bins. We

additionally allow computation of spectra sourced by both parity-conserving and parity-breaking physics; the first time, to our knowledge, that this has been performed for higher-point tensor observables. The latter could occur due to a range of new physics in inflation (such as couplings to vector fields, Chern-Simons interactions, loops, and beyond [61–66]), or via some form of chiral gravity modifying the observables at late times [29]. In all cases, our estimators are unbiased (such that they can be straightforwardly compared to theoretical models), and can be implemented using spin-weighted spherical harmonic transforms. We additionally provide a general purpose implementation in the PolyBin code [67].¹ This should allow for a range of novel theoretical tests, both for CMB and weak lensing analyses, such as a first measurement of the large-scale polarized trispectrum.

The remainder of this paper is as follows. In Sec. II we set out our conventions for spinning fields on the two-sphere and the various correlators, before discussing the general form of N -point function estimators in Sec. III. Sections IV–VI discuss practical computation of the power spectrum, bispectrum, and trispectrum respectively, before the estimators are validated with a suite of tests in Sec. VII. Finally, we conclude in Sec. VIII. We caution that the remainder of this paper is (quite necessarily) mathematically dense. Much of the material builds upon the spin-0 work presented in [52], but we recapitulate the relevant material here for completeness.

II. TENSORS AND CORRELATORS

A. Definition

A spin- s field defined on the two-sphere, ${}_s a(\hat{\mathbf{n}})$, can be expanded in spin-weighted spherical harmonics as

$${}_s a(\hat{\mathbf{n}}) = \sum_{\ell m} {}_s a_{\ell m} Y_{\ell m}(\hat{\mathbf{n}}), \quad {}_s a_{\ell m} = \int d\hat{\mathbf{n}} [{}_s Y_{\ell m}(\hat{\mathbf{n}})]^* {}_s a(\hat{\mathbf{n}}). \quad (1)$$

Assuming the underlying fields to be real with $[{}_s a(\hat{\mathbf{n}})]^* = -{}_s a(\hat{\mathbf{n}})$, the spin-components satisfy

$$[{}_s a_{\ell m}]^* = (-1)^{s+m} {}_{-s} a_{\ell(-m)}, \quad \mathbb{P}[{}_s a_{\ell m}] = (-1)^\ell {}_{-s} a_{\ell m}, \quad (2)$$

where \mathbb{P} is the parity operator. For spin- ± 2 fields, it is conventional to expand in E and B modes:

$${}_{\pm 2} a_{\ell m} = -[a_{\ell m}^E \pm i a_{\ell m}^B], \quad (3)$$

where a^E and a^B give parity-even and parity-odd contributions respectively (both of which are real in map space) and the negative sign appears by convention. These satisfy

¹Available at [GitHub.com/OliverPhilcox/PolyBin](https://github.com/OliverPhilcox/PolyBin).

$$\left[a_{\ell m}^{E,B} \right]^* = (-1)^m a_{\ell(-m)}^{E,B}, \quad \mathbb{P} \left[a_{\ell m}^{E,B} \right] = p_{E,B} (-1)^\ell a_{\ell m}^{E,B}, \quad (4)$$

where $p_E = 1$ and $p_B = -1$ specifies the parity. A similar decomposition is possible for spin-1 fields, in terms of gradient (parity-odd) and curl (parity-even) contributions.

For the cosmic microwave background, there are three fields of interest; the spin-0 temperature, ${}_0a = T$, and the spin- ± 2 polarization fields ${}_{\pm 2}a = Q \pm iU$. These are related to T , E , and B fields via a linear transformation

$$\begin{pmatrix} {}_0a_{\ell m} \\ -{}_2a_{\ell m} \\ +{}_2a_{\ell m} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & -i \\ 0 & -1 & i \end{pmatrix} \begin{pmatrix} a_{\ell m}^T \\ a_{\ell m}^E \\ a_{\ell m}^B \end{pmatrix} \\ \Rightarrow {}_s a_{\ell m} \equiv \sum_X {}_s \mathcal{R}_X a_{\ell m}^X, \quad (5)$$

where $X \in \{T, E, B\}$, and the transformation matrix (which is not a rotation) will be of use later. An analogous form can be written for galaxy positions and shapes, whence ${}_0a = \delta_g$ (the galaxy density), and ${}_{\pm 2}a = \gamma_1 \pm i\gamma_2$ (the galaxy shear).

B. Power spectra

1. Definition

The correlators of spin-2 fields are fundamentally described by the power spectra of T , E , B ,

$$\langle a_{\ell_1 m_1}^X a_{\ell_2 m_2}^{Y*} \rangle = C_{\ell_1 \ell_2}^{XY} \delta_{\ell_1 \ell_2}^K \delta_{m_1 m_2}^K \quad X, Y \in \{T, E, B\}, \quad (6)$$

assuming rotation and translation invariance and noting that autospectra include noise. We additionally ignore any spatially-varying window function, but will return to this later. In terms of the spin fields in harmonic space, we have the ideal correlators

$$\langle {}_{s_1} a_{\ell_1 m_1} [{}_{s_2} a_{\ell_2 m_2}]^* \rangle = {}_{s_1 s_2} C_{\ell_1 \ell_2} \delta_{\ell_1 \ell_2}^K \delta_{m_1 m_2}^K, \quad (7)$$

where the (Hermitian) ${}_{ss}C_\ell$ matrix of correlators can be written in terms of T , E , B as

$${}_{ss}C_\ell = \sum_{XY} ({}_s \mathcal{R}_X) C_\ell^{XY} ({}_{s'} \mathcal{R}_Y)^\dagger = \begin{pmatrix} C_\ell^{TT} & -(C_\ell^{TE} - iC_\ell^{TB}) & -(C_\ell^{TE} + iC_\ell^{TB}) \\ -(C_\ell^{TE} + iC_\ell^{TB}) & C_\ell^{EE} + C_\ell^{BB} & C_\ell^{EE} - C_\ell^{BB} + 2iC_\ell^{EB} \\ -(C_\ell^{TE} - iC_\ell^{TB}) & C_\ell^{EE} - C_\ell^{BB} - 2iC_\ell^{EB} & C_\ell^{EE} + C_\ell^{BB} \end{pmatrix}, \quad (8)$$

where we note that $C_\ell^{XY} = C_\ell^{YX}$ are real. The full set of nontrivial correlators involve the following fields

$$U_2 \equiv \{TT, TE, EE, BB\} + \{TB, EB\}, \quad (9)$$

separating out parity-even and parity-odd spectra, with the latter contributing to the spin- (s, s') covariance multiplied by i . In general, the two-point covariance is nontrivial, but significantly simplifies if we assume parity-conservation and zero-intrinsic B mode,

$$({}_{ss'}C_\ell)^{\text{ideal}} = \begin{pmatrix} C_\ell^{TT} & -C_\ell^{TE} & -C_\ell^{TE} \\ -C_\ell^{TE} & C_\ell^{EE} & C_\ell^{EE} \\ -C_\ell^{TE} & C_\ell^{EE} & C_\ell^{EE} \end{pmatrix} + \begin{pmatrix} N_\ell^T & 0 & 0 \\ 0 & 2N_\ell^P & 0 \\ 0 & 0 & 2N_\ell^P \end{pmatrix}. \quad (10)$$

We have additionally separated out contributions from noise, assuming the impact on the temperature and E/B fields to be independent and described by spectra N_ℓ^T and N_ℓ^P respectively.

2. Application

To form optimal estimators, we will require both the pixel-space covariance matrix and its inverse. From (7), these are given by

$$\begin{aligned} {}_{s_1 s_2} \mathbf{C}(\hat{\mathbf{n}}_1, \hat{\mathbf{n}}_2) &\equiv \langle {}_{s_1} a(\hat{\mathbf{n}}_1) [{}_{s_2} a(\hat{\mathbf{n}}_2)]^* \rangle \\ &= \sum_{\ell m} {}_{s_1} Y_{\ell m}(\hat{\mathbf{n}}_1) [{}_{s_1 s_2} C_\ell]_{s_2} Y_{\ell m}^*(\hat{\mathbf{n}}_2), \\ {}_{s_1 s_2} \mathbf{C}^{-1}(\hat{\mathbf{n}}_1, \hat{\mathbf{n}}_2) &= \sum_{\ell m} {}_{s_1} Y_{\ell m}(\hat{\mathbf{n}}_1) [{}_{s_1 s_2} (C^{-1})_\ell]_{s_2} Y_{\ell m}^*(\hat{\mathbf{n}}_2), \end{aligned} \quad (11)$$

in terms of spherical harmonics. The latter can be written in terms of the matrix of $T/E/B$ correlators $\mathbb{C}_\ell \equiv \{C_\ell^{XY}\}$

$${}_{s_1 s_2} \mathbf{C}^{-1}(\hat{\mathbf{n}}_1, \hat{\mathbf{n}}_2) = \sum_{\ell m} {}_{s_1} Y_{\ell m}(\hat{\mathbf{n}}_1) {}_{s_1 s_2} [\mathcal{R}^{-\dagger} \mathbb{C}_\ell^{-1} \mathcal{R}^{-1}]_{s_2} Y_{\ell m}^*(\hat{\mathbf{n}}_2). \quad (12)$$

Whilst this may seem unduly complicated, it can be straightforwardly applied to any map in question, noting that \mathcal{R} and \mathbb{C} are just 3×3 matrices. Acting on some map ${}_s a(\hat{\mathbf{n}})$, we find

$$s_1[\mathbf{C}^{-1}x](\hat{\mathbf{n}}_1) = \sum_{\ell m} s_1 Y_{\ell m}(\hat{\mathbf{n}}_1)_{s_1 s_2} [\mathcal{R}^{-\dagger} \mathbf{C}_{\ell}^{-1} \mathcal{R}^{-1}]_{s_2} x_{\ell m}, \quad (13)$$

replacing the $\hat{\mathbf{n}}_2$ integral by a harmonic transform of x . The first \mathcal{R}^{-1} operator transforms $s^x x_{\ell m}$ into the $T/E/B$ representation $x_{\ell m}^X$, which is then Wiener filtered by \mathbf{C}_{ℓ}^{-1} and transformed back to the spin representation by \mathcal{R}^{-1} , before being projected into real space by an inverse harmonic transform. Similarly, the inner product can be written

$$\begin{aligned} x^{\dagger} \mathbf{C}^{-1} y &\equiv \sum_{s_1 s_2} \int d\hat{\mathbf{n}}_1 d\hat{\mathbf{n}}_2 s_1 x^*(\hat{\mathbf{n}}_1)_{s_1 s_2} \mathbf{C}^{-1}(\hat{\mathbf{n}}_1, \hat{\mathbf{n}}_2)_{s_2} y(\hat{\mathbf{n}}_2) \\ &= \sum_{\ell m} \sum_{XY} [x_{\ell m}^X]^* (\mathbf{C}_{\ell})_{XY}^{-1} y_{\ell m}^Y, \end{aligned} \quad (14)$$

which is straightforward to compute.

C. Bispectra

The bispectra of three isotropic and homogeneous harmonic space fields can be written

$$\begin{aligned} \langle a_{\ell_1 m_1}^X a_{\ell_2 m_2}^Y a_{\ell_3 m_3}^Z \rangle &\equiv w_{m_1 m_2 m_3}^{\ell_1 \ell_2 \ell_3} b_{\ell_1 \ell_2 \ell_3}^{XYZ}, \\ \{X, Y, Z\} &\in \{T, E, B\}, \end{aligned} \quad (15)$$

where b is the reduced bispectrum and $|\ell_1 - \ell_2| \leq \ell_3 \leq \ell_1 + \ell_2$. In this work, we adopt the following weights (which differ from those of [52])

$$\begin{aligned} w_{m_1 m_2 m_3}^{\ell_1 \ell_2 \ell_3} &= \sqrt{\frac{(2\ell_1 + 1)(2\ell_2 + 1)(2\ell_3 + 1)}{4\pi}} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \left[\frac{1}{3} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ -1 & -1 & 2 \end{pmatrix} + 2 \text{ cyc.} \right] \\ &= \frac{1}{3} \int d\hat{\mathbf{n}} [{}_{+1}Y_{\ell_1 m_1}(\hat{\mathbf{n}}) {}_{+1}Y_{\ell_2 m_2}(\hat{\mathbf{n}}) {}_{-2}Y_{\ell_3 m_3}(\hat{\mathbf{n}}) + 2 \text{ cyc.}], \end{aligned} \quad (16)$$

expressing in terms of the Gaunt integral in the second line. Usually, one fixes $w_{m_1 m_2 m_3}^{\ell_1 \ell_2 \ell_3}$ to the Gaunt symbol; however, this nulls any contributions from parity-odd trispectra (as discussed in [46]), since it is zero for odd $\ell_1 + \ell_2 + \ell_3$. Here, the only restriction is that $\ell_i \geq 2$, which is usually true (after monopole and dipole subtraction). Furthermore, the bispectrum is symmetric under interchange of any $\{\ell_i, X\}$ pairs; similarly the weights are symmetric under exchange of any pair of $\{\ell_i, m_i\}$. Given these symmetries, the bispectrum is fully specified by the following components:

$$\begin{aligned} U_3 &\equiv \{TTT, TTE, TEE, EEE, TBB, EBB\} \\ &+ \{TTB, TEB, EEB, BBB\}, \end{aligned} \quad (17)$$

where the two groups have $p_{XYZ} = p_X p_Y p_Z = 1$ and -1 respectively. An additional result is that not all ℓ -bins are independent. To fully specify the spectra we require the triplets of (17) with

$$\ell_1 \leq \ell_2 \quad (X = Y), \quad \ell_2 \leq \ell_3 \quad (Y = Z), \quad (18)$$

with no restrictions on ℓ_i (except from the triangle conditions).²

²Our convention differs from that of *Planck* [7], which asserted the $X = Y = Z$ conditions for all correlators, but included also permutations e.g., TET in their analysis. Both approaches are equivalent, though we find it useful to minimize the number of output spectra, particularly for higher-order correlators.

Finally, we note that, due to the symmetries of the weighting, the bispectrum has the following properties under conjugation and parity:

$$\begin{aligned} (b_{\ell_1 \ell_2 \ell_3}^{XYZ})^* &= (-1)^{\ell_1 + \ell_2 + \ell_3} b_{\ell_1 \ell_2 \ell_3}^{XYZ}, \\ \mathbb{P}[b_{\ell_1 \ell_2 \ell_3}^{XYZ}] &= p_{XYZ} (-1)^{\ell_1 + \ell_2 + \ell_3} b_{\ell_1 \ell_2 \ell_3}^{XYZ}, \end{aligned} \quad (19)$$

using (4). As such, configurations of fields with $p_{XYZ} = 1$ (the first group in (17) have real bispectra for even $\ell_1 + \ell_2 + \ell_3$ and imaginary else, whilst the situation reverses for $p_{XYZ} = -1$. Note that rotational invariance forbids scalar physics from generating a parity-odd three-point function, thus TTT etc. do not contain contributions with odd $\ell_1 + \ell_2 + \ell_3$, and cannot be used to probe primordial scalar parity-violation. Parity-even spectra with a single B mode can be generated from scalars however and appear at second order in cosmic perturbation theory [68]. As such, scalar physics will generate only four spectra at leading order (ignoring, e.g., lensing); ($b^{TTT}, b^{TTE}, b^{TEE}, b^{EEE}$).

In our case, we require the bispectrum of the spin- s components directly, since these can be straightforwardly related to the real-space windowed quantities. Noting that $s a_{\ell m} = s \mathcal{R}_X a_{\ell m}^X$, in the notation of (5), we can write

$$\begin{aligned} \langle s_1 a_{\ell_1 m_1} s_2 a_{\ell_2 m_2} s_3 a_{\ell_3 m_3} \rangle &= \sum_{XYZ} w_{m_1 m_2 m_3}^{\ell_1 \ell_2 \ell_3} (s_1 \mathcal{R}_X) (s_2 \mathcal{R}_Y) (s_3 \mathcal{R}_Z) b_{\ell_1 \ell_2 \ell_3}^{XYZ}. \end{aligned} \quad (20)$$

Unlike the trispectrum, we do not require permutations in this definition, since the weights are symmetric under (ℓ, m) exchange.

D. Trispectra

The relation for trispectra follows similarly. In this case, we write the general definition

$$\langle a_{\ell_1 m_1}^X a_{\ell_2 m_2}^Y a_{\ell_3 m_3}^Z a_{\ell_4 m_4}^W \rangle \equiv \sum_{LM} (-1)^M w_{\ell_1 \ell_2 m_1 m_2}^{L(-M)} w_{\ell_3 \ell_4 m_3 m_4}^{LM} t_{\ell_1 \ell_2, \ell_3 \ell_4}^{XY, ZW}(L) + 23 \text{ perms.}, \quad \{X, Y, Z, W\} \in \{T, E, B\}, \quad (21)$$

where the permutations are over the joint set $\{X, \ell, m\}$, and the relevant weights are similar as before

$$\begin{aligned} w_{\ell_1 \ell_2 m_1 m_2}^{LM} &= \sqrt{\frac{(2\ell_1 + 1)(2\ell_2 + 1)(2L + 1)}{4\pi}} \begin{pmatrix} \ell_1 & \ell_2 & L \\ m_1 & m_2 & M \end{pmatrix} \begin{pmatrix} \ell_1 & \ell_2 & L \\ -1 & -1 & 2 \end{pmatrix} \\ &= \int d\hat{\mathbf{n}} {}_{+1}Y_{\ell_1 m_1}(\hat{\mathbf{n}}) {}_{+1}Y_{\ell_2 m_2}(\hat{\mathbf{n}}) {}_{-2}Y_{LM}(\hat{\mathbf{n}}), \end{aligned} \quad (22)$$

requiring $\ell_i \geq 1$ and $L \geq 2$. The trispectrum has the following symmetries:

$$t_{\ell_1 \ell_2, \ell_3 \ell_4}^{XY, ZW}(L) = t_{\ell_2 \ell_1, \ell_3 \ell_4}^{YX, ZW}(L) = t_{\ell_3 \ell_4, \ell_1 \ell_2}^{ZW, XY}(L), \quad (23)$$

with other pairs nontrivially related due to the degeneracy in the definition of tetrahedron diagonals. This additionally requires the triangle conditions $|\ell_1 - \ell_2| \leq L \leq \ell_1 + \ell_2$, $|\ell_3 - \ell_4| \leq L \leq \ell_3 + \ell_4$.

The relevant trispectra are defined by the following 21 quadruplets of fields

$$\begin{aligned} U_4 &\equiv \{TTTT, TTTE, TTEE, TTBB, TETE, TEEE, TEBB, TBTB, TBEB, EEEE, EEBB, EBEB, BBBB\} \\ &+ \{TTTB, TTEB, TETB, TEEB, TBEE, TBBB, EEEB, EBBB\}, \end{aligned} \quad (24)$$

with $p_{XYZW} = 1$ and -1 , respectively. As before, the trispectrum symmetries of (23) imply that we can apply the following restrictions on ℓ_i to yield a complete set of bins

$$\begin{aligned} \ell_1 \leq \ell_2 \quad (X = Y), \quad \ell_3 \leq \ell_4 \quad (Z = W) \\ \ell_1 \leq \ell_3 \quad (X = Z \cup Y = W), \quad \ell_2 \leq \ell_4 \quad \text{if } \ell_1 = \ell_3 \quad (X = Z \cup Y = W). \end{aligned} \quad (25)$$

Here, the trispectrum satisfies

$$(t_{\ell_1 \ell_2, \ell_3 \ell_4}^{XY, ZW}(L))^* = (-1)^{\ell_1 + \ell_2 + \ell_3 + \ell_4} t_{\ell_1 \ell_2, \ell_3 \ell_4}^{XY, ZW}(L), \quad \mathbb{P}[t_{\ell_1 \ell_2, \ell_3 \ell_4}^{XY, ZW}(L)] = p_{XYZW} (-1)^{\ell_1 + \ell_2 + \ell_3 + \ell_4} t_{\ell_1 \ell_2, \ell_3 \ell_4}^{XY, ZW}(L) \quad (26)$$

such that spectra with $p_{XYZW} = 1$ are parity-even for even $\ell_1 + \ell_2 + \ell_3 + \ell_4$ and parity-odd for odd $\ell_1 + \ell_2 + \ell_3 + \ell_4$ whilst this is reversed for $p_{XYZW} = -1$. In linear CMB theory, scalar physics only sources spectra containing T and E modes at leading order, which reduces the number of nontrivial spectra to six (though B -modes appear at second order and above, and through lensing). Finally, in terms of the spin-weighted fields we can write

$$\langle a_{s_1 \ell_1 m_1 s_2} a_{s_2 \ell_2 m_2 s_3} a_{s_3 \ell_3 m_3 s_4} a_{s_4 \ell_4 m_4} \rangle \equiv \sum_{LM} (-1)^M w_{\ell_1 \ell_2 m_1 m_2}^{L(-M)} w_{\ell_3 \ell_4 m_3 m_4}^{LM} (s_1 \mathcal{R}_X)(s_2 \mathcal{R}_Y)(s_3 \mathcal{R}_Z)(s_4 \mathcal{R}_W) t_{\ell_1 \ell_2, \ell_3 \ell_4}^{XY, ZW}(L) + 23 \text{ perms.} \quad (27)$$

E. Binning

To form estimators, we must relate the continuous spectra defined above to their binned forms (which we measure directly). This is somewhat nontrivial in the presence of polarization, since only some of the various correlators are independent, with, for example, the BT power spectrum being fully defined by TB .

For the power spectrum, we begin by expressing C_ℓ^{XY} in terms of the binned quantities of interest; $C^u(b)$, where b

specifies an ℓ -bin and $u \in U_2$ specifies a pair of fields from the set given in (9). For arbitrary XY pairs, we can write

$$C_\ell^{XY} = \sum_{b, u \in U_2} \frac{C^u(b)}{\Delta_2^u} [\delta_K^{u, XY} + \delta_K^{u, YX}] \Theta_\ell(b). \quad (28)$$

where $\Theta_\ell(b)$ is unity if ℓ is in bin b and zero else. Here, the symmetry factor Δ_2^u is two if both components of u are

equal and unity else (i.e., $\Delta_u^2 = 1 + \delta_K^{u_1 u_2}$ for $u \equiv u_1 u_2$). In each case, the sum over pairs and bins picks out the relevant spectrum, for example, C_ℓ^{BT} is mapped to $-C_\ell^{TB}$. Strictly, the above result is valid only in the narrow bin limit, but this makes little difference for the higher-point functions. The

above result can be straightforwardly inserted into (7), and allows extraction of the derivatives of the ${}_s a_{\ell m}$ covariance with respect to the desired bandpowers.

A similar procedure is possible for the bispectrum. In this case, we sum over all triplets of fields $s \in S_3$ (17),

$$b_{\ell_1 \ell_2 \ell_3}^{XYZ} = \frac{1}{2} \sum_{\mathbf{b}, u \in U_3, \chi \in \pm 1} \frac{b_\chi^u(\mathbf{b})}{\Delta_3^u(\mathbf{b})} \left[\delta_K^{u, XYZ} \Theta_{\ell_1}(b_1) \Theta_{\ell_2}(b_2) \Theta_{\ell_3}(b_3) + 5 \text{ perms.} \right] \left[1 + \chi p_u (-1)^{\ell_1 + \ell_2 + \ell_3} \right], \quad (29)$$

where the allowed values of the bins $\mathbf{b} \equiv \{b_1, b_2, b_3\}$ are set by the symmetries of (18) (i.e., $b_1 \leq b_2$ if $u_1 = u_2$ and $b_2 \leq b_3$ if $u_2 = u_3$, for $u \equiv u_1 u_2 u_3$).³ Here the degeneracy factor $\Delta_3^u(\mathbf{b})$ is defined by

$$\Delta_3^u(\mathbf{b}) = \begin{cases} 6 & b_1 = b_2 = b_3 \text{ and } u_1 = u_2 = u_3 \\ 2 & b_1 = b_2 \text{ and } u_1 = u_2 \\ 2 & b_2 = b_3 \text{ and } u_2 = u_3 \\ 1 & \text{else,} \end{cases} \quad (30)$$

where the conditions should be read sequentially. Finally, we have inserted a parity-factor $[1 + \chi p_u (-1)^{\ell_1 + \ell_2 + \ell_3}]/2$;

this allows us to separate parity-even and parity-odd components which source even and odd $\ell_1 + \ell_2 + \ell_3$ (depending on the spectrum of interest). Here, we denote conventional parity-conserving contributions by $b_+^u(\mathbf{b})$ and those sourced by parity-odd effects (e.g., chiral gravitational waves) by $b_-^u(\mathbf{b})$. Note that this is not required for the power spectrum, since spectra with $p_u = -1$ (TB and EB) can only be sourced by parity-breaking physics (and likewise for $p_u = 1$ and parity-conserving physics).

Finally, the binned trispectrum for fields X, Y, Z, W can be written as a sum over all trispectra listed in (24), for bins $\mathbf{b} \equiv \{b_1, b_2, b_3, b_4\}$ in ℓ_i and bin B in L ,

$$\langle a_{\ell_1 m_1}^X a_{\ell_2 m_2}^Y a_{\ell_3 m_3}^Z a_{\ell_4 m_4}^W \rangle = \frac{1}{2} \sum_{LM} (-1)^M w_{\ell_1 \ell_2 m_1 m_2}^{L(-M)} w_{\ell_3 \ell_4 m_3 m_4}^{LM} \sum_{\mathbf{b}, u \in U_4, \chi \in \pm 1} \frac{t_\chi^u(\mathbf{b}, B)}{\Delta_4^u(\mathbf{b})} \Theta_L(B) [1 + \chi p_u (-1)^{\ell_1 + \ell_2 + \ell_3 + \ell_4}] \\ \times \left[\delta_K^{u, XYZW} \Theta_{\ell_1}(b_1) \Theta_{\ell_2}(b_2) \Theta_{\ell_3}(b_3) \Theta_{\ell_4}(b_4) + 7 \text{ perms.} \right] + (2 \leftrightarrow 3) + (2 \leftrightarrow 4), \quad (31)$$

where we include the m variables in this definition (unlike (29), to yield the correct behavior under permutations. This can be straightforwardly related to that of the spin- s fields ${}_s a_{\ell m}$ via (27). Here, the eight permutations in the bracket are over exchanges of the following sets of fields-and-bins, $\{b_i, u_i\}$:

$$\{1234\}, \{1243\}, \{2134\}, \{2143\}, \{3412\}, \{3421\}, \{4312\}, \{4321\}, \quad (32)$$

and the final two terms permute all relevant ℓ_i, m_i, u_i . As before, the sum over bins in (31) is limited to nontrivial bins, as defined by (25). In this case, the degeneracy factor is defined as

$$\Delta_4^u(\mathbf{b}) = \begin{cases} 8 & b_1 = b_2 = b_3 = b_4 \text{ and } u_1 = u_2 = u_3 = u_4 \\ 4 & b_1 = b_2 \text{ and } b_3 = b_4 \text{ and } u_1 = u_2 \text{ and } u_3 = u_4 \\ 2 & b_1 = b_2 \text{ and } u_1 = u_2 \\ 2 & b_3 = b_4 \text{ and } u_3 = u_4 \\ 2 & b_1 = b_3 \text{ and } b_2 = b_4 \text{ and } u_1 = u_3 \text{ and } u_2 = u_4 \\ 1 & \text{else.} \end{cases} \quad (33)$$

Finally, we insert a parity-factor, as for the bispectrum, to separate out contributions sourced by parity-even and parity-odd physics, here denoted by t_+ and t_- .

³In the narrow-bin limit, bins with $b_1 = b_2 = b_3$ have even $\ell_1 + \ell_2 + \ell_3$, thus cannot yield signals for $\chi p_u = -1$. In [52], such bins were dropped; here, we keep them for consistency between parity-even and parity-odd definitions, but note that they will have large variances. A similar conclusion holds for certain trispectrum components.

III. SPECTRAL ESTIMATORS

A. Ideal case

In this section, we derive optimal estimators for weakly non-Gaussian statistics, starting from a suitable expansion of the data likelihood and maximizing with respect to the signal of interest. Such estimators can be robustly defined and implemented, and, in the limit of Gaussian statistics, are minimum variance, i.e., they saturate their Cramér-Rao bound.

We begin with the Edgeworth expansion for the likelihood of data d_i in some set of pixels $\hat{\mathbf{n}}$ and spins s (jointly denoted by the index i). Assuming Einstein summation convention, we can write

$$L[d] \propto \exp \left[-\frac{1}{2} d^{i*} \mathbf{C}_{ij}^{-1} d^j - \frac{1}{2} \text{Tr} \log \mathbf{C} \right] \times \left\{ 1 + \frac{1}{3!} \langle d_i d_j d_k \rangle (\mathcal{H}^{ijk})^* + \frac{1}{4!} \langle d_i d_j d_k d_l \rangle (\mathcal{H}^{ijkl})^* + \dots \right\}. \quad (34)$$

Here $\mathbf{C}_{ij} \equiv \langle d_i d_j^* \rangle$ is the map-space covariance matrix, whose form was given in (11) as a function of two positions and two spins. We have additionally included the three- and four-point expectations $\langle d_i d_j d_k \rangle$ and $\langle d_i d_j d_k \rangle$ which can be similarly computed in terms of the bispectrum and trispectrum of Sec. II. The objects denoted $\mathcal{H}_{ij\dots}$ are *Hermite tensors*, given by

$$\begin{aligned} \mathcal{H}_{ijk} &= h_i h_j h_k - (h_i \langle h_j h_k \rangle + 2 \text{ perms.}) \\ \mathcal{H}_{ijkl} &= h_i h_j h_k h_l - (h_i h_j \langle h_k h_l \rangle + 5 \text{ perms.}) \\ &\quad + (\langle h_i h_j \rangle \langle h_k h_l \rangle + 2 \text{ perms.}), \end{aligned} \quad (35)$$

in terms of the Wiener-filtered data $h_i \equiv [\mathbf{C}^{-1} d]_i$. This expansion is similar to that used in previous works [52,59], but allows for complex data as required for fields with nonzero spin (e.g., $Q \pm iU$ polarization states).

To obtain estimators for the N -point correlators, we maximize the logarithm of (34) with respect to some coefficients, $x_\chi^u(\mathbf{b})$, contained only within the $(N > 2)$ -point function (e.g., the parity even/odd bandpowers in some set of bins and fields). This yields an estimator of the form

$$[\mathcal{F}_N \hat{x}_\chi^u](\mathbf{b}) = \frac{1}{N!} \frac{\partial \langle d_{i_1} \dots d_{i_N} \rangle}{\partial x_\chi^u(\mathbf{b})} (\mathcal{H}^{i_1 \dots i_N})^*, \quad (36)$$

involving a data-independent normalization matrix, defined as

$$\begin{aligned} \mathcal{F}_{N, \chi \chi'}^{uu'}(\mathbf{b}, \mathbf{b}') &= \frac{1}{N!} \left[\frac{\partial \langle d^{i_1} \dots d^{i_N} \rangle^*}{\partial x_\chi^u(\mathbf{b})} \mathbf{C}_{i_1 j_1}^{-1} \right. \\ &\quad \left. \dots \mathbf{C}_{i_N j_N}^{-1} \frac{\partial \langle d^{j_1} \dots d^{j_N} \rangle}{\partial x_{\chi'}^{u'}(\mathbf{b}')} \right]^*. \end{aligned} \quad (37)$$

Assuming weak non-Gaussianity, this is an optimal estimator for $x_\chi^u(\mathbf{b})$, i.e., it is minimum variance. Furthermore, the choice of \mathcal{F}_N ensures that the estimator is unbiased, such that $\mathbb{E}[\hat{x}_\chi^u(\mathbf{b})] = x_\chi^u(\mathbf{b})$ (via the Hermite polynomial orthogonality).⁴ Furthermore, in the Gaussian limit, the covariance matrix of $\hat{x}_\chi^u(\mathbf{b})$ is given by the inverted normalization $\mathcal{F}_{N, \chi \chi'}^{-1, uu'}(\mathbf{b}, \mathbf{b}')$. Finally, denoting the Cholesky factorized Fisher matrix as $\mathcal{F}^{1/2}$, we note that e.g., [55,56]

$$[\mathcal{F}_N^{1/2} \hat{x}_\chi^u](\mathbf{b}) \sim \mathcal{N}(\mathbf{0}, \mathbb{I}), \quad (38)$$

i.e., the rescaled estimator obeys a standardized normal distribution in the limit of weak non-Gaussianity. This allows for efficient data compression, as in [60].

B. Realistic case

In practice, the estimators we will discuss in this work are a little more nuanced than the above. First, the data is often masked or windowed, such that the true map, d , is related to the observed map, \tilde{d} , by ${}_s \tilde{d}(\hat{\mathbf{n}}) = W(\hat{\mathbf{n}}) {}_s d(\hat{\mathbf{n}})$, where W is some pixel-space mask which we will generally take to be spin-independent and real.⁵ As such, we should formulate the estimators of (36) and (37) in terms of masked fields \tilde{d} , and their respective correlators. Since the mask is multiplicative, this is straightforward to implement in practice.

Secondly, the above estimators require Wiener-filtered data, $h \equiv \mathbf{C}^{-1} \tilde{d}$, where $\mathbf{C} = \langle \tilde{d} \tilde{d}^\dagger \rangle$. The full covariance matrix is high dimensional (containing $9N_{\text{pix}}^2$ elements for spins $0, \pm 2$) and thus hard to compute and harder to invert, particularly in the presence of translation-dependent noise and holes in the window (such as a Galactic sky mask). A more practical estimator is wrought by abandoning the exact Wiener filter and instead applying some custom weighting \mathbf{S}^{-1} , with h redefined to $h \equiv \mathbf{S}^{-1} \tilde{d}$. The weighting is not required to be invertible (i.e., \mathbf{S} does not have to exist), and can be chosen to allow for simple implementation. In CMB analyses, a typical approach would be to choose the linear operator \mathbf{S}^{-1} such that it (a) (noninvertibly) projects out undesired areas of the map (from e.g., the Milky Way and point sources), and (b) filters the data in harmonic space by the ideal power spectrum (including noise). Assuming a diagonal-in- $\hat{\mathbf{n}}$ projection operator Π and a diagonal-in- ℓ filtering by the idealized $T/E/B$ -space covariance \mathcal{C}_ℓ , this is applied to a map x as

⁴We have implicitly assumed that the N -point correlators are fully described by the binned polyspectra. Any contributions from nonprimordial sources, e.g., noise, are thus included within the binned correlators.

⁵One could straightforwardly use a different mask for each spin and similarly deconvolve its effect from the output correlators.

$$\begin{aligned}
 {}_s x(\hat{\mathbf{n}}) &\rightarrow \Pi(\hat{\mathbf{n}}) {}_s x(\hat{\mathbf{n}}) \rightarrow {}_s [\Pi x]_{\ell m} \rightarrow [\mathbf{C}_{\ell}^{-1} \mathcal{R}^{-1} [\Pi x]]_{\ell m}^X \\
 &\rightarrow {}_s [\mathcal{R}^{-\dagger} \mathbf{C}_{\ell}^{-1} \mathcal{R}^{-1} [\Pi x]](\hat{\mathbf{n}}), \quad (39)
 \end{aligned}$$

which requires only two harmonic transforms to implement.

For a general filter \mathbf{S}^{-1} and mask, the binned-polyspectrum estimator becomes

$$\begin{aligned}
 [\mathcal{F}_N \hat{x}]_{\chi}^u(\mathbf{b}) &= \frac{1}{N!} \frac{\partial \langle \tilde{d}_{i_1} \cdots \tilde{d}_{i_N} \rangle}{\partial x_{\chi}^u(\mathbf{b})} (\mathcal{H}^{i_1 \cdots i_N})^* \\
 \mathcal{F}_{N, \chi \chi'}^{uu'}(\mathbf{b}, \mathbf{b}') &= \frac{1}{N!} \left[\frac{\partial \langle \tilde{d}^{i_1} \cdots \tilde{d}^{i_N} \rangle^*}{\partial x_{\chi}^u(\mathbf{b})} \mathbf{S}_{i_1 j_1}^{-1} \right. \\
 &\quad \left. \cdots \mathbf{S}_{i_N j_N}^{-1} \frac{\partial \langle \tilde{d}^{j_1} \cdots \tilde{d}^{j_N} \rangle}{\partial x_{\chi'}^{u'}(\mathbf{b}')} \right]^*, \quad (40)
 \end{aligned}$$

where $h \equiv \mathbf{S}^{-1} \tilde{d}$ and \mathcal{F}_N is Hermitian only if \mathbf{S}^{-1} is. This remains unbiased, such that $\mathbb{E}[\hat{x}_{\chi}^u(\mathbf{b})] = x_{\chi}^u(\mathbf{b})$ for any filter \mathbf{S}^{-1} (due to the definition of \mathcal{F}_N). This implies that the estimator deconvolves the effect of any linear masking and inpainting contained within \mathbf{S}^{-1} . In the limit of $\mathbf{S}^{-1} \rightarrow \mathbf{C}^{-1}$, the estimator becomes optimal, such that \mathcal{F}_N^{-1} is equal to the estimator's covariance matrix; any difference second order in $\mathbf{S} - \mathbf{C}$ (if \mathbf{S}^{-1} is invertible). The remainder of this work is dedicated to exploring how such estimators can be applied in practice.

$$\begin{aligned}
 {}_{s_1 s_2} \mathbf{C}(\hat{\mathbf{n}}_1, \hat{\mathbf{n}}_2) &= W(\hat{\mathbf{n}}_1) W(\hat{\mathbf{n}}_2) \sum_{\ell m} {}_{s_1} Y_{\ell m}(\hat{\mathbf{n}}_1) {}_{s_2} [\mathcal{R} \mathbf{C} \mathcal{R}^{\dagger}]_{s_2} Y_{\ell m}^*(\hat{\mathbf{n}}_2) \\
 &= W(\hat{\mathbf{n}}_1) W(\hat{\mathbf{n}}_2) \sum_{\ell m XY} {}_{s_1} Y_{\ell m}(\hat{\mathbf{n}}_1) {}_{s_1} \mathcal{R}_X \left(\sum_{b, u \in U_2} \frac{C^u(b)}{\Delta_2^u} [\delta_K^{u, XY} + \delta_K^{u, YX}] \Theta_{\ell}(b) B_{\ell}^X B_{\ell}^Y \right) {}_{s_2} \mathcal{R}_Y^{\dagger} Y_{\ell m}^*(\hat{\mathbf{n}}_2) \quad (43)
 \end{aligned}$$

using the definition of the binned power spectrum (28) in the second line. We have additionally inserted an isotropic beam, B_{ℓ} , which differ between temperature and polarization. Inserting this into the numerator of (41) and simplifying gives the explicitly real form

$$\begin{aligned}
 [\mathcal{F}_2 \hat{C}]^u(b) &= \frac{1}{2 \Delta_2^u} \sum_{\ell m XY} \Theta_{\ell}(b) B_{\ell}^X B_{\ell}^Y [\mathcal{R}^{\dagger} W h]_{\ell m}^{X*} \\
 &\quad \times [\delta_K^{u, XY} + \delta_K^{u, YX}] [\mathcal{R}^{\dagger} W h]_{\ell m}^Y, \quad (44)
 \end{aligned}$$

⁶If we include an additive noise term with covariance \mathbf{N} , such that $\mathbf{C} = \sum_{b, u} C^u(b) \partial \mathbf{C} / \partial C^u(b) + \mathbf{N}$, the estimator numerator is modified to

$$[\mathcal{F}_2 \hat{C}]^u(b) = \frac{1}{2} h_i^* \frac{\partial C^{ij}}{\partial C^u(b)} h_j - \frac{1}{2} \text{Tr} \left[\mathbf{S}^{-\dagger} \frac{\partial \mathbf{C}}{\partial C^u(b)} \mathbf{S}^{-1} \mathbf{N} \right]. \quad (42)$$

Here, we absorb the noise into the observed spectrum.

IV. POWER SPECTRUM

The above derivation of the N -point spectrum estimators strictly applies only for $N > 2$. From the Gaussian part of the likelihood (34), we can derive a similar estimator for the bandpower of some pair of fields u in bin b :

$$\begin{aligned}
 [\mathcal{F}_2 \hat{C}]^u(b) &= \frac{1}{2} h_i^* \frac{\partial C^{ij}}{\partial C^u(b)} h_j \\
 \mathcal{F}_2^{uu'}(b, b') &= \frac{1}{2} \text{Tr} \left[\mathbf{S}^{-\dagger} \frac{\partial \mathbf{C}}{\partial C^u(b)} \mathbf{S}^{-1} \frac{\partial \mathbf{C}}{\partial C^{u'}(b')} \right], \quad (41)
 \end{aligned}$$

again assuming a general weighting \mathbf{S}^{-1} and denoting the windowed covariance as $\mathbf{C}_{ij} \equiv \langle \tilde{d}_i \tilde{d}_j^* \rangle$. To derive this we have expanded the likelihood around some fiducial set of bandpowers (since we cannot assume the fiducial value to be zero), the value of which exactly cancels in the estimator, once the $\text{Tr} \log \mathbf{C}$ term is considered.⁶ This bears strong similarities to the higher-point estimators of (40), differing only due to the complex conjugate present in the definition of the power spectrum and the absence of a parity index.

A. Numerator

We now simplify this expression, starting first with the numerator. Starting with the explicit definition of \mathbf{C} in pixel space cf., (11), incorporating the window function, we can write,

where $[\mathcal{R}^{\dagger} W h]_{\ell m}^X \equiv \sum_s ({}_s \mathcal{R}_X^{\dagger})_s [W h]_{\ell m}$, and we have replaced the integrals over $\hat{\mathbf{n}}_{1,2}$ by a harmonic transform. We may additionally compute the sum using only modes with $m \geq 0$ by adding a factor of $(1 + \delta_{m \geq 0}^K)$, which evaluates to unity if $m = 0$ and two else. In the absence of a mask and with an ideal weighting $\mathbf{S}^{-1} \rightarrow \mathbf{C}^{-1}$, $[\mathcal{R}^{\dagger} h]_{\ell m}^X = [\mathbf{C}^{-1} a]_{\ell m}^X$; in general, the various $T/E/B$ components mix due to masking, however. Following these manipulations (44) is relatively straightforward to compute: one performs harmonic transforms to yield the filtered maps $[\mathcal{R}^{\dagger} W h]_{\ell m}^X$ then combines pairwise to yield the estimator, summing over all ℓ modes in the bin and fields of interest.⁷ We note that, even without the mask, the XY power spectrum does not reduce to the canonical form $a_{\ell m}^X a_{\ell m}^{Y*}$ due to the nontrivial correlations between spectra (e.g., a nonzero C_{ℓ}^{TE} prior).

⁷In the practical implementation of these estimators, we replace $[\mathcal{R}^{\dagger} W h]_{\ell m}^X$ by $h_{\ell m}^X$ if the mask is uniform, thus avoiding any unnecessary harmonic transforms.

B. Normalization

Computing the Fisher matrix in (41) is nontrivial due to the trace operation, which naively requires $\mathcal{O}(9N_{\text{pix}}^2)$ operations to implement. To sidestep this, we adopt a similar procedure to [48,52], introducing a set of Gaussian random field (GRF) maps a_i with known and invertible covariance $\mathbf{A} \equiv \langle aa^\dagger \rangle_a$:

$$\mathcal{F}_2^{uu'}(b, b') = \frac{1}{2} \left\langle a^\dagger \mathbf{S}^{-\dagger} \frac{\partial \mathbf{C}}{\partial C^u(b)} \mathbf{S}^{-1} \frac{\partial \mathbf{C}}{\partial C^{u'}(b')} \mathbf{A}^{-1} a \right\rangle_a, \quad (45)$$

where $\langle a^\dagger \mathbf{A}^{-1} a \rangle_a = 1$. Noting that the various covariances and weightings are linear operators, this can be computed in $\mathcal{O}(3N_{\text{pix}})$ time, with the expectation (which replaces the trace) implemented as Monte Carlo average over realizations of a_i . Assuming \mathbf{A}^{-1} is close to \mathbf{S}^{-1} , this converges quickly (roughly as $1 + 1/N_{\text{sim}}$). The above expression can be rewritten succinctly as

$$\mathcal{F}_2^{uu'}(b, b') = \frac{1}{2} \left\langle (Q_2^u[\mathbf{S}^{-1}a](b))^* \cdot [W\mathbf{S}^{-1}W] \cdot Q_2^{u'}[\mathbf{A}^{-1}a](b') \right\rangle_a, \quad (46)$$

defining

$${}_s Q_2^u[x](\hat{\mathbf{n}}; b) = {}_s \left[\frac{\partial \langle dd^\dagger \rangle}{\partial C^u(b)} Wx \right](\hat{\mathbf{n}}), \quad (47)$$

where we have explicitly separated out the mask W . This is simply a product to two maps, Q_2^* and $[W\mathbf{S}^{-1}WQ_2]$, summed over each pixel in real- or harmonic-space (via the Wiener-Khinchin theorem). To compute the normalization we thus perform the following set of operations:

- (1) Generate a GRF a from some fiducial covariance \mathbf{A} .
- (2) Filter this map to obtain $\mathbf{A}^{-1}a$ and $\mathbf{S}^{-1}a$.
- (3) Compute $Q_2^u[x](b)$ maps for each bin of interest.
- (4) Combine pairwise to form the Fisher matrix.
- (5) Average over N_{sim} realizations of a .

Inserting the explicit form for the two-point function (28), we find the following form for Q_2 , expressed in harmonic space:

$$\begin{aligned} {}_s Q_{2,\ell m}^u[x](b) &= \frac{1}{\Delta_2^u} \sum_{XY} \Theta_\ell(b) B_\ell^X B_{\ell s}^Y \mathcal{R}_X \left[\delta_K^{u,XY} + \delta_K^{u,YX} \right] [\mathcal{R}^\dagger Wx]_{\ell m}^Y \\ Q_{2,\ell m}^{u,X}[x](b) &= \frac{1}{\Delta_2^u} \Theta_\ell(b) B_\ell^{u_1} B_\ell^{u_2} \left([\mathcal{R}^\dagger Wx]_{\ell m}^{u_2} \delta_K^{u_1 X} + [\mathcal{R}^\dagger Wx]_{\ell m}^{u_1} \delta_K^{u_2 X} \right), \end{aligned} \quad (48)$$

transforming to $T/E/B$ space in the second line. This is similar to the numerator term of (44), and can be simply implemented via spin-weighted spherical harmonic transforms.⁸

C. Ideal limits

We now consider the limiting form of the above estimators to gain some intuition for their form. In ideal scenarios, we can drop the mask and beam and assume $\mathbf{S}^{-1} = \mathbf{C}^{-1}$, which is translationally invariant. In this case, the $[\mathcal{R}^\dagger Wh]_{\ell m}^X$ term becomes $h_{\ell m}^X \equiv [\mathbf{C}_\ell^{-1} a_{\ell m}]^X$ (since \mathbf{C}^{-1}

involves $\mathcal{R}^{-\dagger} \mathbf{C}^{-1} \mathcal{R}^{-1}$), thus the numerator simplifies to

$$[\mathcal{F}_2 \hat{C}]^u(b) \rightarrow \frac{1}{\Delta_2^u} \sum_{\ell m} \Theta_\ell(b) \text{Re}[h_{\ell m}^{u_1*} h_{\ell m}^{u_2}], \quad (49)$$

where the bracket takes the real part. This is just the standard power spectrum estimator, albeit with Wiener filtering through \mathbf{C}^{-1} .

In the same limit, the normalization can be written in terms of harmonic-space $T/E/B$ quantities as

$$\begin{aligned} \mathcal{F}_2^{uu'}(b, b') &\rightarrow \frac{1}{2} \text{Tr} \left[\mathbf{C}^{-1} \frac{\partial \mathbf{C}}{\partial C^u(b)} \mathbf{C}^{-1} \frac{\partial \mathbf{C}}{\partial C^{u'}(b')} \right] \\ &= \frac{1}{2\Delta_2^u \Delta_2^{u'}} \sum_{\ell} (2\ell + 1) \Theta_\ell(b) \Theta_\ell(b') \sum_{XYZW} \left[\mathbf{C}_\ell^{-1,XY} (\delta_K^{u,YZ} + \delta_K^{u,ZY}) \mathbf{C}_\ell^{-1,ZW} (\delta_K^{u',WX} + \delta_K^{u',XW}) \right] \\ &= \frac{1}{\Delta_2^u \Delta_2^{u'}} \sum_{\ell} (2\ell + 1) \Theta_\ell(b) \Theta_\ell(b') \left[\mathbf{C}_\ell^{-1,u_1 u_1'} \mathbf{C}_\ell^{-1,u_2 u_2'} + \mathbf{C}_\ell^{-1,u_2' u_1} \mathbf{C}_\ell^{-1,u_1 u_2'} \right] \end{aligned} \quad (50)$$

⁸If we had additionally included a noise term in the quadratic estimator, this could be computed in a similar manner, replacing the random fields a by fields drawn from the noise covariance.

inserting the explicit derivatives in the second line, and noting that the covariance is symmetric. Assuming that the bins are disjoint, this is diagonal in b, b' , but not in u, u' . This implies that each ℓ -bin is uncorrelated, but there are nontrivial correlations between the different power spectrum estimators, since \mathbf{C}^{-1} is not generally assumed to be diagonal. If we ignored such correlations, we would find a familiar form involving $\sum_{\ell}(2\ell+1)\Theta_{\ell}^2(b)/C_{\ell}^{XX}C_{\ell}^{ZZ}$. A subtle point is that the correlation matrix (defined by \mathcal{F}_2^{-1} under ideal assumptions) is block diagonal, thus the spectra bifurcate into uncorrelated groups; parity-even

($p_u = 1$) and parity-odd ($p_u = -1$). This is a consequence of rotational symmetry and does not hold for the general case, since masks and filtering can induce leakage between parity-odd and parity-even components.

V. BISPECTRUM

We now turn our attention to three-point functions. Computation proceeds analogously to the power spectrum, beginning with the general estimator for $N = 3$ cf., (40):

$$\begin{aligned} [\mathcal{F}_3 \hat{b}]_{\chi}^u(\mathbf{b}) &= \frac{1}{3!} \frac{\partial \langle \tilde{d}^{i_1} \tilde{d}^{i_2} \tilde{d}^{i_3} \rangle}{\partial b_{\chi}^u(\mathbf{b})} \left(h_{i_1} h_{i_2} h_{i_3} - [h_{i_1} \langle h_{i_2} h_{i_3} \rangle + 2 \text{ perms.}] \right)^*, \\ \mathcal{F}_{3,XX'}^{uu'}(\mathbf{b}, \mathbf{b}') &= \frac{1}{3!} \left[\frac{\partial \langle \tilde{d}^{i_1} \tilde{d}^{i_2} \tilde{d}^{i_3} \rangle^*}{\partial b_{\chi}^u(\mathbf{b})} \mathbf{S}_{i_1 j_1}^{-1} \mathbf{S}_{i_2 j_2}^{-1} \mathbf{S}_{i_3 j_3}^{-1} \frac{\partial \langle \tilde{d}^{j_1} \tilde{d}^{j_2} \tilde{d}^{j_3} \rangle}{\partial b_{\chi'}^{u'}(\mathbf{b}')} \right]^*. \end{aligned} \quad (51)$$

The principal difference between this and the power spectrum estimator is the presence of the linear term in the numerator: this can significantly reduce the large scale variance of the estimator [52] and may be computed via Monte Carlo methods. To this end, we introduce a set of maps $\{\alpha\}$ with covariance equal to that of the data, i.e., $\langle \alpha \alpha^{\dagger} \rangle = \mathbf{C}$.⁹ The linear term is then given by an average over α ;

$$[\mathcal{F}_3 \hat{b}]_{\chi}^u(\mathbf{b}) = \frac{1}{3!} \frac{\partial \langle \tilde{d}^{i_1} \tilde{d}^{i_2} \tilde{d}^{i_3} \rangle}{\partial b_{\chi}^u(\mathbf{b})} \left(h_{i_1} h_{i_2} h_{i_3} - [h_{i_1} \langle [\mathbf{S}^{-1} \alpha]_{i_2} [\mathbf{S}^{-1} \alpha]_{i_3} \rangle_{\alpha} + 2 \text{ perms.}] \right)^*, \quad (52)$$

which can be estimated via Monte Carlo averaging.

A. Numerator

To massage the bispectrum estimator into a practically computable form, we first require the explicit definition of the map-space three-point correlation function in terms of the reduced bispectrum (20),

$$\begin{aligned} \langle s_1 \tilde{d}(\hat{\mathbf{n}}_1)_{s_2} \tilde{d}(\hat{\mathbf{n}}_2)_{s_3} \tilde{d}(\hat{\mathbf{n}}_3) \rangle &= \prod_{i=1}^3 \left[W(\hat{\mathbf{n}}_i) \sum_{\ell_i m_i X_i} B_{\ell_i s_i}^{X_i} Y_{\ell_i m_i}(\hat{\mathbf{n}}_i)_{s_i} \mathcal{R}_{X_i} \right] w_{m_1 m_2 m_3}^{\ell_1 \ell_2 \ell_3} b_{\ell_1 \ell_2 \ell_3}^{X_1 X_2 X_3} \\ &= \frac{1}{2} \prod_{i=1}^3 \left[W(\hat{\mathbf{n}}_i) \sum_{\ell_i m_i X_i} B_{\ell_i s_i}^{X_i} Y_{\ell_i m_i}(\hat{\mathbf{n}}_i)_{s_i} \mathcal{R}_{X_i} \right] w_{m_1 m_2 m_3}^{\ell_1 \ell_2 \ell_3} \\ &\quad \times \sum_{\mathbf{b}, u \in U_3} \frac{b_{\chi}^u(\mathbf{b})}{\Delta_3^u(\mathbf{b})} \left[1 + \chi p_u (-1)^{\ell_1 + \ell_2 + \ell_3} \right] \left[\delta_{\mathbf{K}}^{u, X_1 X_2 X_3} \Theta_{\ell_1}(b_1) \Theta_{\ell_2}(b_2) \Theta_{\ell_3}(b_3) + 5 \text{ perms.} \right], \end{aligned} \quad (53)$$

using the binned bispectrum definition (29) in the second line, and inserting an isotropic beam B_{ℓ}^X . Our next step is to insert this into the estimator numerator and simplify, which leads to the form,

$$\begin{aligned} [\mathcal{F}_3 \hat{b}]_{\chi}^u(\mathbf{b}) &= \frac{1}{2 \Delta_3^u(\mathbf{b})} \sum_{\ell_i m_i} w_{m_1 m_2 m_3}^{\ell_1 \ell_2 \ell_3} \Theta_{\ell_1}(b_1) \Theta_{\ell_2}(b_2) \Theta_{\ell_3}(b_3) B_{\ell_1}^{u_1} B_{\ell_2}^{u_2} B_{\ell_3}^{u_3} [1 + \chi p_u (-1)^{\ell_1 + \ell_2 + \ell_3}] \\ &\quad \times \left\{ [\mathcal{R}^{\dagger} W h]_{\ell_1 m_1}^{u_1} [\mathcal{R}^{\dagger} W h]_{\ell_2 m_2}^{u_2} [\mathcal{R}^{\dagger} W h]_{\ell_3 m_3}^{u_3} - ([\mathcal{R}^{\dagger} W h]_{\ell_1 m_1}^{u_1} \langle [\mathcal{R}^{\dagger} W \alpha]_{\ell_2 m_2}^{u_2} [\mathcal{R}^{\dagger} W \alpha]_{\ell_3 m_3}^{u_3} \rangle_{\alpha} + 2 \text{ perms.}) \right\}^*, \end{aligned} \quad (54)$$

⁹Note that the bispectrum estimator remains unbiased if $\langle \alpha \alpha^{\dagger} \rangle_{\alpha} \neq \mathbf{C}$, though its variance generically increases.

where the spatial integrals (from the i summations) have been replaced with harmonic transforms and we have invoked symmetry to absorb one of the permutation sums. This form of the estimator is not easy to implement, since it involves a coupled sum over three (ℓ, m) pairs, with a

prohibitive $\mathcal{O}(\ell_{\max}^6)$ scaling. To proceed, we must rewrite the expression in a factorizable form, which is made possible by writing the weighting function $w_{m_1 m_2 m_3}^{\ell_1 \ell_2 \ell_3}$ as an integral over spin-weighted spherical harmonics, as in (16). Following some simplification, we find

$$\begin{aligned} [\mathcal{F}_3 \hat{b}]_{\chi}^u(\mathbf{b}) &= \hat{\beta}_{\chi}^u[d, d, d](\mathbf{b}) - \left[\langle \hat{\beta}_{\chi}^u[d, \alpha, \alpha](\mathbf{b}) \rangle_{\alpha} + 2 \text{ perms.} \right] \\ \hat{\beta}_{\chi}^u[\alpha, \beta, \gamma](\mathbf{b}) &\equiv \frac{1}{6\Delta_3^u(\mathbf{b})} \int d\hat{\mathbf{n}} \left\{ {}_{+1}H^{u_1}[\alpha](b_1, \hat{\mathbf{n}}) {}_{+1}H^{u_2}[\beta](b_2, \hat{\mathbf{n}}) {}_{-2}H^{u_3}[\gamma](b_3, \hat{\mathbf{n}}) \right. \\ &\quad \left. + (\chi p_u) {}_{+1}\bar{H}^{u_1}[\alpha](b_1, \hat{\mathbf{n}}) {}_{+1}\bar{H}^{u_2}[\beta](b_2, \hat{\mathbf{n}}) {}_{-2}\bar{H}^{u_3}[\gamma](b_3, \hat{\mathbf{n}}) \right\} + 2 \text{ perms.}, \end{aligned} \quad (55)$$

defining the filtered maps

$${}_s H^X[x](\hat{\mathbf{n}}; b) = \sum_{\ell m} \Theta_{\ell}(b) B_{\ell}^X[\mathcal{R}^{\dagger} \mathbf{W} \mathbf{S}^{-1} x]_{\ell m s}^{X*} Y_{\ell m}(\hat{\mathbf{n}}), \quad {}_s \bar{H}^X[x](\hat{\mathbf{n}}; b) = \sum_{\ell m} (-1)^{\ell} \Theta_{\ell}(b) B_{\ell}^X[\mathcal{R}^{\dagger} \mathbf{W} \mathbf{S}^{-1} x]_{\ell m s}^{X*} Y_{\ell m}(\hat{\mathbf{n}}), \quad (56)$$

which satisfy the conjugation relation

$$[{}_s H^X[x](\hat{\mathbf{n}}; b)]^* = (-1)^s {}_{-s} H^X[x](\hat{\mathbf{n}}; b) = (-1)^s {}_s \bar{H}^X[x](\hat{\mathbf{n}}; b). \quad (57)$$

The conjugate property implies that $\{{}_{+1}H, {}_{-1}H\}$ are a spin- ± 1 pair of fields (rather than ${}_{\pm 1}H$). In (55), we symmetrize both over the choice of fields appearing in the expectation, and over the position of the external spin index (i.e., over nontrivial exchanges of $\{u_1, b_1\}$ pairs). Via the conjugation relations, $(\hat{\beta}_{\chi}^u)^* = \chi p_u \hat{\beta}_{\chi}^u$, thus the estimator simplifies to

$$\begin{aligned} \hat{\beta}_{+}^u[\alpha, \beta, \gamma](\mathbf{b}) &= \frac{1}{3\Delta_3^u(\mathbf{b})} \int d\hat{\mathbf{n}} \text{Re} \{ {}_{+1}H^{u_1}[\alpha](b_1, \hat{\mathbf{n}}) {}_{+1}H^{u_2}[\beta](b_2, \hat{\mathbf{n}}) {}_{-2}H^{u_3}[\gamma](b_3, \hat{\mathbf{n}}) + 2 \text{ perms.} \}, \\ \hat{\beta}_{-}^u[\alpha, \beta, \gamma](\mathbf{b}) &= \frac{i}{3\Delta_3^u(\mathbf{b})} \int d\hat{\mathbf{n}} \text{Im} \{ {}_{+1}H^{u_1}[\alpha](b_1, \hat{\mathbf{n}}) {}_{+1}H^{u_2}[\beta](b_2, \hat{\mathbf{n}}) {}_{-2}H^{u_3}[\gamma](b_3, \hat{\mathbf{n}}) + 2 \text{ perms.} \}. \end{aligned} \quad (58)$$

Despite all the numerous definitions, (55) is straightforward to compute in $\mathcal{O}(3N_{\text{pix}}N_{\text{it}})$ operations, using the following set of operations:

- (1) Draw a GRF α from the (assumed) true covariance \mathbf{C} .
- (2) Filter both α and the data d by the weighting scheme and the mask to form $\mathbf{W} \mathbf{S}^{-1} d$, $\mathbf{W} \mathbf{S}^{-1} \alpha$ maps, then use a spherical harmonic transform to compute $\mathcal{R}^{\dagger} \mathbf{W} \mathbf{S}^{-1} d$ and $\mathcal{R}^{\dagger} \mathbf{W} \mathbf{S}^{-1} \alpha$ in $T/E/B$ space.
- (3) Compute the ${}_s H^X$ and ${}_s \bar{H}^X$ maps for each bin of interest with $X \in \{T, E, B\}$ and $s \in \{+1, -2\}$ via a further spin-weighted spherical harmonic transform.
- (4) Compute $\hat{\beta}$ for each configuration of interest as a real-space summation.
- (5) Iterate over N_{it} random fields α to form the Monte Carlo average and use this to assemble the estimator numerator.

B. Normalization

The bispectrum Fisher matrix can be computed using similar techniques applied to the power spectrum (Sec. IV B, cf., [48]). First, we rewrite the $\mathcal{O}(N_{\text{pix}}^3)$ map-space sum as a Monte Carlo average by introducing some fiducial covariance \mathbf{A} and its inverse as follows:

$$\begin{aligned} \mathbf{S}_{l_2 j_2}^{-1} \mathbf{S}_{l_3 j_3}^{-1} &\rightarrow \mathbf{S}_{l_2 j_2}^{-1} \mathbf{S}_{l_3 j_3}^{-1} \times \frac{1}{2} (\mathbf{A}_{l_2 j_2} \mathbf{A}_{l_3 j_3} + \mathbf{A}_{l_2 j_3} \mathbf{A}_{l_3 j_2}) \\ &\quad \times \mathbf{A}_{j_2 j_2}^{-1} \mathbf{A}_{j_3 j_3}^{-1}, \end{aligned} \quad (59)$$

noting that the Fisher matrix is symmetric under $j_2 \leftrightarrow j_3$. As in Sec. IV, we write each covariance as the expectation of a random field $a^{(i)}$, $\mathbf{A} \equiv \langle a^{(i)} a^{(i)\dagger} \rangle$, allowing the sum to be rewritten as an expectation:

$$\mathbf{S}_{i_2 j_2}^{-1} \mathbf{S}_{i_3 j_3}^{-1} \rightarrow \frac{1}{4} \mathbf{S}_{i_2' j_2'}^{-1} \mathbf{S}_{i_3' j_3'}^{-1} \left\langle a_{i_2'}^{(1)} a_{i_3'}^{(1)} a_{j_2'}^{(1)*} a_{j_3'}^{(1)*} + a_{i_2'}^{(2)} a_{i_3'}^{(2)} a_{j_2'}^{(2)*} a_{j_3'}^{(2)*} - a_{i_2'}^{(1)} a_{i_3'}^{(1)} a_{j_2'}^{(2)*} a_{j_3'}^{(2)*} - a_{i_2'}^{(2)} a_{i_3'}^{(2)} a_{j_2'}^{(1)*} a_{j_3'}^{(1)*} \right\rangle_a \mathbf{A}_{j_2' j_2'}^{-1} \mathbf{A}_{j_3' j_3'}^{-1}, \quad (60)$$

where $a^{(1)}$ and $a^{(2)}$ are independent and identically distributed.¹⁰ Applied to (51), we find

$$\begin{aligned} \mathcal{F}_{3, \chi \chi'}^{uu'}(\mathbf{b}, \mathbf{b}') &= \frac{1}{2} \left(F_{3, \chi \chi'}^{uu', 1111}(\mathbf{b}, \mathbf{b}') + F_{3, \chi \chi'}^{uu', 2222}(\mathbf{b}, \mathbf{b}') - F_{3, \chi \chi'}^{uu', 1122}(\mathbf{b}, \mathbf{b}') - F_{3, \chi \chi'}^{uu', 2211}(\mathbf{b}, \mathbf{b}') \right) \\ F_{3, \chi \chi'}^{uu', \mu\nu\rho\sigma}(\mathbf{b}, \mathbf{b}') &\equiv \frac{1}{12} \left\langle \frac{\partial \langle \tilde{d}^{i_1} \tilde{d}^{i_2} \tilde{d}^{i_3} \rangle^*}{\partial b_{\chi'}^u(\mathbf{b})} \mathbf{S}_{i_1 j_1}^{-1} [\mathbf{S}^{-1} a^{(\mu)}]_{i_2} [\mathbf{S}^{-1} a^{(\nu)}]_{i_3} \times [\mathbf{A}^{-1} a^{(\rho)}]_{j_2}^* [\mathbf{A}^{-1} a^{(\sigma)}]_{j_3}^* \frac{\partial \langle \tilde{d}^{j_1} \tilde{d}^{j_2} \tilde{d}^{j_3} \rangle}{\partial b_{\chi'}^u(\mathbf{b}')} \right\rangle_a \\ &= \frac{1}{12} \left\langle \left(Q_{3, \chi}^u [\mathbf{S}^{-1} a^{(\mu)}, \mathbf{S}^{-1} a^{(\nu)}](\mathbf{b}) \right)^* \cdot [W \mathbf{S}^{-1} W] \cdot \left(Q_{3, \chi'}^u [\mathbf{A}^{-1} a^{(\rho)}, \mathbf{A}^{-1} a^{(\sigma)}] \right) \right\rangle_a^* \end{aligned} \quad (61)$$

defining the pixel-space derivative maps

$$Q_{3, \chi, i}^u[x, y](\mathbf{b}) = \frac{\partial \langle d^i d^j d^k \rangle}{\partial b_{\chi}^u(\mathbf{b})} [W x]_j^* [W y]_k^*. \quad (62)$$

As for the power spectrum, this is the inner product of two maps, for each combination of bins, fields, and parities.

Inserting the definition of the three-point function and simplifying, the Q_3 derivatives can be written in harmonic space

$$\begin{aligned} s_1 Q_{3, \chi, \ell_1 m_1}^u[x, y](\mathbf{b}) &= \frac{1}{2\Delta_3^u(\mathbf{b})} \sum_{\ell_2 \ell_3 m_2 m_3 X_1 X_2 X_3} s_1 \mathcal{R}_{X_1} w_{m_1 m_2 m_3}^{\ell_1 \ell_2 \ell_3} B_{\ell_1}^{X_1} B_{\ell_2}^{X_2} B_{\ell_3}^{X_3} \left[1 + \chi p_u (-1)^{\ell_1 + \ell_2 + \ell_3} \right] \\ &\times \left[\delta_{\mathbf{K}}^{u, X_1 X_2 X_3} \Theta_{\ell_1}(b_1) \Theta_{\ell_2}(b_2) \Theta_{\ell_3}(b_3) + 5 \text{ perms.} \right] [\mathcal{R}^\dagger W x]_{\ell_2 m_2}^{X_2*} [\mathcal{R}^\dagger W y]_{\ell_3 m_3}^{X_3*}. \end{aligned} \quad (63)$$

As for the estimator numerator, this can be simplified by inserting the integral form of the weights:

$$\begin{aligned} s_1 Q_{3, \chi, \ell_1 m_1}^u[x, y](\mathbf{b}) &= \left(\frac{\Theta_{\ell_1}(b_1) B_{\ell_1}^{u_1}}{6\Delta_3^u(\mathbf{b})} s_1 \mathcal{R}_{u_1} \int d\hat{\mathbf{n}} \left\{ {}_{+1} Y_{\ell_1 m_1}(\hat{\mathbf{n}}) {}_{+1} H^{u_2}[x](\hat{\mathbf{n}}; b_2) {}_{-2} H^{u_3}[y](\hat{\mathbf{n}}; b_3) \right. \right. \\ &\left. \left. + \chi p_u (-1)^{\ell_1} {}_{+1} Y_{\ell_1 m_1}(\hat{\mathbf{n}}) {}_{+1} \bar{H}^{u_2}[x](\hat{\mathbf{n}}; b_2) {}_{-2} \bar{H}^{u_3}[y](\hat{\mathbf{n}}; b_3) \right\} + 2 \text{ perms.} \right) + 5 \text{ perms.} \end{aligned} \quad (64)$$

where the first set of permutations are over the positions of the $\{+1, +1, -2\}$ spins and the second over the $\{b, u\}$ pairs. Using the parity properties of spin-weighted spherical harmonics, this can be further rewritten in $T/E/B$ space as

$$\begin{aligned} Q_{3, \chi, \ell_1 m_1}^{u, X}[x, y](\mathbf{b}) &= \left(\frac{\Theta_{\ell_1}(b_1) B_{\ell_1}^{u_1}}{3\Delta_3^u(\mathbf{b})} \delta_{\mathbf{K}}^{u_1 X} \int d\hat{\mathbf{n}} \left\{ {}_{-+1} Y_{\ell_1 m_1}^*(\hat{\mathbf{n}}) {}_{-1} H^{u_2}[x](\hat{\mathbf{n}}; b_2) {}_{+2} H^{u_3}[y](\hat{\mathbf{n}}; b_3) \right. \right. \\ &\left. \left. - (\chi p_u) {}_{-1} Y_{\ell_1 m_1}^*(\hat{\mathbf{n}}) {}_{+1} H^{u_2}[x](\hat{\mathbf{n}}; b_2) {}_{-2} H^{u_3}[y](\hat{\mathbf{n}}; b_3) \right\}^* + 2 \text{ perms.} \right) + 2 \text{ perms.,} \end{aligned} \quad (65)$$

noting that two permutations are equivalent. We additionally note the conjugate relation,

$$\left(Q_{3, \chi, \ell_1 m_1}^{u, X}[x, y](\mathbf{b}) \right)^* = (\chi p_u) \times (-1)^{m_1} Q_{3, \chi, \ell_1 (-m_1)}^{u, X}[x, y](\mathbf{b}), \quad (66)$$

thus Q_3 is real (imaginary) in map space if $\chi p_u = 1$ ($\chi p_u = -1$). The above can be efficiently computed using spin-weighted spherical harmonic transforms, first to define the ${}_{\pm 1} H$ and ${}_{\pm 2} H$ fields, then to perform the $\hat{\mathbf{n}}$ integrals, noting that $Q_{\ell m}$ transforms as a triplet of real fields (i.e., T, Q, U) if we add a factor of i when $\chi p_u = -1$. As such, the derivative, and thus the full Fisher matrix, can be efficiently computed as a summation in real space, or, for a local-in- ℓ weighting and $W = 1$, harmonic space.

¹⁰Whilst we could simply write the expectation as $\mathbf{A}_{i_2' j_2'} \mathbf{A}_{i_3' j_3'} = \langle a_{i_2'}^{(1)} a_{i_3'}^{(1)} a_{j_2'}^{(2)*} a_{j_3'}^{(2)*} \rangle$, the form above makes more efficient use of the random fields.

C. Ideal limits

In the ideal limit of homogeneous noise, unit mask and beam, and ideal Wiener filtering, the bispectrum estimators significantly simplify. As before, we start by noting that $[\mathcal{R}^\dagger W h]_{\ell m}^X \rightarrow h_{\ell m}^X$, thus the bispectrum numerator can be written as

$$\begin{aligned} [\mathcal{F}_3 \hat{b}]_{\chi}^u(\mathbf{b}) &\rightarrow \frac{1}{2\Delta_3^u(\mathbf{b})} \sum_{\ell_i m_i} w_{m_1 m_2 m_3}^{\ell_1 \ell_2 \ell_3} \Theta_{\ell_1}(b_1) \Theta_{\ell_2}(b_2) \Theta_{\ell_3}(b_3) [1 + \chi p_u (-1)^{\ell_1 + \ell_2 + \ell_3}] (h_{\ell_1 m_1}^{u_1} h_{\ell_2 m_2}^{u_2} h_{\ell_3 m_3}^{u_3})^* \\ &= \frac{\chi p_u}{\Delta_3^u(\mathbf{b})} \sum_{\ell_1 \ell_2 \ell_3} \sum_{m_1 m_2 m_3} w_{m_1 m_2 m_3}^{\ell_1 \ell_2 \ell_3} \Theta_{\ell_1}(b_1) \Theta_{\ell_2}(b_2) \Theta_{\ell_3}(b_3) h_{\ell_1 m_1}^{u_1} h_{\ell_2 m_2}^{u_2} h_{\ell_3 m_3}^{u_3}, \end{aligned} \quad (67)$$

with the additional restriction $(-1)^{\ell_1 + \ell_2 + \ell_3} = \chi p_u$. Here, we have dropped the expectation terms, since this fixes $\ell_2 = \ell_3$ which requires $\ell_1 = 0$ after summing over m_i . In the second line, we have replaced terms with their complex conjugates, finding a familiar bispectrum estimator; simply the product of three harmonic-space fields multiplied by the weights. For efficient implementation, one can express the weights in integral form as before,

$$\begin{aligned} [\mathcal{F}_3 \hat{b}]_{\chi}^u(\mathbf{b}) &\rightarrow \frac{1}{6\Delta_3^u(\mathbf{b})} \int d\hat{\mathbf{n}} \left\{ H^{u_1}(\hat{\mathbf{n}}; b_1)_{+1} H^{u_2}(\hat{\mathbf{n}}; b_2)_{-2} H^{u_3}(\hat{\mathbf{n}}; b_3) \right. \\ &\quad \left. + (\chi p_u)_{+1} \bar{H}^{u_1}(\hat{\mathbf{n}}; b_1)_{+1} \bar{H}^{u_2}(\hat{\mathbf{n}}; b_2)_{-2} \bar{H}^{u_3}(\hat{\mathbf{n}}; b_3) \right\} + 2 \text{ perms.}, \end{aligned} \quad (68)$$

with the ideal definitions

$${}_s H^X(\hat{\mathbf{n}}; b) \rightarrow \sum_{\ell m} \Theta_{\ell}(b) h_{\ell m}^{X*} Y_{\ell m}(\hat{\mathbf{n}}), \quad {}_s \bar{H}^X(\hat{\mathbf{n}}; b) \rightarrow \sum_{\ell m} (-1)^{\ell} \Theta_{\ell}(b) h_{\ell m}^{X*} Y_{\ell m}(\hat{\mathbf{n}}). \quad (69)$$

This matches standard prescriptions e.g., [46,69]. Furthermore, from the conjugation properties of (57), this is explicitly real if $\chi p_u = 1$ and imaginary if $\chi p_u = -1$, and can be written succinctly in the form of (58).

For the normalization, we first rewrite in harmonic space in terms of $T/E/B$ fields,

$$\mathcal{F}_{3, \chi \chi'}^{uu'}(\mathbf{b}, \mathbf{b}') \rightarrow \frac{1}{3!} \sum_{\ell_i m_i X_i X_i'} \left[\frac{\partial \langle d_{\ell_1 m_1}^{X_1} d_{\ell_2 m_2}^{X_2} d_{\ell_3 m_3}^{X_3} \rangle^*}{\partial b_{\chi}^u(\mathbf{b})} (\mathbb{C}_{\ell_1}^{-1})^{X_1 X_1'} (\mathbb{C}_{\ell_2}^{-1})^{X_2 X_2'} (\mathbb{C}_{\ell_3}^{-1})^{X_3 X_3'} \frac{\partial \langle d_{\ell_1 m_1}^{X_1'} d_{\ell_2 m_2}^{X_2'} d_{\ell_3 m_3}^{X_3'} \rangle}{\partial b_{\chi'}^{u'}(\mathbf{b}')} \right]^*. \quad (70)$$

Inserting the explicit polyspectrum derivatives, we find

$$\begin{aligned} \mathcal{F}_{3, \chi \chi'}^{uu'}(\mathbf{b}, \mathbf{b}') &\rightarrow \frac{\chi p_u}{6\Delta_3^u(\mathbf{b}) \Delta_3^{u'}(\mathbf{b}')} \sum_{\ell_1 \ell_2 \ell_3} \left(\sum_{m_1 m_2 m_3} -[w_{\ell_1 \ell_2 \ell_3}^{m_1 m_2 m_3}]^2 \right) \sum_{X_i X_i'} \left[\mathbb{C}_{\ell_1}^{-1, X_1 X_1'} \mathbb{C}_{\ell_2}^{-1, X_2 X_2'} \mathbb{C}_{\ell_3}^{-1, X_3 X_3'} \right] \\ &\quad \times \left[\delta_{\mathbb{K}}^{u, X_1 X_2 X_3} \Theta_{\ell_1}(b_1) \Theta_{\ell_2}(b_2) \Theta_{\ell_3}(b_3) + 5 \text{ perms.} \right] \left[\delta_{\mathbb{K}}^{u', X_1' X_2' X_3'} \Theta_{\ell_1}(b_1') \Theta_{\ell_2}(b_2') \Theta_{\ell_3}(b_3') + 5 \text{ perms.} \right], \end{aligned} \quad (71)$$

recalling that \mathbb{C}^{-1} is real. This has the additional restriction $(-1)^{\ell_1 + \ell_2 + \ell_3} = \chi p_u = \chi' p_{u'}$, implying that there are no correlations between physics sourced by different parities. The m_i summation yields

$$\sum_{m_1 m_2 m_3} [w_{\ell_1 \ell_2 \ell_3}^{m_1 m_2 m_3}]^2 = \frac{1}{9} \frac{(2\ell_1 + 1)(2\ell_2 + 1)(2\ell_3 + 1)}{4\pi} \left[\begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ -1 & -1 & 2 \end{pmatrix} + 2 \text{ cyc.} \right]^2, \quad (72)$$

via Wigner $3j$ orthogonality. Finally, we can absorb one of the permutation summations, giving a factor of six, thus

$$\begin{aligned} \mathcal{F}_{3, \chi \chi'}^{uu'}(\mathbf{b}, \mathbf{b}') &\rightarrow \frac{1}{9} \frac{(2\ell_1 + 1)(2\ell_2 + 1)(2\ell_3 + 1)}{4\pi} \left[\begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ -1 & -1 & 2 \end{pmatrix} + 2 \text{ cyc.} \right]^2 \frac{\chi p_u \delta_{\mathbb{K}}^{\chi p_u \chi' p_{u'}}}{\Delta_3^u(\mathbf{b}) \Delta_3^{u'}(\mathbf{b}')} \sum_{\ell_1 \ell_2 \ell_3} \Theta_{\ell_1}(b_1') \Theta_{\ell_2}(b_2') \Theta_{\ell_3}(b_3') \\ &\quad \times \left[\mathbb{C}_{\ell_1}^{-1, u_1 u_1'} \mathbb{C}_{\ell_2}^{-1, u_2 u_2'} \mathbb{C}_{\ell_3}^{-1, u_3 u_3'} \delta_{b_1 b_1'}^{\mathbb{K}} \delta_{b_2 b_2'}^{\mathbb{K}} \delta_{b_3 b_3'}^{\mathbb{K}} + 5 \text{ perms.} \right], \end{aligned} \quad (73)$$

where the permutations are over $\{u_i, b_i\}$ pairs and again we restrict to $(-1)^{\ell_1+\ell_2+\ell_3} = \chi p_u$. Notably, this is not diagonal in fields, with, for example TTT and TEE nontrivially correlated (as for the power spectrum). Some covariances do vanish however, for example, there are no TBB - TTT correlations if $C_\ell^{TB} = 0$. Finally, we note that pairs of bispectra are correlated only if their bins \mathbf{b} and \mathbf{b}' are permutations of each other. Although Sec. II E gave restrictions on which bins are required for each spectrum (which depend on the fields in question), there can be nontrivial correlations with $\mathbf{b} \neq \mathbf{b}'$, for example $b^{TTT}(b_1, b_2, b_3)$ and $b^{TTE}(b_1, b_3, b_2)$ are related. This lies

in contrast with the spin-zero case, where the restriction of $b_1 \leq b_2 \leq b_3$ gave a factor $\delta_{\mathbf{b}, \mathbf{b}'}$, and thus a diagonal ideal Fisher matrix.

VI. TRISPECTRUM

Finally, we turn to the estimation of the polarized four-point function.¹¹ This is essentially analogous to the bispectrum estimator, but made somewhat more involved by the addition of new fields, internal momenta, and a more complex permutation structure. Our starting point is the general $N = 4$ estimator, from (40)

$$\begin{aligned} [\mathcal{F}_4 \hat{t}]_\chi^u(\mathbf{b}, B) &= \frac{1}{4!} \frac{\partial \langle \tilde{d}^{i_1} \dots \tilde{d}^{i_4} \rangle}{\partial t_\chi^u(\mathbf{b}, B)} \left(h_{i_1} h_{i_2} h_{i_3} h_{i_4} - [h_{i_1} h_{i_2} \langle h_{i_3} h_{i_4} \rangle + 5 \text{ perms.}] + [\langle h_{i_1} h_{i_2} \rangle \langle h_{i_3} h_{i_4} \rangle + 2 \text{ perms.}] \right)^* \\ \mathcal{F}_{4, \chi \chi'}^{uu'}(\mathbf{b}, B, \mathbf{b}', B') &= \frac{1}{4!} \left[\frac{\partial \langle \tilde{d}^{i_1} \dots \tilde{d}^{i_4} \rangle}{\partial t_\chi^u(\mathbf{b}, B)} \mathbf{S}_{i_1 j_1}^{-1} \dots \mathbf{S}_{i_4 j_4}^{-1} \frac{\partial \langle \tilde{d}^{j_1} \dots \tilde{d}^{j_4} \rangle}{\partial t_{\chi'}^{u'}(\mathbf{b}', B')} \right]^*, \end{aligned} \quad (74)$$

where the binning depends on four sides $\mathbf{b} \equiv \{b_1, b_2, b_3, b_4\}$ and a diagonal B , as specified in Sec. II E. As in Sec. V, the expectations in the numerator can be computed via Monte Carlo averaging over GRFs with fiducial covariance \mathbf{C} ,

$$\begin{aligned} [\mathcal{F}_4 \hat{t}]_\chi^u(\mathbf{b}, B) &= \frac{1}{4!} \frac{\partial \langle \tilde{d}^{i_1} \dots \tilde{d}^{i_4} \rangle}{\partial t_\chi^u(\mathbf{b}, B)} \left(h_{i_1} h_{i_2} h_{i_3} h_{i_4} - [h_{i_1} h_{i_2} \langle [\mathbf{S}^{-1} \alpha]_{i_3} [\mathbf{S}^{-1} \alpha]_{i_4} \rangle_\alpha + 5 \text{ perms.}] \right. \\ &\quad \left. + \left[\langle [\mathbf{S}^{-1} \alpha^{(1)}]_{i_1} [\mathbf{S}^{-1} \alpha^{(1)}]_{i_2} [\mathbf{S}^{-1} \alpha^{(2)}]_{i_3} [\mathbf{S}^{-1} \alpha^{(2)}]_{i_4} \rangle_\alpha + 2 \text{ perms.} \right] \right)^*, \end{aligned} \quad (75)$$

where $\alpha^{(1)}$ and $\alpha^{(2)}$ are independent and identically distributed. As we show in Sec. VI C, these terms vanish for parity-odd correlators in the ideal limit, but, for parity-even scenarios, perform an important role subtracting off the Gaussian part of the four-point correlator. As such, it is important that the random fields have a covariance close to the true value to avoid biasing the estimator.

A. Numerator

As before, we begin by considering the explicit form of the map-space four-point correlation function (27),

$$\begin{aligned} \langle s_1 \tilde{d}(\hat{\mathbf{n}}_1) \dots s_4 \tilde{d}(\hat{\mathbf{n}}_4) \rangle &= \prod_{i=1}^4 \left[W(\hat{\mathbf{n}}_i) \sum_{\ell_i m_i X_i} B_{\ell_i s_i}^{X_i} Y_{\ell_i m_i}(\hat{\mathbf{n}}_i)_{s_i} \mathcal{R}_{X_i} \right] \sum_{LM} (-1)^M w_{\ell_1 \ell_2 m_1 m_2}^{L(-M)} w_{\ell_3 \ell_4 m_3 m_4}^{LM} t_{\ell_1 \ell_2, \ell_3 \ell_4}^{X_1 X_2, X_3 X_4}(L) + 23 \text{ perms.} \\ &= \frac{1}{2} \prod_{i=1}^4 \left[W(\hat{\mathbf{n}}_i) \sum_{\ell_i m_i X_i} B_{\ell_i s_i}^{X_i} Y_{\ell_i m_i}(\hat{\mathbf{n}}_i)_{s_i} \mathcal{R}_{X_i} \right] \sum_{LM} (-1)^M w_{\ell_1 \ell_2 m_1 m_2}^{L(-M)} w_{\ell_3 \ell_4 m_3 m_4}^{LM} \\ &\quad \times \sum_{\mathbf{b}, u \in U_4, \chi \in \pm 1} \frac{t_\chi^u(\mathbf{b}, B)}{\Delta_4^u(\mathbf{b})} \Theta_L(B) [1 + \chi p_u (-1)^{\ell_1 + \ell_2 + \ell_3 + \ell_4}] \\ &\quad \times \left[\delta_K^{u, XYZW} \Theta_{\ell_1}(b_1) \Theta_{\ell_2}(b_2) \Theta_{\ell_3}(b_3) \Theta_{\ell_4}(b_4) + 7 \text{ perms.} \right] + (2 \leftrightarrow 3) + (2 \leftrightarrow 4). \end{aligned} \quad (76)$$

¹¹Whilst one *could* extend the techniques of this work to higher-point correlators still, this would be somewhat academic since, if the likelihood is quasi-Gaussian, there is a strong descending hierarchy of information content in the N -point correlators. Quadspectra and beyond are similarly of little use for inflationary studies, since they are sourced only by three-particle interactions or higher loops. In the nonlinear regime, higher-order correlators can dominate; however, the N -point correlations functions are perturbative statistics at heart, and thus far from optimal summary statistics in this setting.

Taking the derivative with respect to the trispectrum components and inserting into the estimator numerator yields

$$\begin{aligned}
[\mathcal{F}_4 \hat{t}]_{\chi}^u(\mathbf{b}, B) &= \hat{t}_{\chi}^u[d, d, d, d](\mathbf{b}, B) - \left[\langle \hat{t}_{\chi}^u[d, d, \alpha, \alpha](\mathbf{b}, B) \rangle_{\alpha} + 5 \text{ perms.} \right] \\
&\quad + \left[\langle \hat{t}_{\chi}^u[\alpha^{(1)}, \alpha^{(1)}, \alpha^{(2)}, \alpha^{(2)}](\mathbf{b}, B) \rangle_{\alpha} + 2 \text{ perms.} \right] \\
\hat{t}_{\chi}^u[\alpha, \beta, \gamma, \delta](\mathbf{b}, B) &= \frac{1}{2\Delta_4^u(\mathbf{b})} \sum_{\ell_i m_i} \sum_{LM} (-1)^M w_{\ell_1 \ell_2 m_1 m_2}^{L(-M)} w_{\ell_3 \ell_4 m_3 m_4}^{LM} \Theta_L(B) B_{\ell_1}^{u_1} B_{\ell_2}^{u_2} B_{\ell_3}^{u_3} B_{\ell_4}^{u_4} [1 + \chi p_u (-1)^{\ell_1 + \ell_2 + \ell_3 + \ell_4}] \\
&\quad \times \Theta_{\ell_1}(b_1) \Theta_{\ell_2}(b_2) \Theta_{\ell_3}(b_3) \Theta_{\ell_4}(b_4) [\mathcal{R}^{\dagger} \mathbf{W} \mathbf{S}^{-1} \alpha]_{\ell_1 m_1}^{u_1*} [\mathcal{R}^{\dagger} \mathbf{W} \mathbf{S}^{-1} \beta]_{\ell_2 m_2}^{u_2*} [\mathcal{R}^{\dagger} \mathbf{W} \mathbf{S}^{-1} \gamma]_{\ell_3 m_3}^{u_3*} [\mathcal{R}^{\dagger} \mathbf{W} \mathbf{S}^{-1} \delta]_{\ell_4 m_4}^{u_4*},
\end{aligned} \tag{77}$$

replacing the map-space integrals by harmonic transforms and absorbing the permutations by symmetry. This factorizes into two terms, linked only by L, M ,

$$\begin{aligned}
\hat{t}_{\chi}^u[\alpha, \beta, \gamma, \delta](\mathbf{b}, B) &= \frac{1}{2\Delta_4^u(\mathbf{b})} \sum_{LM} (-1)^M \Theta_L(B) \\
&\quad \times \left\{ A_{b_1 b_2}^{u_1 u_2}[\alpha, \beta](L, -M) A_{b_3 b_4}^{u_3 u_4}[\gamma, \delta](L, M) + (\chi p_u) \bar{A}_{b_1 b_2}^{u_1 u_2}[\alpha, \beta](L, -M) \bar{A}_{b_3 b_4}^{u_3 u_4}[\gamma, \delta](L, M) \right\}
\end{aligned} \tag{78}$$

defining the harmonic space maps

$$\begin{aligned}
A_{b_1 b_2}^{u_1 u_2}[\alpha, \beta](L, M) &= \sum_{\ell_1 \ell_2 m_1 m_2} w_{\ell_1 \ell_2 m_1 m_2}^{LM} \Theta_{\ell_1}(b_1) \Theta_{\ell_2}(b_2) B_{\ell_1}^{u_1} B_{\ell_2}^{u_2} [\mathcal{R}^{\dagger} \mathbf{W} \mathbf{S}^{-1} \alpha]_{\ell_1 m_1}^{u_1*} [\mathcal{R}^{\dagger} \mathbf{W} \mathbf{S}^{-1} \beta]_{\ell_2 m_2}^{u_2*} \\
\bar{A}_{b_1 b_2}^{u_1 u_2}[\alpha, \beta](L, M) &= \sum_{\ell_1 \ell_2 m_1 m_2} (-1)^{\ell_1 + \ell_2 + L} w_{\ell_1 \ell_2 m_1 m_2}^{LM} \Theta_{\ell_1}(b_1) \Theta_{\ell_2}(b_2) B_{\ell_1}^{u_1} B_{\ell_2}^{u_2} [\mathcal{R}^{\dagger} \mathbf{W} \mathbf{S}^{-1} \alpha]_{\ell_1 m_1}^{u_1*} [\mathcal{R}^{\dagger} \mathbf{W} \mathbf{S}^{-1} \beta]_{\ell_2 m_2}^{u_2*}.
\end{aligned} \tag{79}$$

This is a consequence of parametrizing the tetrahedron by sides and diagonals, which splits the shape up into two triangles with a shared diagonal. These have the symmetry property

$$[A_{b_1 b_2}^{u_1 u_2}[\alpha, \beta](L, M)]^* = (-1)^M \bar{A}_{b_1 b_2}^{u_1 u_2}[\alpha, \beta](L, -M), \tag{80}$$

implying that the \hat{t} estimator can be rewritten as

$$\begin{aligned}
\hat{t}_+^u[\alpha, \beta, \gamma, \delta](\mathbf{b}, B) &= \frac{1}{2\Delta_4^u(\mathbf{b})} \sum_{LM \geq 0} (1 + \delta_{M>0}^K) \Theta_L(B) \text{Re} \left\{ \left(\bar{A}_{b_1 b_2}^{u_1 u_2}[\alpha, \beta](L, M) \right)^* A_{b_3 b_4}^{u_3 u_4}[\gamma, \delta](L, M) \right. \\
&\quad \left. + \left(\bar{A}_{b_3 b_4}^{u_3 u_4}[\gamma, \delta](L, M) \right)^* A_{b_1 b_2}^{u_1 u_2}[\alpha, \beta](L, M) \right\}, \\
\hat{t}_-^u[\alpha, \beta, \gamma, \delta](\mathbf{b}, B) &= \frac{i}{\Delta_4^u(\mathbf{b})} \sum_{LM \geq 0} (1 + \delta_{M>0}^K) \Theta_L(B) \text{Im} \left\{ \left(\bar{A}_{b_1 b_2}^{u_1 u_2}[\alpha, \beta](L, M) \right)^* A_{b_3 b_4}^{u_3 u_4}[\gamma, \delta](L, M) \right. \\
&\quad \left. + \left(\bar{A}_{b_3 b_4}^{u_3 u_4}[\gamma, \delta](L, M) \right)^* A_{b_1 b_2}^{u_1 u_2}[\alpha, \beta](L, M) \right\},
\end{aligned} \tag{81}$$

which is considerably simpler to implement, involving only modes with $M \geq 0$ (i.e., those stored by HEALPix).

Next, we insert the integral forms of the trispectrum weights given in (22), yielding

$$\begin{aligned}
A_{b_1 b_2}^{u_1 u_2}[\alpha, \beta](L, M) &= \int d\hat{\mathbf{n}}_{-2} Y_{LM}(\hat{\mathbf{n}})_{+1} H^{u_1}[\alpha](\hat{\mathbf{n}}; b_1)_{+1} H^{u_2}[\beta](\hat{\mathbf{n}}; b_2) \\
&= \left(\int d\hat{\mathbf{n}}_{-2} Y_{LM}^*(\hat{\mathbf{n}})_{-1} H^{u_1}[\alpha](\hat{\mathbf{n}}; b_1)_{-1} H^{u_2}[\beta](\hat{\mathbf{n}}; b_2) \right)^*, \\
\bar{A}_{b_1 b_2}^{u_1 u_2}[\alpha, \beta](L, M) &= (-1)^L \int d\hat{\mathbf{n}}_{-2} Y_{LM}(\hat{\mathbf{n}})_{+1} \bar{H}^{u_1}[\alpha](\hat{\mathbf{n}}; b_1)_{+1} \bar{H}^{u_2}[\beta](\hat{\mathbf{n}}; b_2) \\
&= \left(\int d\hat{\mathbf{n}}_{+2} Y_{LM}^*(\hat{\mathbf{n}})_{+1} H^{u_1}[\alpha](\hat{\mathbf{n}}; b_1)_{+1} H^{u_2}[\beta](\hat{\mathbf{n}}; b_2) \right)^*,
\end{aligned} \tag{82}$$

in terms of the H maps of (56) using the symmetry properties of (57). Despite the chain of definitions, this form is straightforward to implement in $\mathcal{O}(3N_{\text{it}}N_{\text{pix}})$ operations via the following prescription:

- (1) Draw two GRFs $\alpha^{(1)}$ and $\alpha^{(2)}$ from the (close-to) true covariance \mathbf{C} .
- (2) Filter both α and the data d by the weighting scheme and the mask to form $W\mathbf{S}^{-1}d$, $W\mathbf{S}^{-1}\alpha$ maps, then use a spherical harmonic transform to compute $\mathcal{R}^\dagger W\mathbf{S}^{-1}d$ and $\mathcal{R}^\dagger W\mathbf{S}^{-1}\alpha$, as before.
- (3) Compute the ${}_{\pm 1}H^X$ maps for each bin of interest with $X \in \{T, E, B\}$ via a further spin-weighted spherical harmonic transform.

- (4) Compute $A(L, M)$ and $\bar{A}(L, M)$ harmonic-space maps for each pair of bins and fields.
- (5) Combine these pairwise to form both $\tau[d, d, \alpha, \alpha]$ and $\tau[\alpha^{(1)}, \alpha^{(1)}, \alpha^{(2)}, \alpha^{(2)}]$ and their permutations.
- (6) Iterate over $N_{\text{it}}/2$ pairs of random fields to form the Monte Carlo average and use this to assemble the estimator numerator.

B. Normalization

Finally, we consider the trispectrum Fisher matrix, which is computed similarly to the bispectrum piece. As before, we start by prudent introduction of the identity matrix in the following manner:

$$\mathbf{S}_{i_2 j_2}^{-1} \mathbf{S}_{i_3 j_3}^{-1} \mathbf{S}_{i_4 j_4}^{-1} + 5 \text{ perms.} = \mathbf{S}_{i_2 i'_2}^{-1} \mathbf{S}_{i_3 i'_3}^{-1} \mathbf{S}_{i_4 i'_4}^{-1} \times \left[\mathbf{A}_{i'_2 j'_2} \mathbf{A}_{i'_3 j'_3} \mathbf{A}_{i'_4 j'_4} + 5 \text{ perms.} \right] \times \mathbf{A}_{j_2 j_2}^{-1} \mathbf{A}_{j_3 j_3}^{-1} \mathbf{A}_{j_4 j_4}^{-1}, \quad (83)$$

where the permutations contribute equivalently to the Fisher matrix due to index interchange symmetry. Introducing two set of random fields $\alpha^{(1)}, \alpha^{(2)}$, we can rewrite this as

$$\begin{aligned} \mathbf{S}_{i_2 j_2}^{-1} \mathbf{S}_{i_3 j_3}^{-1} \mathbf{S}_{i_4 j_4}^{-1} + 5 \text{ perms.} &= \frac{1}{8} \mathbf{S}_{i_2 i'_2}^{-1} \mathbf{S}_{i_3 i'_3}^{-1} \mathbf{S}_{i_4 i'_4}^{-1} \times \left\langle \left(a_{i'_2}^{(1)} a_{i'_3}^{(1)} a_{i'_4}^{(1)} a_{j_2}^{(1)*} a_{j_3}^{(1)*} a_{j_4}^{(1)*} + a_{i'_2}^{(2)} a_{i'_3}^{(2)} a_{i'_4}^{(2)} a_{j_2}^{(2)*} a_{j_3}^{(2)*} a_{j_4}^{(2)*} \right) \right. \\ &+ 9 \left(a_{i'_2}^{(1)} a_{i'_3}^{(1)} a_{i'_4}^{(2)} a_{j_2}^{(1)*} a_{j_3}^{(1)*} a_{j_4}^{(2)*} + a_{i'_2}^{(1)} a_{i'_3}^{(2)} a_{i'_4}^{(1)} a_{j_2}^{(1)*} a_{j_3}^{(2)*} a_{j_4}^{(1)*} \right) \\ &\left. - 6 \left(a_{i'_2}^{(1)} a_{i'_3}^{(1)} a_{i'_4}^{(1)} a_{j_2}^{(1)*} a_{j_3}^{(2)*} a_{j_4}^{(2)*} + a_{i'_2}^{(2)} a_{i'_3}^{(2)} a_{i'_4}^{(1)} a_{j_2}^{(1)*} a_{j_3}^{(1)*} a_{j_4}^{(2)*} \right) \right\rangle_a \times \mathbf{A}_{j_2 j_2}^{-1} \mathbf{A}_{j_3 j_3}^{-1} \mathbf{A}_{j_4 j_4}^{-1}, \quad (84) \end{aligned}$$

which makes most efficient use of the pairs of simulations, as discussed in [47]. This allows the Fisher matrix to be written in the form:

$$\begin{aligned} \mathcal{F}_{4, \chi \chi'}^{uu'}(\mathbf{b}, B, \mathbf{b}', B') &= \frac{1}{48} \left[\left(F_{4, \chi \chi'}^{uu', 111, 111}(\mathbf{b}, B, \mathbf{b}', B') + F_{4, \chi \chi'}^{uu', 222, 222}(\mathbf{b}, B, \mathbf{b}', B') \right) \right. \\ &+ 9 \left(F_{4, \chi \chi'}^{uu', 112, 112}(\mathbf{b}, B, \mathbf{b}', B') + F_{4, \chi \chi'}^{uu', 122, 122}(\mathbf{b}, B, \mathbf{b}', B') \right) \\ &\left. - 6 \left(F_{4, \chi \chi'}^{uu', 111, 122}(\mathbf{b}, B, \mathbf{b}', B') + F_{4, \chi \chi'}^{uu', 222, 112}(\mathbf{b}, B, \mathbf{b}', B') \right) \right] \\ F_{4, \chi \chi'}^{uu', abc, def}(\mathbf{b}, B, \mathbf{b}', B') &= \frac{1}{4!} \left\langle \frac{\partial \langle \tilde{d}^{i_1} \dots \tilde{d}^{i_4} \rangle^*}{\partial t_{\chi'}^u(\mathbf{b}, B)} \mathbf{S}_{i_1 j_1}^{-1} [\mathbf{S}^{-1} a^{(a)}]_{i_2} [\mathbf{S}^{-1} a^{(b)}]_{i_3} [\mathbf{S}^{-1} a^{(c)}]_{i_4} \right. \\ &\times \left. [\mathbf{A}^{-1} a^{(d)}]_{j_2}^* [\mathbf{A}^{-1} a^{(e)}]_{j_3}^* [\mathbf{A}^{-1} a^{(f)}]_{j_4}^* \frac{\partial \langle \tilde{d}^{j_1} \dots \tilde{d}^{j_4} \rangle^*}{\partial t_{\chi'}^u(\mathbf{b}', B')} \right\rangle_a \\ &= \frac{1}{4!} \left\langle \left(Q_{4, \chi}^u [\mathbf{S}^{-1} a^{(a)}, \mathbf{S}^{-1} a^{(b)}, \mathbf{S}^{-1} a^{(c)}](\mathbf{b}, B) \right)^* \cdot [W\mathbf{S}^{-1}W] \right. \\ &\left. \cdot \left(Q_{4, \chi'}^u [\mathbf{A}^{-1} a^{(d)}, \mathbf{A}^{-1} a^{(e)}, \mathbf{A}^{-1} a^{(f)}](\mathbf{b}', B') \right) \right\rangle_a^*, \quad (85) \end{aligned}$$

with the derivative maps

$$Q_{4, \chi, i}^u[x, y, z](\mathbf{b}, B) = \frac{\partial \langle d^i d^j d^k d^l \rangle}{\partial t_{\chi}^u(\mathbf{b}, B)} [Wx]_j^* [Wy]_k^* [Wz]_l^*. \quad (86)$$

Given the Q_4 maps, the Fisher matrix can thus be computed as an inner product (though we caution that there will be a large number of components, unless the binning is very coarse).

Inserting the binned trispectrum definition (31), the derivative maps take the explicit form in harmonic space

$$\begin{aligned}
s_1 Q_{4,\chi,\ell_1 m_1}^u[x, y, z](\mathbf{b}, B) &= \frac{1}{2\Delta_4^u(\mathbf{b})} \sum_{\ell_i m_i X_i} s_1 \mathcal{R}_{X_1} \sum_{LM} (-1)^M w_{\ell_1 \ell_2 m_1 m_2}^{L(-M)} w_{\ell_3 \ell_4 m_3 m_4}^{LM} \Theta_L(B) B_{\ell_1}^{X_1} B_{\ell_2}^{X_2} B_{\ell_3}^{X_3} B_{\ell_4}^{X_4} \\
&\times [1 + \chi P_u (-1)^{\ell_1 + \ell_2 + \ell_3 + \ell_4}] \left[\delta_K^{u, X_1 X_2 X_3 X_4} \Theta_{\ell_1}(b_1) \Theta_{\ell_2}(b_2) \Theta_{\ell_3}(b_3) \Theta_{\ell_4}(b_4) + 7 \text{ perms.} \right] \\
&\times [\mathcal{R}^\dagger W x]_{\ell_2 m_2}^{X_2*} [\mathcal{R}^\dagger W y]_{\ell_3 m_3}^{X_3*} [\mathcal{R}^\dagger W z]_{\ell_4 m_4}^{X_4*} + (2 \leftrightarrow 3) + (2 \leftrightarrow 4).
\end{aligned} \tag{87}$$

To simplify these further, we can replace the second trispectrum weight by its integral representation and rearrange, finding

$$\begin{aligned}
s_1 Q_{4,\chi,\ell_1 m_1}^u[x, y, z](\mathbf{b}, B) &= \left(\frac{\Theta_{\ell_1}(b_1) B_{\ell_1}^{u_1}}{2\Delta_4^u(\mathbf{b})} s_1 \mathcal{R}_{u_1} \sum_{\ell_2 m_2 LM} (-1)^M w_{\ell_1 \ell_2 m_1 m_2}^{L(-M)} \Theta_L(B) \Theta_{\ell_2}(b_2) B_{\ell_2}^{u_2} [\mathcal{R}^\dagger W x]_{\ell_2 m_2}^{u_2*} \right. \\
&\times \left[A_{b_3 b_4}^{u_3 u_4}[y, z](L, M) + \chi P_u (-1)^{\ell_1 + \ell_2 + L} \bar{A}_{b_3 b_4}^{u_3 u_4}[y, z](L, M) \right] \\
&\left. + 7 \text{ perms.} \right) + (2 \leftrightarrow 3) + (2 \leftrightarrow 4),
\end{aligned} \tag{88}$$

in terms of the A functions of (79). Replacing also the first weighting function yields

$$\begin{aligned}
s_1 Q_{4,\chi,\ell_1 m_1}^u[x, y, z](\mathbf{b}, B) &= \left(\frac{\Theta_{\ell_1}(b_1) B_{\ell_1}^{u_1}}{2\Delta_4^u(\mathbf{b})} s_1 \mathcal{R}_{u_1} \int d\hat{\mathbf{n}} \left\{ {}_{+1} Y_{\ell_1 m_1}(\hat{\mathbf{n}}) {}_{+1} H^{u_2}[x](\hat{\mathbf{n}}; b_2) {}_{-2} A_{b_3 b_4}^{u_3 u_4}[y, z](\hat{\mathbf{n}}; B) \right. \right. \\
&+ \left. \left. (\chi P_u) {}_{-1} Y_{\ell_1 m_1}(\hat{\mathbf{n}}) {}_{-1} H^{u_2}[x](\hat{\mathbf{n}}; b_2) {}_{+2} A_{b_3 b_4}^{u_3 u_4}[y, z](\hat{\mathbf{n}}; B) \right\} \right. \\
&\left. + 7 \text{ perms.} \right) + (x \leftrightarrow y) + (x \leftrightarrow z),
\end{aligned} \tag{89}$$

invoking the A symmetry given in (80), noting that ${}_s \bar{H}^X[x](-\hat{\mathbf{n}}; b) = {}_{-s} H^X[x](\hat{\mathbf{n}}; b)$, and defining the spin- ± 2 maps

$$\begin{aligned}
{}_{+2} A_{b_3 b_4}^{u_3 u_4}[y, z](\hat{\mathbf{n}}; B) &= \sum_{LM} \Theta_L(B) {}_{+2} Y_{LM}(\hat{\mathbf{n}}) A_{b_3 b_4}^{u_3 u_4*}[y, z](L, M), \\
{}_{-2} A_{b_3 b_4}^{u_3 u_4}[y, z](\hat{\mathbf{n}}; B) &= \sum_{LM} \Theta_L(B) {}_{-2} Y_{LM}(\hat{\mathbf{n}}) \bar{A}_{b_3 b_4}^{u_3 u_4*}[y, z](L, M),
\end{aligned} \tag{90}$$

(noting the complex conjugates). Here, the eight permutations in (89) are over pairs of bins and fields, preserving the (b_1, b_2) and (b_3, b_4) pairings. Finally, we can simplify a little further using the above symmetries to yield the $T/E/B$ -space representation,

$$\begin{aligned}
Q_{4,\chi,\ell_1 m_1}^{u,X}[x, y, z](\mathbf{b}, B) &= - \left(\frac{\Theta_{\ell_1}(b_1) B_{\ell_1}^{u_1}}{2\Delta_4^u(\mathbf{b})} \delta_K^{u_1 X} \int d\hat{\mathbf{n}} \left\{ {}_{+1} Y_{\ell_1 m_1}^*(\hat{\mathbf{n}}) {}_{-1} H^{u_2}[x](\hat{\mathbf{n}}; b_2) {}_{+2} A_{b_3 b_4}^{u_3 u_4}[y, z](\hat{\mathbf{n}}; B) \right. \right. \\
&+ \left. \left. (\chi P_u) {}_{-1} Y_{\ell_1 m_1}^*(\hat{\mathbf{n}}) {}_{+1} H^{u_2}[x](\hat{\mathbf{n}}; b_2) {}_{-2} A_{b_3 b_4}^{u_3 u_4}[y, z](\hat{\mathbf{n}}; B) \right\}^* + 7 \text{ perms.} \right) \\
&+ (x \leftrightarrow y) + (x \leftrightarrow z),
\end{aligned} \tag{91}$$

analogous to the Q_3 form of (65). This can be computed as a further harmonic transform, noting that $(Q_{4,\chi,\ell_1 m_1}^{u,X})^* = (\chi P_u) (-1)^{m_1} Q_{4,\chi,\ell_1(-m_1)}^{u,X}$, thus Q_4 consists of three real fields in map space, once we take the imaginary part of any parity-odd correlators.

Given the above manipulations, our route to computing the full Fisher matrix in $\mathcal{O}(3N_{\text{pix}} N_{\text{it}})$ operations becomes:

- (1) Generate random fields $a^{(1)}$ and $a^{(2)}$ from the fiducial covariance \mathbf{A} (which should be close to \mathbf{C} for efficient convergence).
- (2) Compute the $[\mathcal{R}^\dagger W S^{-1} a]$ and $[\mathcal{R}^\dagger W A^{-1} a]$ fields via harmonic transforms for both random fields.
- (3) Compute ${}_{\pm 1} H$ for each field and bin of interest via a harmonic transform.

- (4) Perform a second harmonic transform to compute A and \bar{A} for each pair of bins and fields, then transform to $_{+2}A$ and $_{-2}\bar{A}$, as a function of the diagonal B .
- (5) Perform another transform to compute ${}_sQ$ in harmonic space, weight by the bin, and (if the mask is nontrivial), transform to real space.
- (6) Compute the Fisher matrix by performing an inner product for each pair of spectra in real or harmonic space.
- (7) Assemble the Monte Carlo average by summing over N_{it} realizations of the above procedure.

The above is computationally demanding, due to the large number of chained harmonic transforms. However, it is not infeasible (as demonstrated in Sec. VII), and not parametrically harder than the scalar case, with the principal added complication coming from the significantly larger number of fields in question.

C. Ideal limits

We now consider the limiting form of the trispectrum estimator in ideal settings. We caution that the resulting form for the normalization is not efficient computationally; a faster numerical approach is to use the Monte Carlo procedures of Sec. VI B.

1. Numerator

Akin to before, we begin by writing the numerator explicitly in $T/E/B$ harmonic space in ideal conditions,

$$\begin{aligned}
 [\mathcal{F}_4 \hat{t}]_{\chi}^u(\mathbf{b}, B) &\rightarrow \frac{1}{4!} \sum_{\ell_i m_i X_i} \frac{\partial \langle d_{\ell_1 m_1}^{X_1} d_{\ell_2 m_2}^{X_2} d_{\ell_3 m_3}^{X_3} d_{\ell_4 m_4}^{X_4} \rangle}{\partial t_{\chi}^u(\mathbf{b}, B)} \left(h_{\ell_1 m_1}^{X_1} h_{\ell_2 m_2}^{X_2} h_{\ell_3 m_3}^{X_3} h_{\ell_4 m_4}^{X_4} \right. \\
 &\quad \left. - \left[h_{\ell_1 m_1}^{X_1} h_{\ell_2 m_2}^{X_2} (-1)^{m_4} \delta_{\ell_3 \ell_4}^K \delta_{m_3(-m_4)}^K C_{\ell_3}^{-1, X_3 X_4} + 5 \text{ perms.} \right] \right. \\
 &\quad \left. + \left[(-1)^{m_2+m_4} \delta_{\ell_1 \ell_2}^K \delta_{\ell_3 \ell_4}^K \delta_{m_1(-m_2)}^K \delta_{m_3(-m_4)}^K C_{\ell_1}^{-1, X_1 X_2} C_{\ell_3}^{-1, X_3 X_4} + 2 \text{ perms.} \right] \right)^*, \quad (92)
 \end{aligned}$$

where we have evaluated the expectations analytically. After inserting the binned trispectrum definition, the first piece, involving four fields, takes the explicit form

$$\begin{aligned}
 [\mathcal{F}_4 \hat{t}]_{\chi}^u(\mathbf{b}, B)|_{4\text{-field}} &\rightarrow \frac{\chi P_u}{\Delta_4^u(\mathbf{b})} \sum_{LM} \Theta_L(B) (-1)^M \left(\sum_{\ell_1 \ell_2 m_1 m_2} \Theta_{\ell_1}(b_1) \Theta_{\ell_2}(b_2) w_{\ell_1 \ell_2 m_1 m_2}^{L(-M)} h_{\ell_1 m_1}^{u_1} h_{\ell_2 m_2}^{u_2} \right) \\
 &\quad \times \left(\sum_{\ell_3 \ell_4 m_3 m_4} \Theta_{\ell_3}(b_3) \Theta_{\ell_4}(b_4) w_{\ell_3 \ell_4 m_3 m_4}^{LM} h_{\ell_3 m_3}^{u_3} h_{\ell_4 m_4}^{u_4} \right), \quad (93)
 \end{aligned}$$

with the additional parity restriction that $(-1)^{\ell_1+\ell_2+\ell_3+\ell_4} = \chi P_u$. Practically, this can be computed analogously to the nonideal case (Sec. VI A) with

$$\begin{aligned}
 [\mathcal{F}_4 \hat{t}]_{\chi}^u(\mathbf{b}, B)|_{4\text{-field}} &\rightarrow \frac{1}{2\Delta_4^u(\mathbf{b})} \sum_{LM} (-1)^M \Theta_L(B) \\
 &\quad \times \left\{ A_{b_1 b_2}^{u_1 u_2}[d, d](L, -M) A_{b_3 b_4}^{u_3 u_4}[d, d](L, M) + (\chi P_u) \bar{A}_{b_1 b_2}^{u_1 u_2}[d, d](L, -M) \bar{A}_{b_3 b_4}^{u_3 u_4}[d, d](L, M) \right\}, \quad (94)
 \end{aligned}$$

where A and \bar{A} are defined as before. This can be written as a real or imaginary part as in (81).

For the term involving two fields, we proceed similarly, but note that we can use the following relation to simplify the m_i summation, given the Kronecker deltas and $3j$ symbols in the weights:

$$\begin{aligned}
 \sum_{m_3 m_4 M} (-1)^M \begin{pmatrix} \ell_1 & \ell_2 & L \\ m_1 & m_2 & -M \end{pmatrix} \begin{pmatrix} \ell_3 & \ell_4 & L \\ m_3 & m_4 & M \end{pmatrix} (-1)^{m_4} \delta_{m_3(-m_4)}^K \delta_{\ell_3 \ell_4}^K &= (-1)^{\ell_1+\ell_3-m_1} \sqrt{\frac{2\ell_3+1}{2\ell_1+1}} \delta_{L0}^K \delta_{\ell_1 \ell_2}^K \delta_{\ell_3 \ell_4}^K \delta_{m_1(-m_2)}^K, \\
 \sum_{m_2 m_4 M} (-1)^M \begin{pmatrix} \ell_1 & \ell_2 & L \\ m_1 & m_2 & -M \end{pmatrix} \begin{pmatrix} \ell_3 & \ell_4 & L \\ m_3 & m_4 & M \end{pmatrix} (-1)^{m_4} \delta_{m_2(-m_4)}^K &= \delta_{\ell_1 \ell_3}^K \delta_{\ell_2 \ell_4}^K \delta_{m_1(-m_3)}^K \frac{(-1)^{\ell_2+\ell_3+L-m_1}}{2\ell_1+1}, \quad (95)
 \end{aligned}$$

assuming triangle conditions to be obeyed in both cases. Importantly, terms of the first form cannot contribute, since the weights vanish for $L=0$. The two-field term can thus be written

$$[\mathcal{F}_4 \hat{t}_\chi^u(\mathbf{b}, B)]_{2\text{-field}} \rightarrow -\frac{\chi p_u}{\Delta_4^u(\mathbf{b})} \sum_{LM} \Theta_L(B) (-1)^M \left(\sum_{\ell_1 \ell_2 m_1 m_2} \Theta_{\ell_1}(b_1) \Theta_{\ell_2}(b_2) w_{\ell_1 \ell_2 m_1 m_2}^{L(-M)} \right) \left(\sum_{\ell_3 \ell_4 m_3 m_4} \Theta_{\ell_3}(b_3) \Theta_{\ell_4}(b_4) w_{\ell_3 \ell_4 m_3 m_4}^{LM} \right) \times \left[\langle h_{\ell_1 m_1}^{u_1} h_{\ell_3 m_3}^{u_3} \rangle h_{\ell_2 m_2}^{u_2} h_{\ell_4 m_4}^{u_4} + 3 \text{ perms.} \right], \quad (96)$$

where $(-1)^{\ell_1 + \ell_2 + \ell_3 + \ell_4} = \chi p_u$ and the remaining permutations contract the pairs 2–4, 1–4, and 2–3. Inserting the second relation of (96) and simplifying, we find the simplified two-field term,

$$[\mathcal{F}_4 \hat{t}_\chi^u(\mathbf{b}, B)]_{2\text{-field}} \rightarrow -\frac{\delta_{\chi p_u, 1}^K}{\Delta_4^u(\mathbf{b})} \sum_L \Theta_L(B) \Theta_{\ell_1}(b_1) \Theta_{\ell_2}(b_2) \begin{pmatrix} \ell_1 & \ell_2 & L \\ -1 & -1 & 2 \end{pmatrix}^2 \frac{(2\ell_1 + 1)(2\ell_2 + 1)(2L + 1)}{4\pi} \times (-1)^{\ell_1 + \ell_2 + L} \left(\delta_{b_1 b_3}^K \delta_{b_2 b_4}^K \left[\mathbb{C}_{\ell_1}^{-1, u_1 u_3} (\widehat{\mathbb{C}}_{\ell_2}^{-1})^{u_2 u_4} + (\widehat{\mathbb{C}}_{\ell_1}^{-1})^{u_1 u_3} \mathbb{C}_{\ell_2}^{-1, u_2 u_4} \right] + \delta_{b_1 b_4}^K \delta_{b_2 b_3}^K \left[\mathbb{C}_{\ell_1}^{-1, u_1 u_4} (\widehat{\mathbb{C}}_{\ell_2}^{-1})^{u_2 u_3} + (\widehat{\mathbb{C}}_{\ell_1}^{-1})^{u_1 u_4} \mathbb{C}_{\ell_2}^{-1, u_2 u_3} \right] \right), \quad (97)$$

introducing the empirical power spectrum estimates

$$(\widehat{\mathbb{C}}_{\ell}^{-1})^{uu'} = \frac{\sum_m (-1)^m h_{\ell m}^u h_{\ell(-m)}^{u'}}{2\ell + 1} \equiv \frac{\sum_m h_{\ell m}^u h_{\ell m}^{u'*}}{2\ell + 1}. \quad (98)$$

Finally, the zero-field term is simply obtained as $(-1/2)$ the expectation of the two-field term, yielding

$$[\mathcal{F}_4 \hat{t}_\chi^u(\mathbf{b}, B)]_{0\text{-field}} \rightarrow \frac{\delta_{\chi p_u, 1}^K}{\Delta_4^u(\mathbf{b})} \sum_L \Theta_L(B) \Theta_{\ell_1}(b_1) \Theta_{\ell_2}(b_2) \begin{pmatrix} \ell_1 & \ell_2 & L \\ -1 & -1 & 2 \end{pmatrix}^2 \frac{(2\ell_1 + 1)(2\ell_2 + 1)(2L + 1)}{4\pi} \times (-1)^{\ell_1 + \ell_2 + L} \left(\delta_{b_1 b_3}^K \delta_{b_2 b_4}^K \mathbb{C}_{\ell_1}^{-1, u_1 u_3} \mathbb{C}_{\ell_2}^{-1, u_2 u_4} + \delta_{b_1 b_4}^K \delta_{b_2 b_3}^K \mathbb{C}_{\ell_1}^{-1, u_1 u_4} \mathbb{C}_{\ell_2}^{-1, u_2 u_3} \right), \quad (99)$$

As expected, the two- and zero-field parts of the estimator just involve the (Wiener filtered) power spectra, with weightings appropriate to the binning scheme. These terms generically vanish due to the requirement that $\chi p_u = (-1)^{\ell_1 + \ell_2 + \ell_3 + \ell_4} \rightarrow 1$; as such, they contribute only to parity-even terms if there are an even number of B -modes in the spectrum, and parity-odd terms else. Under the additional assumption of parity-conserving power spectra, we find that the disconnected terms are important only for parity-even physics ($\chi = 1$) in spectra with $p_u = 1$ (e.g., $TTTE$, $EBEB$, etc.).

2. Normalization

We now turn to the Fisher matrix, which can be written explicitly in harmonic $T/E/B$ space as

$$\mathcal{F}_{4, \chi \chi'}^{uu'}(\mathbf{b}, B, \mathbf{b}', B') \rightarrow \frac{1}{4!} \sum_{\ell_i m_i X_i X_i'} \left[\frac{\partial \langle d_{\ell_1 m_1}^{X_1} \cdots d_{\ell_4 m_4}^{X_4} \rangle^*}{\partial t_\chi^u(\mathbf{b}, B)} \mathbb{C}_{\ell_1}^{-1, X_1 X_1'} \cdots \mathbb{C}_{\ell_4}^{-1, X_4 X_4'} \frac{\partial \langle d_{\ell_1 m_1}^{X_1'} \cdots d_{\ell_4 m_4}^{X_4'} \rangle}{\partial t_{\chi'}^{u'}(\mathbf{b}', B')} \right]^*, \quad (100)$$

just as for the bispectrum. To proceed, we insert the reduced trispectrum definition and simplify via the following 3j relations:

$$\begin{aligned} \sum_{m_1 m_2 m_3 m_4 M M'} (-1)^{M+M'} \begin{pmatrix} \ell_1 & \ell_2 & L \\ m_1 & m_2 & -M \end{pmatrix} \begin{pmatrix} \ell_3 & \ell_4 & L \\ m_3 & m_4 & M \end{pmatrix} \begin{pmatrix} \ell_1 & \ell_2 & L' \\ m_1 & m_2 & -M' \end{pmatrix} \begin{pmatrix} \ell_3 & \ell_4 & L' \\ m_3 & m_4 & M' \end{pmatrix} &= \frac{\delta_{LL'}^K}{2L + 1}, \\ \sum_{m_1 m_2 m_3 m_4 M M'} (-1)^{M+M'} \begin{pmatrix} \ell_1 & \ell_2 & L \\ m_1 & m_2 & -M \end{pmatrix} \begin{pmatrix} \ell_3 & \ell_4 & L \\ m_3 & m_4 & M \end{pmatrix} \begin{pmatrix} \ell_1 & \ell_3 & L' \\ m_1 & m_3 & -M' \end{pmatrix} \begin{pmatrix} \ell_2 & \ell_4 & L' \\ m_2 & m_4 & M' \end{pmatrix} &= (-1)^{\ell_2 + \ell_3} \begin{Bmatrix} L & \ell_1 & \ell_2 \\ L' & \ell_4 & \ell_3 \end{Bmatrix}, \\ \sum_{m_1 m_2 m_3 m_4 M M'} (-1)^{M+M'} \begin{pmatrix} \ell_1 & \ell_2 & L \\ m_1 & m_2 & -M \end{pmatrix} \begin{pmatrix} \ell_3 & \ell_4 & L \\ m_3 & m_4 & M \end{pmatrix} \begin{pmatrix} \ell_1 & \ell_4 & L' \\ m_1 & m_4 & -M' \end{pmatrix} \begin{pmatrix} \ell_3 & \ell_2 & L' \\ m_3 & m_2 & M' \end{pmatrix} &= (-1)^{L+L'} \begin{Bmatrix} L & \ell_1 & \ell_2 \\ L' & \ell_3 & \ell_4 \end{Bmatrix}, \end{aligned} \quad (101)$$

cf., [52], which cover the three nontrivial permutations in the trispectrum definition (21). The Fisher matrix can thus be written as a sum of three parts, the first of which involves the first line of (101) and is thus diagonal in B, B' :

$$\mathcal{F}_{4,\chi\chi'}^{uu'}(\mathbf{b}, B, \mathbf{b}', B')|_I \rightarrow \delta_{BB'}^K \frac{(\chi P_u) \delta_{\chi P_u \chi' P_{u'}}^K}{\Delta_4^u(\mathbf{b}) \Delta_4^{u'}(\mathbf{b}')} \sum_{\ell_i L} \Theta_{\ell_1}(b_1) \cdots \Theta_{\ell_4}(b_4) \Theta_L(B) \frac{(2\ell_1 + 1) \cdots (2\ell_4 + 1)(2L + 1)}{(4\pi)^2} \\ \times \begin{pmatrix} \ell_1 & \ell_2 & L \\ -1 & -1 & 2 \end{pmatrix}^2 \begin{pmatrix} \ell_3 & \ell_4 & L \\ -1 & -1 & 2 \end{pmatrix}^2 \left[C_{\ell_1}^{-1, u_1 u'_1} \cdots C_{\ell_4}^{-1, u_4 u'_4} \delta_{b_1 b'_1}^K \cdots \delta_{b_4 b'_4}^K + 7 \text{ perms.} \right], \quad (102)$$

where we have additionally inserted the binned trispectrum definition and asserted that $(-1)^{\ell_1 + \ell_2 + \ell_3 + \ell_4} = \chi P_u = \chi' P_{u'}$. Here, the permutations are those that preserve the 1–2, 3–4 pairings.

The second and third parts of the Fisher matrix are similar, but are not diagonal in L, L' , due to the $6j$ symbols present in (101). This implies that they mix different binned components (as in the scalar scale), though we will find that \mathbf{b} and \mathbf{b}' must always be permutations of each other. The second part takes the explicit form,

$$\mathcal{F}_{4,\chi\chi'}^{uu'}(\mathbf{b}, B, \mathbf{b}', B')|_{II} \rightarrow \frac{(\chi P_u) \delta_{\chi P_u \chi' P_{u'}}^K}{\Delta_4^u(\mathbf{b}) \Delta_4^{u'}(\mathbf{b}')} \sum_{\ell_i L L'} \Theta_{\ell_1}(b_1) \cdots \Theta_{\ell_4}(b_4) \Theta_L(B) \Theta_{L'}(B') \frac{(2\ell_1 + 1) \cdots (2\ell_4 + 1)(2L + 1)(2L' + 1)}{(4\pi)^2} \\ \times \begin{pmatrix} \ell_1 & \ell_2 & L \\ -1 & -1 & 2 \end{pmatrix} \begin{pmatrix} \ell_3 & \ell_4 & L \\ -1 & -1 & 2 \end{pmatrix} \begin{pmatrix} \ell_1 & \ell_3 & L' \\ -1 & -1 & 2 \end{pmatrix} \begin{pmatrix} \ell_2 & \ell_4 & L' \\ -1 & -1 & 2 \end{pmatrix} (-1)^{\ell_2 + \ell_3} \left\{ \begin{matrix} L \ell_1 \ell_2 \\ L' \ell_4 \ell_3 \end{matrix} \right\} \\ \times \left[C_{\ell_1}^{-1, u_1 u'_1} C_{\ell_1}^{-1, u_2 u'_2} C_{\ell_1}^{-1, u_3 u'_3} C_{\ell_4}^{-1, u_4 u'_4} \delta_{b_1 b'_1}^K \delta_{b_2 b'_2}^K \delta_{b_3 b'_3}^K \delta_{b_4 b'_4}^K + 7 \text{ perms.} \right], \quad (103)$$

with permutations exchanging, e.g., $b'_1 \leftrightarrow b'_2$, again asserting $(-1)^{\ell_1 + \ell_2 + \ell_3 + \ell_4} = \chi P_u = \chi' P_{u'}$. Finally, the third contribution is given by

$$\mathcal{F}_{4,\chi\chi'}^{uu'}(\mathbf{b}, B, \mathbf{b}', B')|_{III} \rightarrow \frac{(\chi P_u) \delta_{\chi P_u \chi' P_{u'}}^K}{\Delta_4^u(\mathbf{b}) \Delta_4^{u'}(\mathbf{b}')} \sum_{\ell_i L L'} \Theta_{\ell_1}(b_1) \cdots \Theta_{\ell_4}(b_4) \Theta_L(B) \Theta_{L'}(B') \frac{(2\ell_1 + 1) \cdots (2\ell_4 + 1)(2L + 1)(2L' + 1)}{(4\pi)^2} \\ \times \begin{pmatrix} \ell_1 & \ell_2 & L \\ -1 & -1 & 2 \end{pmatrix} \begin{pmatrix} \ell_3 & \ell_4 & L \\ -1 & -1 & 2 \end{pmatrix} \begin{pmatrix} \ell_1 & \ell_4 & L' \\ -1 & -1 & 2 \end{pmatrix} \begin{pmatrix} \ell_3 & \ell_2 & L' \\ -1 & -1 & 2 \end{pmatrix} (-1)^{L+L'} \left\{ \begin{matrix} L & \ell_1 & \ell_2 \\ L' & \ell_3 & \ell_4 \end{matrix} \right\} \\ \times \left[C_{\ell_1}^{-1, u_1 u'_1} C_{\ell_1}^{-1, u_2 u'_2} C_{\ell_1}^{-1, u_3 u'_3} C_{\ell_4}^{-1, u_4 u'_4} \delta_{b_1 b'_1}^K \delta_{b_2 b'_2}^K \delta_{b_3 b'_3}^K \delta_{b_4 b'_4}^K + 7 \text{ perms.} \right]. \quad (104)$$

As for the bispectrum, \mathcal{F}_4 encodes correlations between different trispectra composed of different fields, though requires $\chi p_u = \chi' p_{u'}$, and, if the power spectra are parity-even, $p_u = p_{u'}$. The correlations between $B' \neq B$ imply that the Fisher matrix (and thus the covariance of the estimator) have a complex structure. This arises from a geometric degeneracy in the tetrahedron definition, since a tetrahedron on the two-sphere can be parametrized by either of two diagonals. This additionally makes the ideal Fisher matrix more work to implement, strictly requiring $\mathcal{O}(\ell_{\max}^6)$ operations. For this reason, the Monte Carlo procedure outlined in Sec. VI B is usually preferred in practice.

VII. VALIDATION

We now validate the power spectrum, bispectrum, and trispectrum algorithms derived above. All estimators are implemented in the publicly available PolyBin package,¹² the

scalar version of which was described in [52]. The estimators make extensive use of the HEALPix [70] and LIBSHARP [71] codes to manipulate data on the two-sphere and perform spin-weighted spherical harmonic transforms. Alongside the addition of tensor fields, the code has been extensively written since its genesis in [52], optimizing for both CPU and memory efficiency, in part by careful avoidance of any unnecessary harmonic transforms (for example, if the mask is uniform). The scalar part of the new code has been explicitly tested against the old, to ensure calculations are consistent to machine precision, where appropriate.¹³ In the code (and in the plots below), we take the imaginary parts of any parity-breaking bispectrum and trispectrum such that the outputs are explicitly real.

¹³Note that we use a different weighting convention for bispectrum in this work to allow for parity-odd contributions, thus the bispectrum outputs differ by the appropriately averaged ratio of the two weights.

¹²[GitHub.com/OliverPhilcox/PolyBin](https://github.com/OliverPhilcox/PolyBin).

A. Setup

We first present an overview of our testing methodology. To validate the estimators, we use simulated Gaussian realizations of the T , Q , and U polarized CMB maps via HEALPix's `synalm` routine [70], with input lensed power spectra derived from the CLASS code [72], at the *Planck* 2018 cosmology [8], and noise via the standard prescription of [73]. We additionally include a parity-breaking signal via the correlation coefficients $r_{TB} = r_{EB} = 0.5$; this facilitates testing of the parity-odd power spectrum sector and covariance matrices. All fields are generated at the desired testing resolution, here set to $N_{\text{side}} = 256$ for fast computation, thus we do not require pixel window functions, and we additionally ignore any harmonic space beam. For the window, we utilize the *Planck* GAL040 mask with 2° apodization [74], which masks all but $\approx 34\%$ of the sky, practically removing the majority of power at $\ell \lesssim 5$.¹⁴ When Monte Carlo simulations are required, these are generated analogously to the mock data, with $N_{\text{it}} = 100$ iterations used by default. Whilst our testing suite is principally focused on CMB applications, we stress that it applies also to any other spin-0/spin-2 fields on the sphere, such as cosmic shear and its correlation with galaxy overdensity.

An important ingredient of the estimators is the weighting function \mathbf{S}^{-1} . As noted in Sec. III, the optimal choice is $\mathbf{S}^{-1} = \mathbf{C}^{-1}$, i.e., the inverse pixel covariance. Here, we shall ignore the impact of the mask (noting that this is unity in the unmasked regions, ignoring apodization), and fix

\mathbf{S}^{-1} to be the diagonal-in- ℓ inverse harmonic-space covariance \mathbf{C}_ℓ^{-1} , including the cross-covariances between all fields. As we see below, this is approximately optimal in practice for $\ell \gtrsim 5$.

When testing the three-point estimators, it is useful to generate simulations with a known bispectrum, rather than restricting to GRFs. For this purpose, we generate random maps, $a_{\ell m}$, via a procedure similar to [46,48], with the $T/E/B$ -space redefinition

$$a_{\ell m}^X \rightarrow a_{\ell m}^X + \sum_{\ell_2 \ell_3 m_2 m_3 YZ} \frac{1}{6} w_{m_1 m_2 m_3}^{\ell_1 \ell_2 \ell_3} b_{\ell_1 \ell_2 \ell_3}^{XYZ} h_{\ell_2 m_2}^{Y*} h_{\ell_3 m_3}^{Z*}. \quad (105)$$

Here, $w_{m_1 m_2 m_3}^{\ell_1 \ell_2 \ell_3}$ is the bispectrum weighting function given in §II, $h_{\ell m}^X \equiv [\mathbf{C}_\ell^{-1} a]_{\ell m}^X$ is the inverse-covariance-weighted field and $b_{\ell_1 \ell_2 \ell_3}^{XYZ}$ is the desired bispectrum. At first order in b^{XYZ} , this satisfies

$$\langle a_{\ell_1 m_1}^X a_{\ell_2 m_2}^Y a_{\ell_3 m_3}^Z \rangle = w_{m_1 m_2 m_3}^{\ell_1 \ell_2 \ell_3} b_{\ell_1 \ell_2 \ell_3}^{XYZ}, \quad (106)$$

as desired, though we note that higher-order corrections may be important if b^{XYZ} is large; furthermore, the power spectrum is distorted by an amount proportional to $(b^{XYZ})^2$. If one assumes a separable bispectrum, $b_{\ell_1 \ell_2 \ell_3}^{XYZ} = \beta_{\ell_1}^X \beta_{\ell_2}^Y \beta_{\ell_3}^Z$, this can be practically implemented as a real-space integral,

$$a_{\ell m}^X \rightarrow a_{\ell m}^X + \frac{1}{18} \beta_\ell^X \int d\hat{\mathbf{n}} [-_1 \mathcal{O}(\hat{\mathbf{n}})_{-1} \mathcal{O}(\hat{\mathbf{n}})_{-2} Y_{\ell m}^*(\hat{\mathbf{n}}) - _{2-1} \mathcal{O}(\hat{\mathbf{n}})_{+2} \mathcal{O}(\hat{\mathbf{n}})_{+1} Y_{\ell m}^*(\hat{\mathbf{n}})], \quad (107)$$

defining

$${}_s \mathcal{O}(\hat{\mathbf{n}}) = \sum_{\ell m X} \beta_\ell^X h_{\ell m}^X Y_{\ell m}(\hat{\mathbf{n}}), \quad (108)$$

making use of the aforementioned conjugate properties. As before, this can be evaluated via spin-weighted spherical harmonic transforms. We may additionally insert a factor of $[1 \pm p_u (-1)^{\ell_1 + \ell_2 + \ell_3}] / 2$ in order to inject input spectra of only a single parity.

In general, our estimators return measurements of the binned polyspectra, which, for comparison to theory, should be connected to the unbinned models.¹⁵ For completeness, we list these relations below, derived by taking expectations of the ideal estimator and inserting the unbinned polyspectrum definitions given in Sec. II. Firstly, the power spectrum becomes

$$\begin{aligned} [\mathcal{F}_2^{-1} C]^{\text{th},u}(b) &= \frac{1}{\Delta_2^u} \sum_{\ell} \Theta_{\ell}(b) (2\ell + 1) \sum_{u'} S_{\ell}^{-1, u_1 u_1'} S_{\ell}^{-1, u_2 u_2'} C_{\ell}^{\text{th}, u'} \\ \mathcal{F}_2^{\text{th}, u u'}(b) &= \frac{1}{\Delta_2^u \Delta_2^{u'}} \sum_{\ell} \Theta_{\ell}(b) (2\ell + 1) \left[S_{\ell}^{-1, u_1 u_1'} S_{\ell}^{-1, u_2 u_2'} + S_{\ell}^{-1, u_1 u_2'} S_{\ell}^{-1, u_2 u_1'} \right], \end{aligned} \quad (109)$$

¹⁴As discussed in [60], an alternative approach is to include the mask only in the weighting function \mathbf{S}^{-1} , setting the window, W , to unity elsewhere in the algorithm. This may be preferable for complex masks containing an abundance of holes, and leads to somewhat expedited computation.

¹⁵A simple approximation is to evaluate the relevant theory polyspectrum in the bin center; this will often suffice in practice if the bins are relatively thin.

with $C^{\text{th},u}(b) \rightarrow C_\ell^{\text{th},u}$ in the narrow-bin limit. Notably, the sum over u' in the first line is over *all* pairs of fields, not just those in (9), since we have not inserted the binned definition; this additionally leads to the permutations in the second line of (109). For the bispectrum, we find

$$\begin{aligned}
 [\mathcal{F}_3^{-1}b]_\chi^{\text{th},u}(\mathbf{b}) &= \frac{1}{9} \frac{\chi P_u}{\Delta_3^u(\mathbf{b})\Delta_3^u(\mathbf{b}')} \sum_{\ell_1\ell_2\ell_3} \Theta_{\ell_1}(b_1)\Theta_{\ell_2}(b_2)\Theta_{\ell_3}(b_3) \frac{(2\ell_1+1)(2\ell_2+1)(2\ell_3+1)}{4\pi} \left[\begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ -1 & -1 & 2 \end{pmatrix} + 2 \text{ cyc.} \right]^2 \\
 &\quad \times \sum_{u'} S_{\ell_1}^{-1,u_1u'_1} S_{\ell_2}^{-1,u_2u'_2} S_{\ell_3}^{-1,u_3u'_3} b_{\ell_1\ell_2\ell_3}^{\text{th},u'}, \\
 \mathcal{F}_{3,\chi\chi'}^{\text{th},uu'}(\mathbf{b},\mathbf{b}') &= \frac{1}{9} \frac{\chi P_u \delta_K^{\chi P_u \chi' P_u}}{\Delta_3^u(\mathbf{b})\Delta_3^u(\mathbf{b}')} \sum_{\ell_1\ell_2\ell_3} \Theta_{\ell_1}(b'_1)\Theta_{\ell_2}(b'_2)\Theta_{\ell_3}(b'_3) \frac{(2\ell_1+1)(2\ell_2+1)(2\ell_3+1)}{4\pi} \left[\begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ -1 & -1 & 2 \end{pmatrix} + 2 \text{ cyc.} \right]^2 \\
 &\quad \times \left[S_{\ell_1}^{-1,u_1u'_1} S_{\ell_2}^{-1,u_2u'_2} S_{\ell_3}^{-1,u_3u'_3} \delta_{b_1b'_1}^K \delta_{b_2b'_2}^K \delta_{b_3b'_3}^K + 5 \text{ perms.} \right], \tag{110}
 \end{aligned}$$

where we restrict to triplets with $(-1)^{\ell_1+\ell_2+\ell_3} = \chi P_u$, and the u' summation is taken over all triplets of fields. As before, this has the thin-bin limit $b^{\text{th},u} \rightarrow b_{\ell_1\ell_2\ell_3}^{\text{th},u}$. Finally, the binned trispectrum takes the form (cf., [52,60] for scalars)

$$\begin{aligned}
 [\mathcal{F}_4\hat{t}]_\chi^{\text{th},u}(\mathbf{b},B) &= 8 \frac{\chi P_u}{\Delta_4^u(\mathbf{b})} \sum_{\ell_1L} \Theta_{\ell_1}(b_1) \cdots \Theta_{\ell_4}(b_4) \Theta_L(B) \frac{(2\ell_1+1) \cdots (2\ell_4+1)(2L+1)}{(4\pi)^2} \\
 &\quad \times \begin{pmatrix} \ell_1 & \ell_2 & L \\ -1 & -1 & 2 \end{pmatrix}^2 \begin{pmatrix} \ell_3 & \ell_4 & L \\ -1 & -1 & 2 \end{pmatrix}^2 \sum_{u'} S_{\ell_1}^{-1,u_1u'_1} \cdots S_{\ell_4}^{-1,u_4u'_4} t_{\ell_1\ell_2\ell_3\ell_4}^{\text{th},u'}(L) \\
 \mathcal{F}_{4,\chi\chi'}^{\text{th},uu'}(\mathbf{b},B,\mathbf{b}',B') &= \delta_{BB'}^K \frac{\chi P_u \delta_K^{\chi P_u \chi' P_u}}{\Delta_4^u(\mathbf{b})\Delta_4^u(\mathbf{b}')} \sum_{\ell_1L} \Theta_{\ell_1}(b_1) \cdots \Theta_{\ell_4}(b_4) \Theta_L(B) \frac{(2\ell_1+1) \cdots (2\ell_4+1)(2L+1)}{(4\pi)^2} \\
 &\quad \times \begin{pmatrix} \ell_1 & \ell_2 & L \\ -1 & -1 & 2 \end{pmatrix}^2 \begin{pmatrix} \ell_3 & \ell_4 & L \\ -1 & -1 & 2 \end{pmatrix}^2 \left[S_{\ell_1}^{-1,u_1u'_1} \cdots S_{\ell_4}^{-1,u_4u'_4} \delta_{b_1b'_1}^K \cdots \delta_{b_4b'_4}^K + 7 \text{ perms.} \right], \tag{111}
 \end{aligned}$$

summing over all quadruplets of fields u' , with the additional parity restriction that $(-1)^{\ell_1+\ell_2+\ell_3+\ell_4} = \chi P_u$. Here, we have assumed that the off-diagonal terms in the Fisher matrix (which mix L and L') are subdominant Sec. VIC). As noted in [38], when applied to real data, these relations are not quite exact, since the true expectation of the estimator includes the variation of the mask across the bins in the numerator and denominator. For higher-order statistics, such corrections are small and can generally be ignored.

B. Power spectrum

To test the two-point estimators, we compute the binned power spectrum multipoles over an array of 40

regularly-spaced bins between $\ell = 2$ and $\ell = 402$, for all six nontrivial combinations of T , E , and B , giving a total of $N_{\text{bin}} = 240$ components. This is performed with and without the *Planck* sky mask; in each case, we compute both the Fisher matrix, \mathcal{F}_2 , and the estimator numerators, via the unwrapped prescription and the ideal limits.

In Table I, we list the runtime of the optimal and ideal estimators. Both with and without the mask function, this is dominated by computation of the Fisher matrix, which must be estimated for each pair of bins and fields from a suite of N_{it} GRF realizations. In the presence of a nontrivial window, this requires harmonic transforms for each element, giving a total of $2 \times 3 \times N_{\text{bin}} \times N_{\text{it}}$ (counting

TABLE I. Computation times for ideal and optimal power spectrum estimators, for $N_{\text{side}} = 256$ and 40 bins with $\ell \in [2, 402]$. We show results for both the estimator numerators and Fisher matrices, with the latter computed from $N_{\text{it}} = 100$ Monte Carlo simulations. Timings are given for both a uniform mask and a *Planck* sky mask; each example requires two harmonic transforms to define the numerator (one for spin 0 and one for spin 2), with the latter requiring an additional ≈ 2000 transforms for each realization of the optimal Fisher matrix. This occurs since each of the 240 elements of the map-space Q_2 derivative is estimated via a reverse harmonic transform (for both spins), then the S^{-1} weighting is applied, which requires a further transformation.

Mask	Ideal Fisher (s)	Optimal Fisher	Ideal numerator (s)	Optimal numerator (s)
Uniform	19	18 s \times 100	1.6	1.6
Galactic	19	45 s \times 100	1.6	1.6

number of spins, number of transforms, number of bins, and number of simulations). Whilst this may be somewhat expensive for high-resolution maps, we stress that it need be computed only once for a given survey geometry and fiducial cosmology, i.e., it is independent of the data. For a uniform mask, no harmonic transforms are needed since all computations can be performed in $T/E/B$ harmonic-space. Computation of the estimator numerators is straightforward, requiring only pair of transforms to define the filtered data, and subsequent harmonic-space summations in each bin. For high-resolution maps at a given N_{side} , computation of this part is essentially independent of N_{bin} (unlike for the higher-point functions).

Figure 1 displays the output binned power spectrum measurements from the PolyBin code against theoretical expectations, all computed using a galactic mask. For each of the six spectra considered, we find excellent agreement between empirical and predicted spectra, both for the unwinded estimators and their ideal equivalents. The last bin of the unwinded estimators is somewhat biased; this is as expected, since we do not correctly account for mask-induced correlations between it and the next bin. In practice, one would always compute the spectrum across a slightly wider range of ℓ -bins than used in the final analysis, allowing such effects to be nulled.

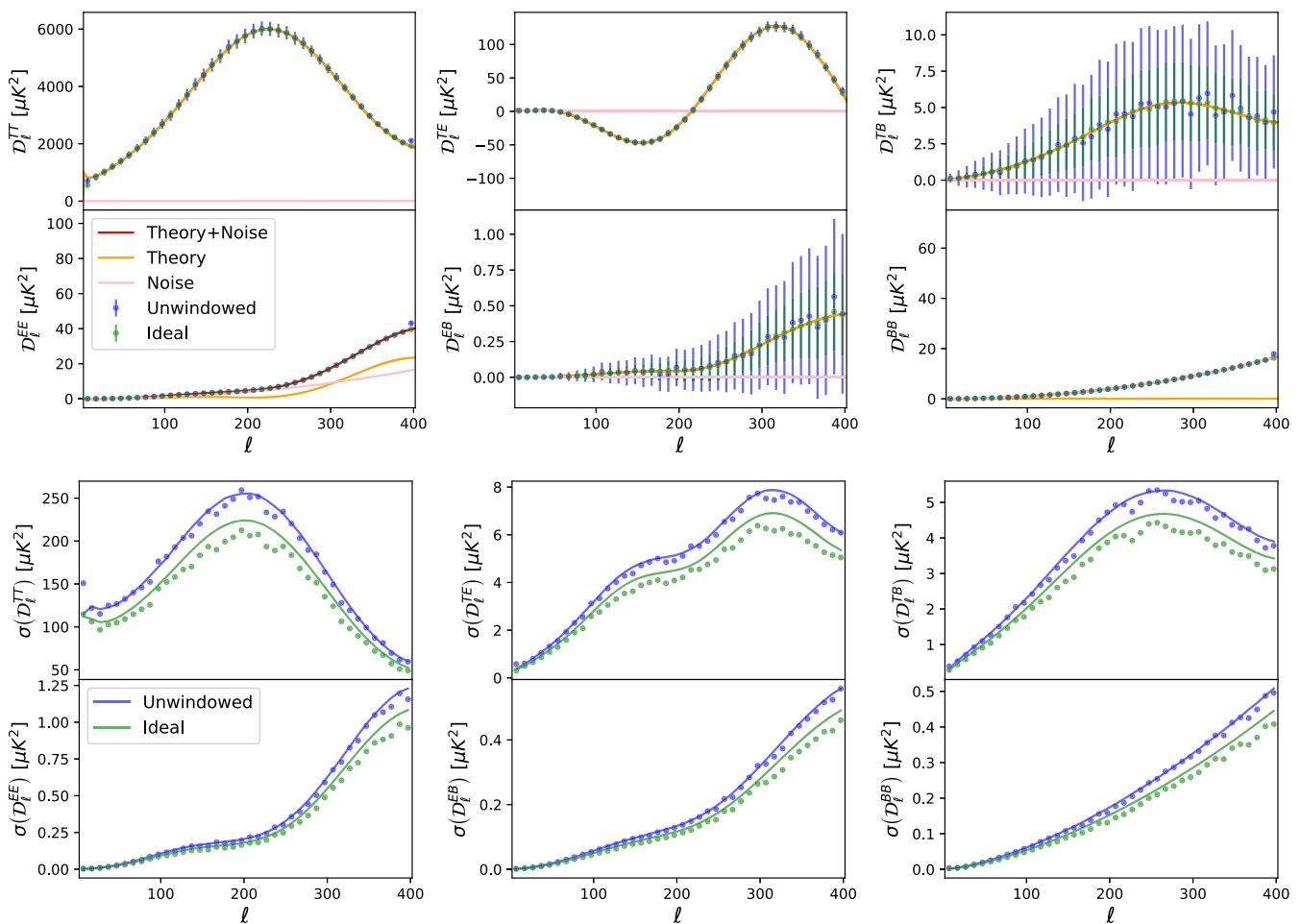


FIG. 1. Measured and theoretical power spectra from a suite of simulations, obtained using the unwinded and idealized estimators of this work implemented in polybin. Simulations are constructed using a *Planck* cosmology and noise model and a galactic-sky mask, but include additional parity-odd correlations for testing. We show measurements from 1000 GRF simulations, with Fisher matrices (needed for the unwinded/optimal estimators) constructed from 100 Monte Carlo realizations. The top panels show the measured spectra using the two estimators alongside the theoretical expectations, rescaled by $\ell(\ell + 1)/2\pi$. In the bottom panel, we plot the estimator errors (points) and the diagonal of the (unwinded or ideal) Fisher matrix. For the ideal estimators, we rescale the means and variances by factors of $\langle W^2 \rangle / \langle W \rangle^2$ and $\langle W^4 \rangle / \langle W^2 \rangle^2$ respectively, to account for the mask effects; this holds only approximately, thus the ideal variances (green) may not precisely match their predictions. If the estimator is unbiased and optimal, the mean of the unwinded estimator mean will match theory, and the variance will match the inverse Fisher matrix; this is demonstrated in the above.

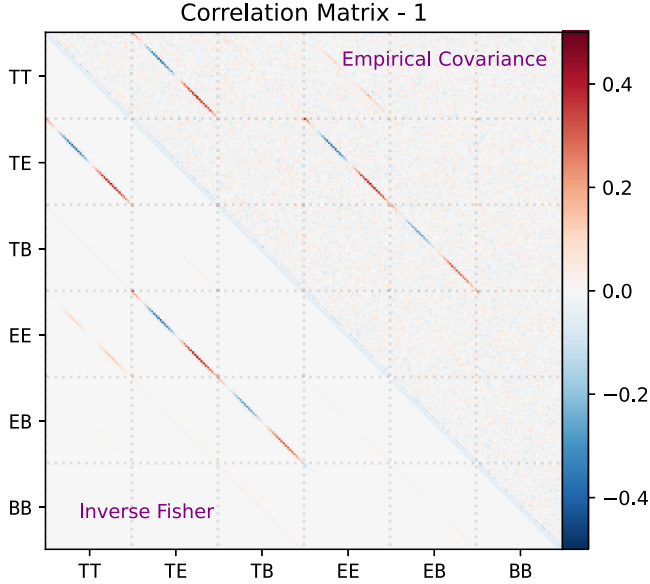


FIG. 2. Correlation matrices corresponding to the unwinded power spectrum measurements shown in Fig. 1. The lower-half matrix shows the theoretical correlation obtained from the inverse Fisher matrix, whilst the upper half shows that extracted from a suite of 1000 mocks. Sub-boxes show correlations between individual spectra, with ℓ -bins stacked from lowest (top left) to highest (bottom-right). We find excellent agreement between theory and data, and observe off-diagonal correlations from both couplings between different polarized components and the survey mask.

Turning to the variances, the unwinded results show excellent agreement between the empirical variances and the inverse Fisher matrix predictions. This indicates that our estimators are (approximately) minimum-variance, and thus optimal in the Gaussian regime. The idealized estimators show fair agreement with their idealized Fisher matrix predictions, though we note that this prediction does not fully account for the spatial dependence of the window, and is thus only approximate. As in [52], the unwinded estimators appear to have higher variances than their idealized equivalents: this is due to an anticorrelation between neighboring bins, and does not lead to reduced signal-to-noise (i.e., it does not imply suboptimality).

In Fig. 2, we show the correlation structure of the optimal estimator, both theoretically (from the inverse Fisher matrix) and empirically (from the simulation suite). As before, we find excellent agreement, with no discernible difference seen between the two matrices (except for noise). We observe strong off-diagonal signals induced by physical TE , TB , and EB correlations (with only the first expected from conventional physics), whilst the effects of masking lead to a leakage of power between adjacent bins, as evidenced by the negative elements alongside the diagonal. The similarity of both variances and correlations is an

important verification of our estimator and its optimality. Finally, we note that our results are stable to the choice of hyperparameters; reducing to just $N_{\text{it}} = 10$ Monte Carlo iterations used to estimate the Fisher matrix yields an error of $(0.0 \pm 0.3)\sigma$ across all bins, implying that our choice of $N_{\text{it}} = 100$ is conservative.

C. Bispectrum

As for the power spectrum, we validate the bispectrum estimators by ensuring that (a) the mean of the estimators matches the theoretical expectation, and (b) the variance is close to the inverse Fisher matrix, i.e., the optimal limit. For this purpose, we analyze simulations both with and without an injected three-point function, noting that the estimators are strictly minimum-variance only in the Gaussian limit. For this test, we compute the bispectrum in $N_\ell = 9\ell$ -bins equally spaced in $\ell^{3/2}$ (chosen to evenly distribute any primordial signal-to-noise [75]), with $\ell \in [2, 440]$.¹⁶ We utilize the same binning for both squeezed and non-squeezed triangles, though note that this is not required by PolyBin; alternative choices may be preferred if one is testing a model that peaks in squeezed configurations (such as local-type primordial non-Gaussianity). In practice, we drop the first bin after the full bispectrum is computed (thus accounting for bin leakage), since this is heavily contaminated by the window function and generally non-Gaussian; a similar procedure was used for the trispectrum in [60]. Gaussian simulations are generated as before, with non-Gaussianity incorporated for even $\ell_1 + \ell_2 + \ell_3$ via the dimensionless factorized bispectrum $b_{\ell_1 \ell_2 \ell_3}^{XYZ} = \beta_{\ell_1}^X \beta_{\ell_2}^Y \beta_{\ell_3}^Z$, with $\beta_\ell^T = (2 \times 10^{-6})e^{-(\ell-2)/40}$, $\beta_\ell^E = \sqrt{C_{\ell=80}^{EE}/C_{\ell=80}^{TT}}\beta_\ell^T$, and $\beta_\ell^B = 0$, set by initial testing to ensure a signal of appropriate strength. We compute bispectra for all 20 combinations of fields and parities (cf., Sec. V), leading to a total of $N_{\text{bin}} = 2140$ bins (after dropping modes with $\ell < 5$).

Before presenting the bispectrum results, it is useful to understand their scalings by estimating the number of spherical harmonic transforms involved in each part of the estimator. The numerator requires $2 \times 3 \times N_\ell$ transforms to define the $_{\pm 1}H$ and $_{\pm 2}H$ fields (one per spin, one per field, and one per ℓ -bin), then the three-field term is evaluated using a real-space summation. When the one-field term is included, this is enhanced by a factor of N_{it} , yielding a total of $\approx 6N_{\text{it}}N_\ell$ transforms for the numerator. For the Fisher matrix, each realization requires N_{bin} harmonic transforms to form the \mathcal{Q}_3 arrays, which are then combined in harmonic- or real-space, with the latter requiring an additional set of transforms to apply the \mathbf{S}^{-1} weighting and the mask. As such, the total Fisher matrix scaling is $\mathcal{O}(N_{\text{it}}N_{\text{bin}})$, whilst the scaling of the numerator is

¹⁶The exact bin edges are fixed to $\{2, 5, 30, 68, 114, 167, 227, 293, 364, 440\}$.

TABLE II. Runtimes of the ideal and unwindowed bispectrum estimators, applied to every combination of fields and parities across 9 linear bins with $\ell \in [2, 440)$, giving 2448 bins (some of which are later dropped). The ideal estimator requires six harmonic transforms for each linear bin (for the two spins, and three $T/E/B$ maps), with the remaining computation involving only real-space summation; however, the normalization requires an $\mathcal{O}(\ell_{\max}^3)$ sum, which takes an additional ≈ 10 hours to compute on a 40-core machine. The same scaling holds for the three-field term of the optimal estimator; however, the one-field term requires the above transformations for each of $N_{\text{it}} = 100$ Monte Carlo realizations (including permutations of data and fields), leading to the significantly larger computation time. For the optimal Fisher matrix, we require around 4600 harmonic transforms per realization for a uniform mask (to define the Q_3 derivatives); this increases to around 54,000 for a nontrivial mask, since one must perform forward and reverse transforms to apply W to Q_3 , for both its spin-0 and spin-2 components, and for two possible weightings.

Mask	Ideal Fisher (h)	Optimal Fisher	Ideal numerator (s)	Optimal numerator [3-field] (s)	Optimal numerator [full] (m)
Uniform	10	770 s \times 100	37	37	230
Galactic	10	1280 s \times 100	36	36	158

$\mathcal{O}(N_{\text{it}}N_{\ell})$.¹⁷ Finally, we note that the computation of the Fisher matrix may be practically limited by computational memory limits rather than runtime. Each Monte Carlo iteration requires Q_3 to be computed in each bin, for two random fields, and for two weightings (\mathbf{S}^{-1} and \mathbf{A}^{-1}). After this, the arrays are combined as an inner product, requiring $4N_{\text{bin}}$ maps to be held in memory. For $N_{\text{side}} = 256$ and the above binning configuration, this requires ≈ 120 GB of memory, which is well within the reach of most clusters.¹⁸ For higher-resolution studies, one may wish to use wider bins to reduce the memory consumption, or, in the most restrictive limit, compute the inner product elementwise, leading to an algorithm with quadratic complexity rather than linear.

Computation times for the bispectrum estimator are shown in Table II, when applied to both a *Planck* galactic mask and a trivial uniform mask. With our choice of binning, each Monte Carlo realization of the Fisher matrices require ≈ 10 – 20 minutes to compute, with the windowed computation taking longer due to the additional harmonic transforms required. Due to our low N_{side} , the transforms are not strongly rate-limiting; this is evidenced by the nontrivial mask computations being only a factor of two slower than the uniform case, despite involving $\approx 10\times$ more harmonic transforms. In this example, computing the Fisher matrix via Monte Carlo summation is actually more efficient than direct computation of the idealized form. The latter involves an $\mathcal{O}(\ell_{\max}^3)$ summation, which took ≈ 10 hours on a 40-core machine. The estimator numerators are fast to compute: the three-field term takes under a minute per simulation (for all ~ 2500 bins), whilst the one-field correction multiplies this by a factor around $N_{\text{it}} = 100$. The latter computation can be somewhat expedited by pre-processing the Monte Carlo simulations and holding them

in memory (which itself required ≈ 4 minutes); if memory is a limitation, this can be dropped, such that each simulation is analyzed only as necessary. As discussed above, the action of the one-field term is to reduce the variance only on ultralarge scales and does not affect the mean; for studies focussing on smaller scales, this contribution can be dropped for significant computational gain. As such, the fastest bispectrum estimator is the optimal form (with a Monte Carlo Fisher matrix), but without the one-field term.

Figure 3 shows the bispectra measured with our pipeline, averaged over 1000 realizations. We find excellent agreement between theory and data for both types of estimator across all fields and parities. This demonstrates that our algorithm can correctly recover an injected bispectrum, i.e., it is unbiased. The various data points show a large dynamic range; this is due to the large range of ℓ ($5 \leq \ell < 440$), and the differences in amplitude between T , E , and B . We note that some $\chi p_u = -1$ data points have very large scatter; this corresponds to bins with $b_1 = b_2 = b_3$, which have no contribution from odd $\ell_1 + \ell_2 + \ell_3$ in the thin-bin limit.

In Fig. 4, we display the empirical bispectrum variances from GRF simulations (always including a Galactic mask). We find excellent agreement between the empirical unwindowed variances and the inverse Fisher predictions for all correlators across a huge range of amplitudes; this implies that our estimator is close to minimum variance, as expected.¹⁹ As for the power spectrum, the ideal estimator shows slightly reduced variances compared to the optimal form; this is again sourced by an anticorrelation between neighboring bins. These are relatively well modeled by the ideal Fisher matrix (multiplied by $\langle W^6 \rangle / \langle W^3 \rangle^2$ for mask W), though exact agreement is not guaranteed in this case. If one pushes to larger scales than plotted ($\ell < 5$), both estimators become suboptimal; this occurs since the bispectrum becomes mask dominated, thus our diagonal-in- ℓ \mathbf{S}^{-1} weighting is far from the true inverse-pixel covariance. In this regime, a more nuanced weighting, for instance invoking conjugate-gradient-descent inversion,

¹⁷Note that the total number of bispectrum bins, N_{bin} , scales as N_{ℓ}^3 if one considers all possible triangle configurations.

¹⁸Note that the algorithm is constructed so as to reduce memory requirements for all steps apart from the (irreducible) Q_3 maps, i.e., it minimizes storage and recomputation of intermediate products.

¹⁹Repeating the exercise with a uniform mask finds exquisite agreement in all bins, as expected.

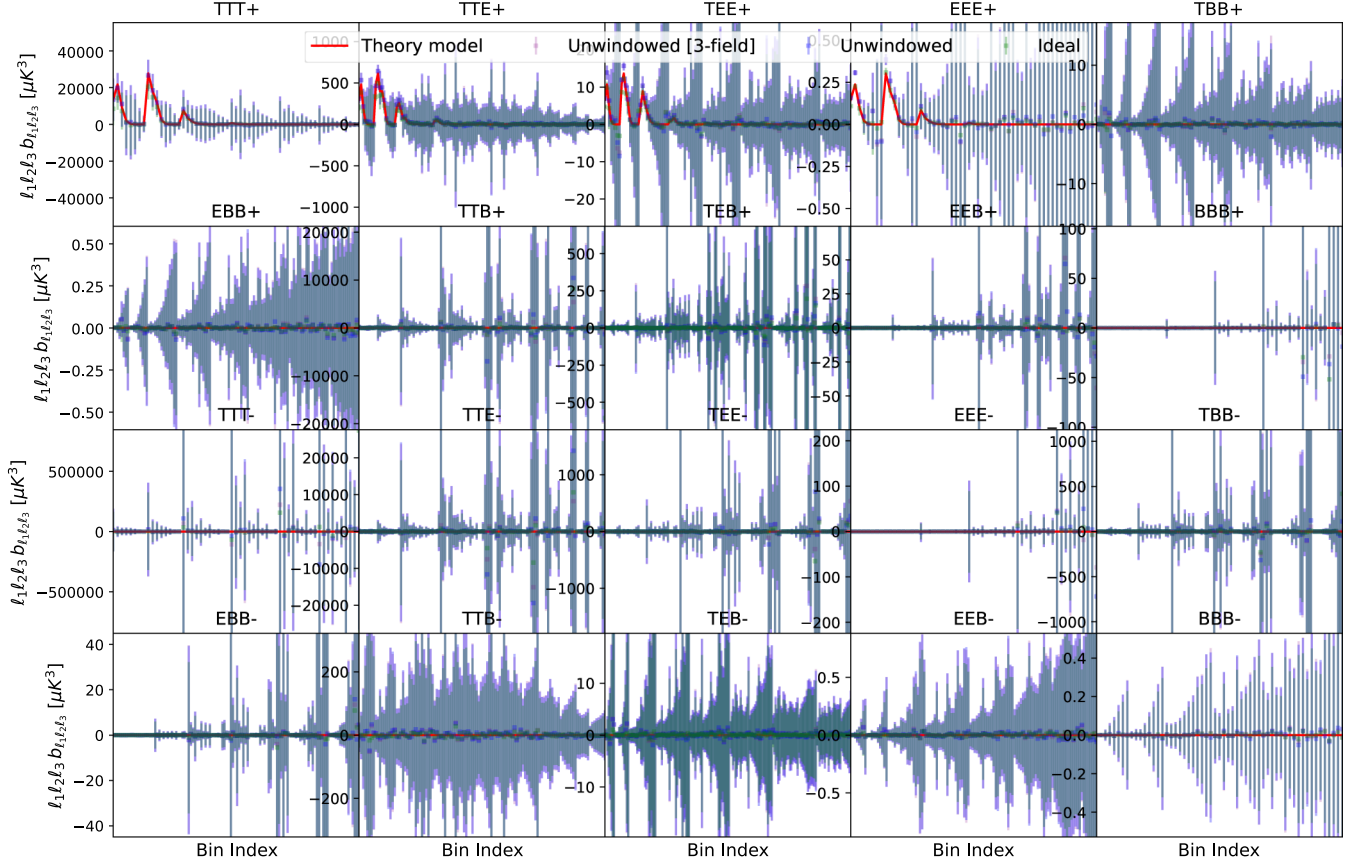


FIG. 3. Measured and theoretical bispectra from a suite of masked simulations, measured using the unwound and idealized estimators of this work. This figure is analogous to Fig. 1, but includes all ten nontrivial bispectrum field configurations and both parities (indicated by the titles, with \pm indicating $\chi = \pm 1$). In each subpanel, the bins are stacked in one dimension and increase in magnitude from the left to the right of the figure. We show the mean bispectra measured from 1000 simulations from both types of estimators, and compare this to the injected parity-even signal (red curve). We find that the theory is well-recovered in all relevant cases, with null results seen in correlators for which it does not contribute. The errors strongly depend on the bin configuration, with particularly large values seen if $p_{\ell\chi} = -1$ and all three bins are equal (since this term vanishes in the thin-bin limit).

may be of use. Finally, we note that on these range of scales, the one-field term has little effect; for more complex masks, it may serve to reduce the large-scale variance, particularly when coupled with an accurate estimate of \mathbf{S}^{-1} .²⁰

Figure 5 shows the structure of the (inverted) optimal Fisher matrix, alongside the empirical correlations estimated from the 1000 GRF simulations. We observe a complex geometric structure coupled across bins and fields; this arises from the power spectrum correlations between T , E , and B (which break parity in our testing suite). That this is far from diagonal implies that computing the polarized binned bispectra is nontrivial; the

results would be biased if one assumes a diagonal normalization term (which is often the case in unpolarized analyses e.g., [69]). If one zooms in sufficiently, a complex structure within each submatrix becomes visible; this is induced both by intrinsic correlations between fields and the nonuniform mask. The latter effect contributes at the $\sim 5\%$ level (coupling adjacent bins), but is significantly amplified if the bin width is reduced. In all cases, the correlation structure is well-recovered from the suite of simulations again validating our estimator. Finally, we note that similar results can be obtained with significantly fewer realizations used to create the Fisher matrix; reducing from $N_{it} = 100$ to $N_{it} = 10$ changes the measured bispectra by at most 0.007σ .

D. Trispectrum

Finally, we validate the trispectrum estimators. Our procedure is essentially identical to that for the lower-order statistics; we check both the mean and the variance of the estimator using suites of simulations with and without a

²⁰The utility of the one-field term depends also on the contributions from various scales within each bin, and thus the Gaunt factor weights $w_{m_1 m_2 m_3}^{\ell_1 \ell_2 \ell_3}$. Whilst this factors out for thin bins with $\Delta\ell = 1$, it may lead to some weighting schemes being preferred in practice [such as the $\binom{\ell_1 \ell_2 \ell_3}{0 \ 0 \ 0}$ weights used for parity-even only analyses.]

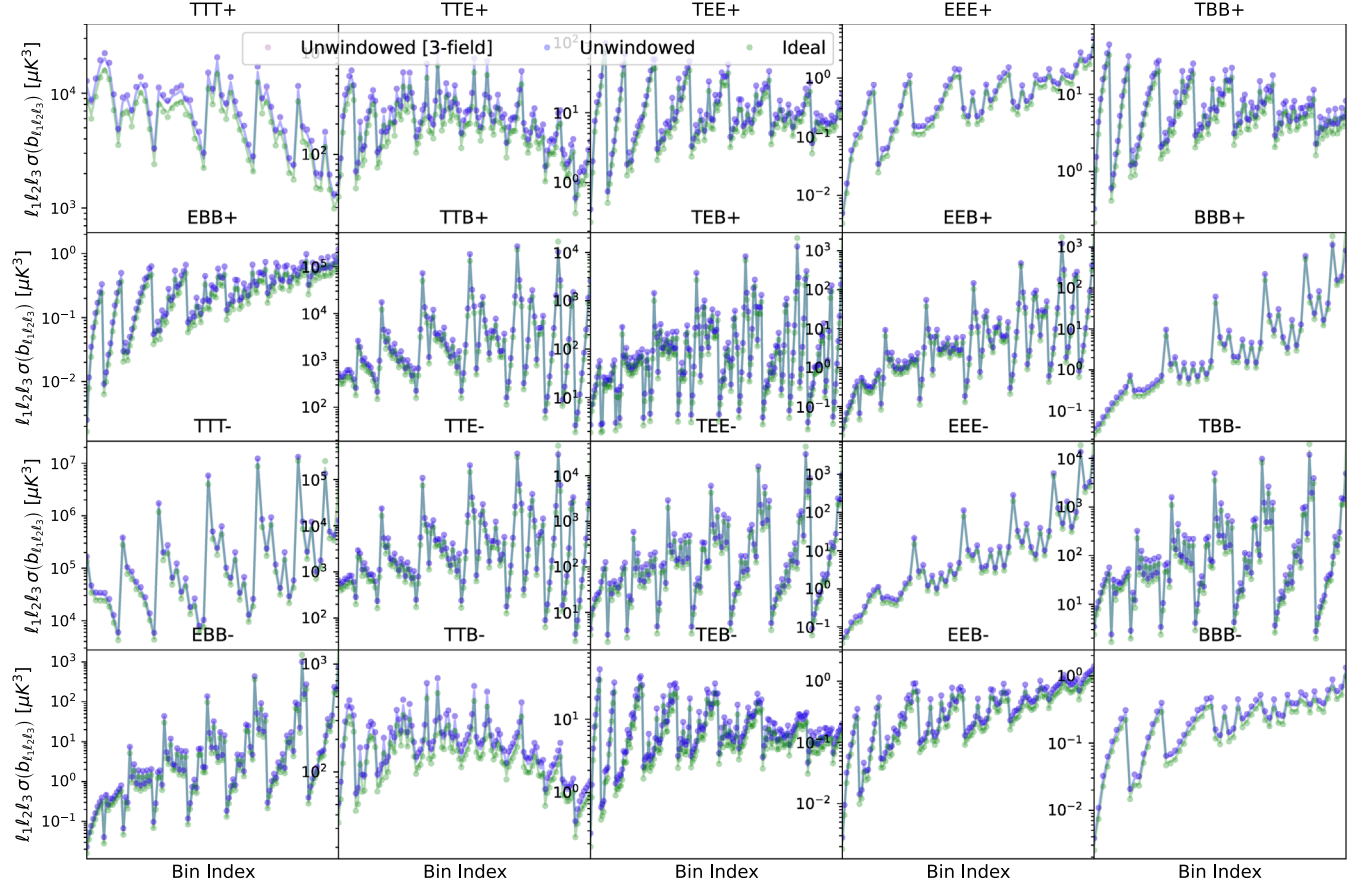


FIG. 4. As Fig. 3, but displaying the empirical bispectrum variances calculated from GRF simulations; these are in good agreement with the expected variances across all correlators, implying that our estimators are close to optimal. We find that removal of the one-field term in the estimators (purple) significantly enhances the variances on large scales.

Galactic mask. To check for bias, we consider the individual four-, two-, and zero-field terms in the estimator explicitly rather than attempting to inject a trispectrum directly; since the full estimator is the sum of these components and the odd-parity estimator differs from the even only by a factor of $(-1)^{\ell_1+\ell_2+\ell_3+\ell_4}$, this is a sufficient test (and indeed, that used in [52]). To keep computation times reasonable, we restrict to 4 ℓ -bins evenly spaced in $\ell^{5/2}$, with $\ell_{\min} = 2$ and $\ell_{\max} = 483$,²¹ and use the same binning for all legs, though note that PolyBin can use different binning to highlight squeezed (large ℓ_2, ℓ_4) and flattened (small L) configurations. As before, we drop the first ℓ -bin in the final analysis (after accounting for bin leakage), since it is mask dominated, and our choice of \mathbf{S}^{-1} weighting is far from optimal in this case. Here, we consider both parities, but now restrict to field configurations containing at most one B mode. This emulates physical scenarios when one is interested in computing scalar physics from T and E modes, but wishes to limit E -to- B mode leakage at leading order. In total, we use

12 fields and 2 parities, giving a total of 820 bins in the output data vector. Finally, we compute the ideal trispectra only from 100 GRF simulations due to its large computational requirements (but use 1000 for the unwindable estimators). In all cases, we use 100 Monte Carlo realizations to compute the disconnected contributions.

The computational scalings of the trispectrum are akin to those of the bispectrum. First, one computes the ${}_{\pm 1}H$ maps, requiring $\mathcal{O}(N_\ell)$ harmonic transforms, then obtains the $A(L, M)$ fields, which scales as N_ℓ^2 . The four-field term of the trispectrum numerator is computed as a harmonic-space summation, yielding an overall complexity of $\mathcal{O}(N_\ell^2)$. The two- and zero-field terms require computation of H and A for $N_{\text{it}}/2$ pairs of simulations, thus the final scaling is $\mathcal{O}(N_{\text{it}}N_\ell^2)$. This is slower than the bispectrum, but notably much reduced compared to the total number of trispectrum bins (which scales as N_ℓ^5). For the Fisher matrix, Q_4 requires a harmonic transform for each A field and L -bin (i.e., a total of $\sim N_\ell^3$), then a harmonic transform to form $Q_{4,\ell m}^X$, giving a total of $\mathcal{O}(N_\ell^4)$ operations. In the presence of a nontrivial mask, these must be additionally transformed into real space, yielding a total complexity of

²¹These are defined by the bin edges $\{2, 5, 62, 212, 483\}$.

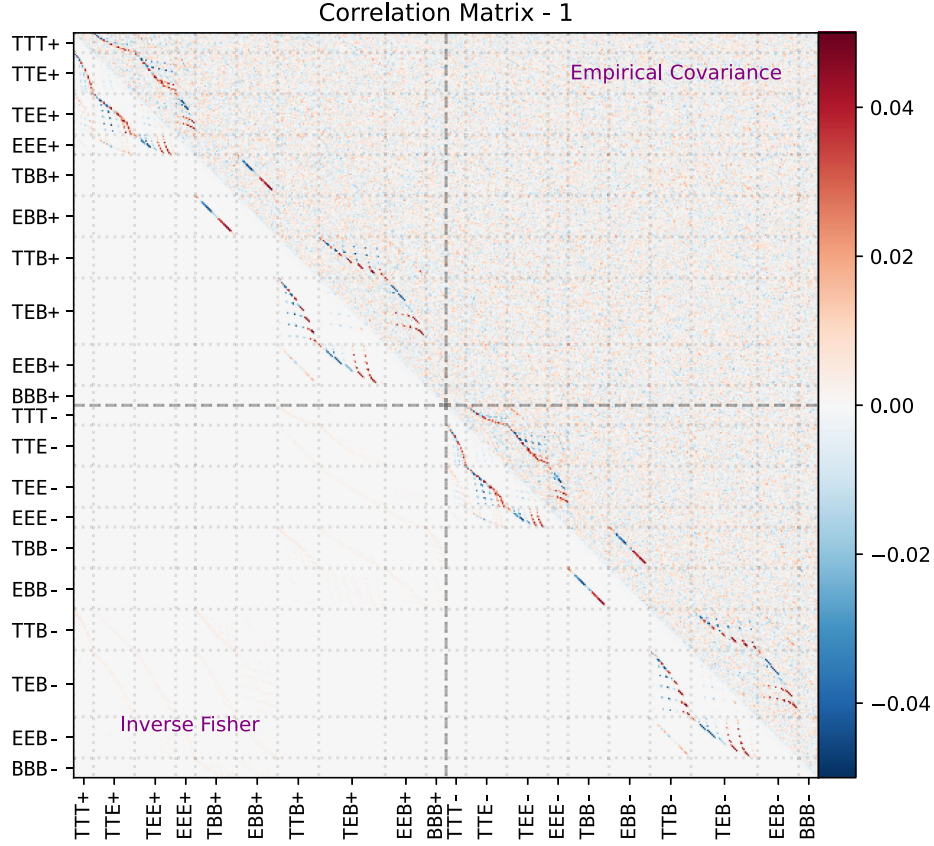


FIG. 5. Correlation matrices corresponding to the unwindowed bispectrum measurements shown in Fig. 3. As in Fig. 2, the lower-half matrix shows the theoretical correlation obtained from the inverse Fisher matrix, whilst the upper half shows that extracted from a suite of 1000 mocks. We demarcate components with $\chi = 1$ and $\chi = -1$ by dashed lines, with individual sub-boxes showing the correlations between pairs of bispectra. The covariance matrix has a complex structure sourced by different polarized components and the survey mask, which is excellently reproduced by the theoretical prediction.

$\mathcal{O}(N_{it}N_{bin})$ from N_{it} realizations. As before, this requires memory proportional to N_{bin} (precisely $8N_{bin}$ times the dimensionality of map, due to the various permutations and weightings required), which may be limiting in practice. Whilst the above demonstrates that trispectrum computation is expensive, it is not parametrically harder than the bispectrum, and, furthermore, computation of the full

Fisher matrix is usually faster than the ideal equivalent, since the latter requires a summation over $\mathcal{O}(\ell_{max}^6)$ individual multipoles (including correlations between L bins).

The trispectrum runtimes are enumerated in Table III. Our conclusions are similar to those of the power spectrum and bispectrum; the four-field term of the estimator is fast

TABLE III. Computation times for ideal and unwindowed trispectrum estimators, as in Tables I and II. Here, we use 4 linear bins with $\ell \in [2, 483]$ across 24 combinations of fields and parities, giving 1996 bins (of which we later remove the low- ℓ contributions). Here, the four-field (disconnected plus connected) term of the trispectrum estimator requires ≈ 100 harmonic transforms to define the numerator (from each computation of H and A fields), with the estimation of the disconnected (i.e., Gaussian) terms increasing this by a factor proportional to $N_{it} = 100$ (with a coefficient larger than one, due to the extra permutations present in these terms). The ideal estimator requires *much* more time to compute than the optimal form; this is due to the internal $\mathcal{O}(\ell_{max}^3)$ summations in the two- and zero-field terms, which make this estimator impractical in realistic settings. The Fisher matrix requires 65,000 harmonic transforms per realization (reducing to 13,000 transforms if the mask is uniform), due to computation of the Q_4 derivatives. In this instance, computation of the ideal Fisher matrix is not feasible (even at $\ell_{max} = 483$), due to the internal $\mathcal{O}(\ell_{max}^6)$ summation over Wigner $6j$ symbols. This remains true even if one assumes an (inexact) diagonal approximation, which reduces the sum to $\mathcal{O}(\ell_{max}^5)$.

Mask	Ideal Fisher (day)	Optimal Fisher	Ideal numerator (m)	Optimal numerator [4-field] (s)	Optimal numerator (m)
Uniform	>7	22 m \times 100	705	15	78
Galactic	>7	36 m \times 100	708	14	92

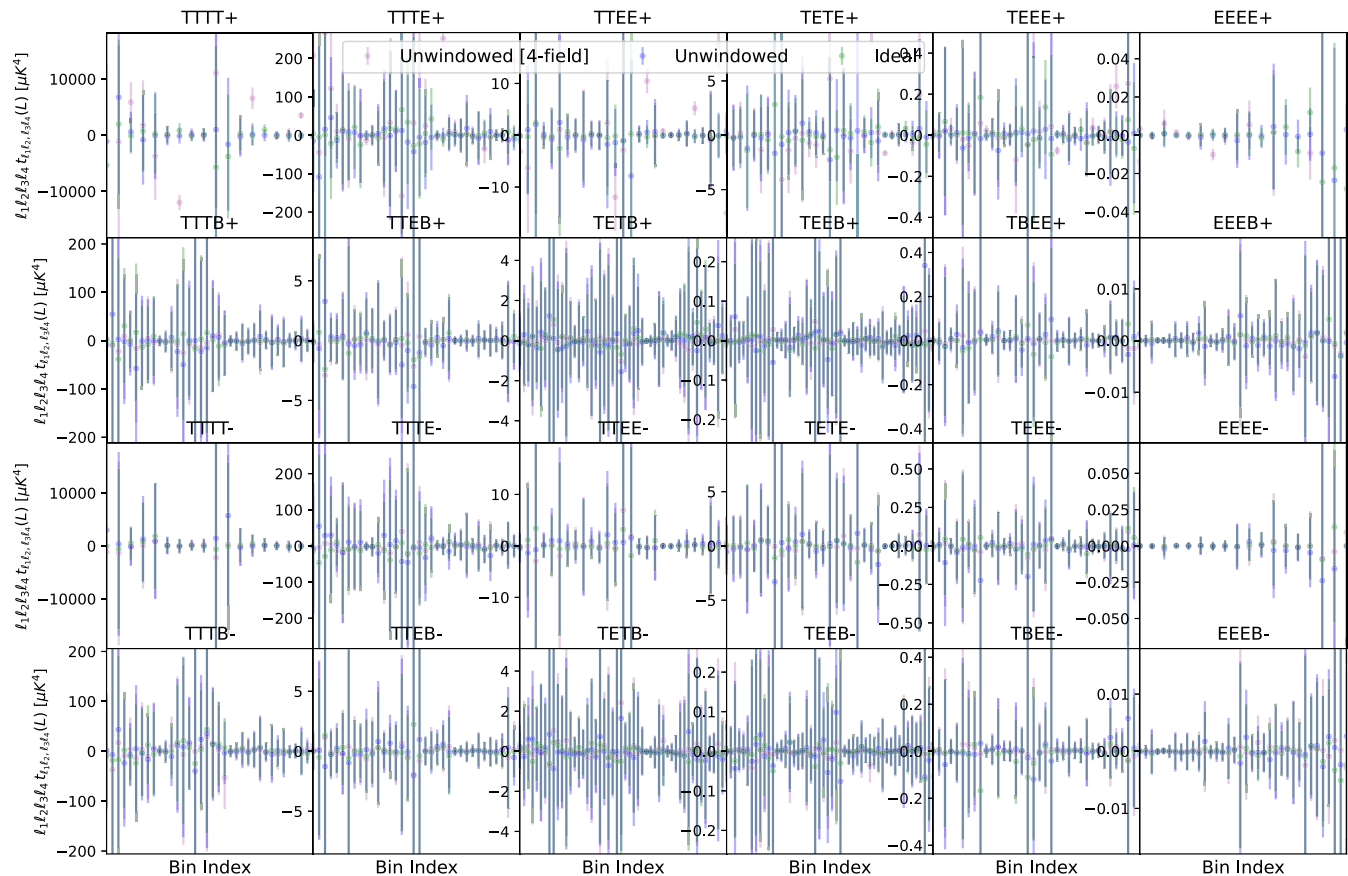


FIG. 6. Comparison of measured and theoretical trispectra from a suite of masked simulations, across a wide range of bins, fields, and parities. Since no trispectrum is injected, the mean of the estimators should be zero; the null results in the figure show this holds true in practice. Strong detections of the disconnected trispectrum can be seen in the four-field term; further analysis of these contributions is shown in Fig. 9. Some parity-odd bins appear to have very large variances; these correspond to configurations that vanish in the exact thin-bin limit due to the restriction to odd $\ell_1 + \ell_2 + \ell_3 + \ell_4$. These were excluded from previous work [52], but included here for full generality.

to compute, with a factor of N_{it} slow-down when we add the two- and zero-field terms. In contrast to the bispectrum, these terms *must* be present in the estimator if $\chi p_u = 1$; else, the trispectrum is dominated by the Gaussian power spectrum contributions. This additionally implies that N_{it} must be relatively large and the fiducial spectrum must be close to the truth, else residual Gaussian pieces will dominate the error (though this is second order in $C_\ell^{\text{sim}} - C_\ell^{\text{true}}$). To compute such terms with the idealized trispectrum estimator, an $\mathcal{O}(\ell_{\text{max}}^3)$ summation over Wigner $3j$ symbols is required; in practice, this is computationally expensive (taking 10 hours in this example), thus the ideal form is unlikely to be of practical use. For our choice of binning, the optimal trispectrum Fisher matrix requires only $\approx 2\times$ the computation time of the bispectrum result; this is well within the capabilities of most clusters. As before, this is usually limited by memory; at $N_{\text{side}} = 256$ with 1996 bins (before removal of low- ℓ bins), a total of 210 GB of data products is stored in memory for efficient

computation. On a high-memory node, computations requiring tenfold more memory would also be practical; beyond this, one must compute the outer products element-wise, leading to an algorithm with quadratic complexity, as for the bispectrum. Finally, we note that the ideal trispectrum Fisher matrix is impractical to compute, even for relatively low ℓ_{max} . With our choice of binning, neither the ideal Fisher matrix nor its diagonal limit could be computed within a few node-days; this further supports the windowed estimators, which avoid having to perform such a sum.

Figures 6 and 7 display the trispectra estimated from our GRF simulations and their variances. In all cases, we find trispectra consistent with zero; this is as expected, and implies that the subtraction of the disconnected trispectrum components is proceeding correctly. Here, we find that $N_{\text{it}} = 100$ is sufficient to recover unbiased trispectra, noting that the error bar on the two- and zero-field contributions scales as $N_{\text{it}}^{-1/2}$. Turning to the variances, we again find that the unwinded estimates are in

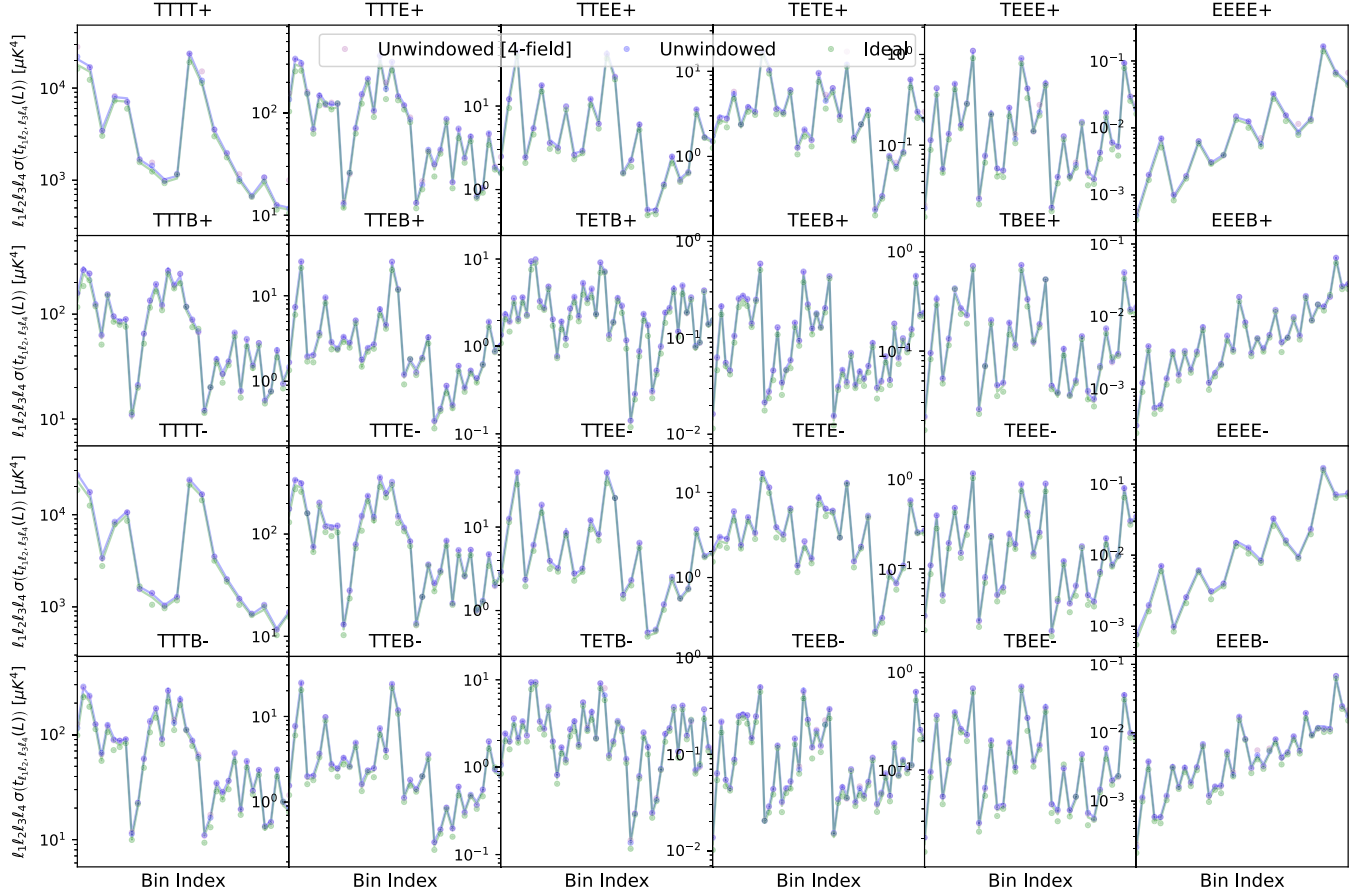


FIG. 7. As Fig. 6, but displaying the theoretical and empirical trispectrum variances. We find excellent agreement across the range of scales considered ($5 \leq \ell < 483$), implying that the unwinded forms are approximately minimum variance, and thus close to optimal.

excellent agreement with predictions from the inverted Fisher matrix, across all scales considered. As discussed above, the ideal Fisher matrices cannot be computed explicitly; here, we obtain them by applying the optimal pipeline without a window function (which is exact, in the limit of infinite Monte Carlo simulations). These are in fair agreement with the empirical variances from the ideal estimators (when weighting by $\langle W^8 \rangle / \langle W^4 \rangle^2$), though we again stress that this relation is only expected to be approximate. For a trivial mask, the ideal and optimal estimator are in excellent agreement, as expected.

The correlation structure is considered in detail in Fig. 8. We find significant covariance at $\mathcal{O}(10\%)$ between different combinations of fields, principally sourced by the intrinsic TE , TB , and EB power spectra in our testing suite. Spectra with $\chi = 1$ and $\chi = -1$ are largely uncorrelated, but we observe highly nontrivial patterns within each trispectrum configuration. As for the scalar case [52], this is partly due to an internal degeneracy in the trispectrum definition due to the two possible tetrahedron diagonals;

further correlations also arise from the masking, which correlates adjacent-in- ℓ bins (which are not always adjacent in one-dimensional projections). We find no evidence for deviation of the inverse Fisher matrix from the empirical correlations; furthermore, the complex structure of this matrix is important to take into account any studies requiring the statistical properties of trispectra. Finally, we note that our Fisher matrices are highly converged; using only $N_{\text{it}} = 10$ simulations to define the Fisher matrix changes the trispectra by at most 0.018σ , thus our choice is overly conservative.

Whilst the above results demonstrate that the trispectrum estimators do not induce spurious signals, they do not sufficiently demonstrate that our pipeline is unbiased (though agreement of empirical and theoretical variances is itself a strong constraint). To further test the approach, we consider the individual components of the trispectrum applied to unwinded GRF data. In this regime, we have a definite prediction of the four-, two-, and zero-field components, with $\mathbb{E}[\mathcal{F}_4 \hat{t}_0] = \mathbb{E}[\mathcal{F}_4 \hat{t}_4] = -(1/2)\mathbb{E}[\mathcal{F}_4 \hat{t}_2]$, and

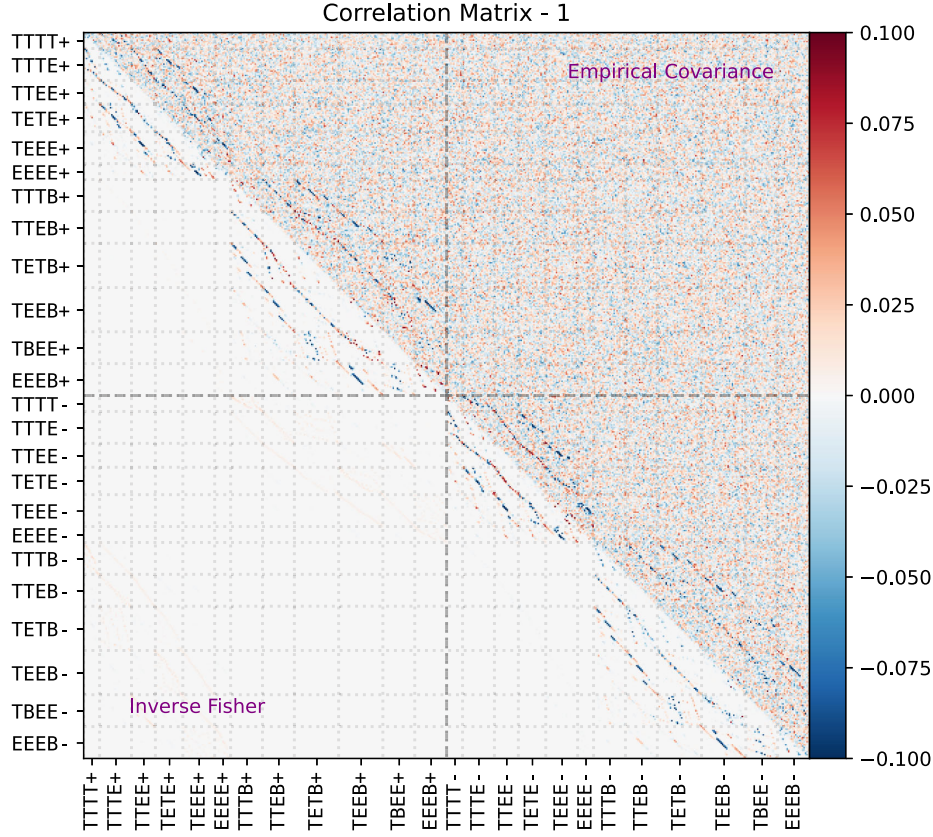


FIG. 8. As Fig. 5, but comparing the correlation matrix of the unwindowed trispectrum measurements to that predicted from the inverse Fisher matrix shown in Figs. 6 and 7. Although the empirical data (obtained from 1000 mocks) is somewhat noisy, we find excellent agreement for all scales and parities considered.

$$\begin{aligned} \mathbb{E}[\mathcal{F}_4 \hat{t}_0]_X^u(\mathbf{b}, B) &= \frac{\delta_{\chi^2 p_{u-1}}^K}{\Delta_4^u(\mathbf{b})} \sum_L \Theta_L(B) \Theta_{\ell_1}(b_1) \Theta_{\ell_2}(b_2) \begin{pmatrix} \ell_1 & \ell_2 & L \\ -1 & -1 & 2 \end{pmatrix}^2 \frac{(2\ell_1 + 1)(2\ell_2 + 1)(2L + 1)}{4\pi} \\ &\times (-1)^{\ell_1 + \ell_2 + L} \left(\delta_{b_1 b_3}^K \delta_{b_2 b_4}^K C_{\ell_1}^{-1, u_1 u_3} C_{\ell_2}^{-1, u_2 u_4} + \delta_{b_1 b_4}^K \delta_{b_2 b_3}^K C_{\ell_1}^{-1, u_1 u_4} C_{\ell_2}^{-1, u_2 u_3} \right), \end{aligned} \quad (112)$$

upon realization averaging. This is nonzero only for particular configurations, in particular pairs of equal bins with nonvanishing two-point correlators. Figure 9 displays all such contributions relative to the theoretical model, and, in every case, we find good agreement, which implies that the sum of the three contributions is consistent with zero (as previously noted). The dominant source of noise here is the zero-field (data-independent term); for a single simulation, this is relatively small compared to the scatter in the four-field term, but contributes an additional source of variance that should be included in any likelihood analysis. Since the parity-breaking spectra not shown above differ only by a factor of $(-1)^{\ell_1 + \ell_2 + \ell_3 + \ell_4}$, this test should be sufficient to convince us that our pipeline is unbiased.

VIII. SUMMARY

In this work, we have presented minimum-variance estimators for the polyspectra of scalar and tensor fields

defined on the two-sphere, derived by extremizing the theoretical likelihood of the dataset. We have considered both the general forms (with an arbitrary weighting scheme \mathbf{S}^{-1}) and the idealized limits, and, in all cases, given practical algorithms and efficient code for their computation. Such approaches can be used to measure binned correlators for a range of spin-0 and spin-2 fields, and take into account the effects of leakage between different bins, polarizations, and parities, due to effects such as intrinsic correlations, survey masks, and inpainting. All estimators have been verified using suites of Gaussian and non-Gaussian simulations, and are found to yield unbiased results, with variances approaching the optimal limits. Furthermore, they can be efficiently computed using spin-weighted spherical harmonic transforms and Monte Carlo summation [with normalization matrices requiring only $\mathcal{O}(10)$ realizations], with the higher-point optimal estimators becoming significantly more practical to implement than their idealized equivalents.

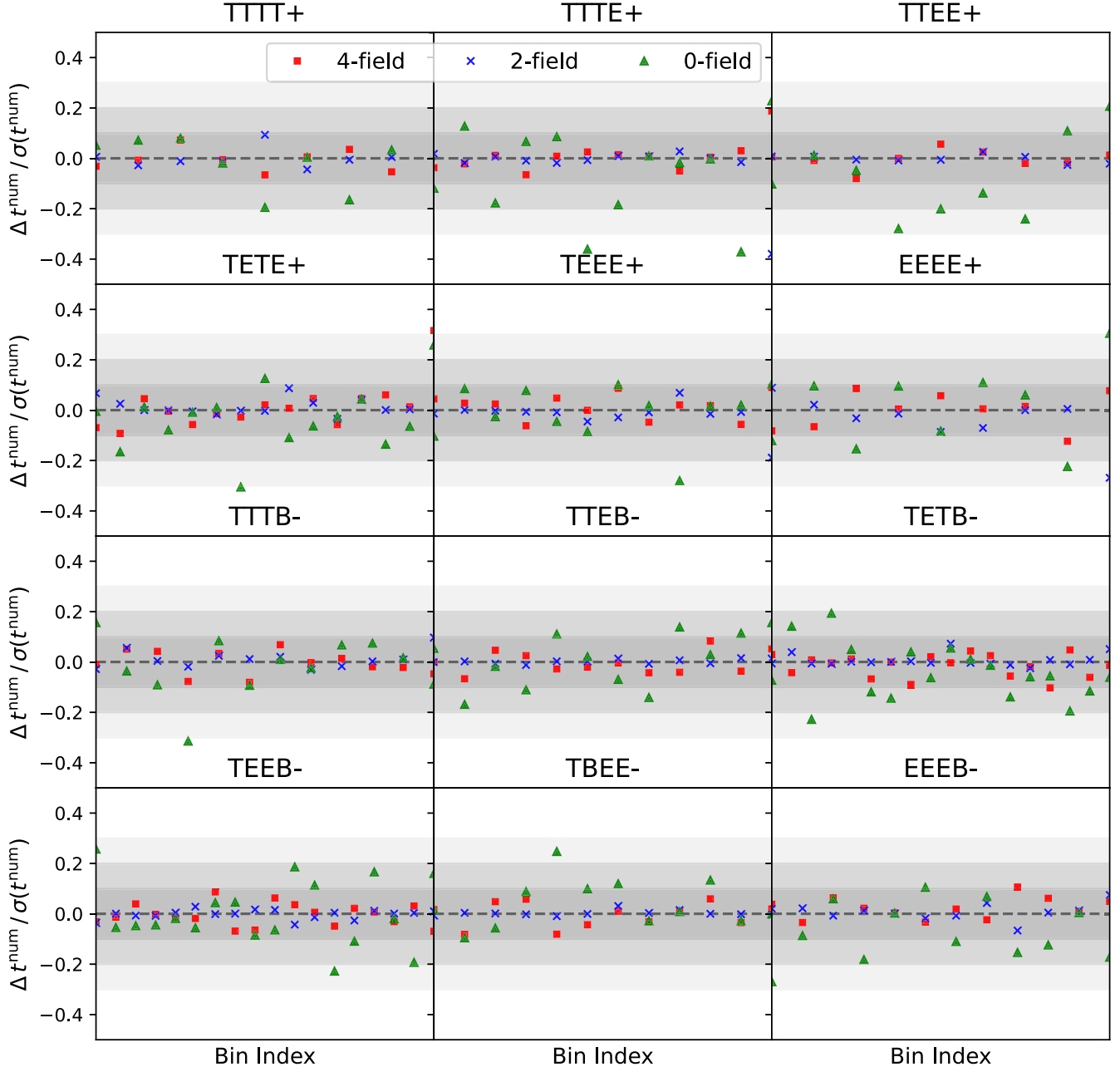


FIG. 9. Comparison of the four-, two-, and zero-field terms entering the unwindowed trispectrum estimator, averaged over 1000 GRFs with a uniform mask. In each case, we compare the measured numerator to the theoretical prediction given in (112), normalizing by the standard deviation of the full statistic. Results are displayed for all combinations of fields and bins where the disconnected term is nonzero. In all cases, we find good agreement though note that the zero-field term exhibits largest variation since this is an average over only $N_{it} = 100$ simulations, as opposed to the 1000 used in the four-field term. This serves to validate that the trispectrum estimators are unbiased (noting that parity-odd estimators differ only by a sign, and the full estimator is the sum of the above components).

It is interesting to consider potential applications of the above techniques. A particularly promising avenue is the investigation of inflationary higher-point functions, as traced by large-scale CMB temperature and polarization anisotropies. Whilst some of this work has already been performed in the context of *Planck* e.g., [7,16,43,60,76–78], many further studies are possible, and allow for novel connections to

theoretical disciplines such as the “Cosmological Collider” program, through the empirical search for signatures of primordial particle interactions e.g., [50]. Much of this work will benefit from the addition of the trispectrum, in particular its polarized components, which to our knowledge, have not been previously been considered. For the *Planck* bispectrum, addition of polarization modes to the temperature three-point

function was found to roughly double the sensitivity to primordial physics, as evidence through bounds on the non-Gaussianity parameter f_{NL} [7]; a similar result can likely be obtained for the associated parameters g_{NL} (and possibly τ_{NL}) arising in the four-point function. Measurements of the parity-odd E mode trispectrum (and its combination with T , e.g., $TTTE$), would bolster the sensitivity to sources of primordial parity-violation, improving upon the results of [60,66]. The B modes could also be of use, and would allow connection between inflationary gravitational wave theory and observations e.g., [79–81].

Though this paper has been framed in the language of CMB analyses, this is not a restriction, and the technology contained within can be similarly applied to any other spin-zero and/or spin-two field defined on the two-sphere. Another interesting cosmological candidate is the measurement of higher-order cosmic shear correlators, and their connection to the spin-zero galaxy overdensity. Historically, most analysis of higher-order shear correlators has been performed in position space, rather than harmonic space; the techniques developed in this work allow for efficient computation of the shear bispectra and

beyond, in a manner that can be simply connected to observations. This obviates common problems such as the difficulty of convolving theory models with masks for statistics beyond the power spectrum, and should facilitate tighter constraints to be wrought on cosmological and astrophysical parameters, in combination with conventional two-point analysis. Finally, we hope that applications may be found beyond cosmology; many other fields consider stochastic spherical data, such as climate science and geology, and could potentially benefit from such techniques.

ACKNOWLEDGMENTS

We thank Adri Duivenvoorden, Maresuke Shiraishi, and Masahiro Takada for insightful discussions. O. H. E. P. is a Junior Fellow of the Simons Society of Fellows and thanks Kavli IPMU for hosting a visit during which this work was started. The author is supported calorically by 7/11 (Japanese edition) and thanks the hospitality team at the Simons Foundation for demonstrating that there is such a thing as a free lunch.

-
- [1] D. N. Spergel and M. Zaldarriaga, *Phys. Rev. Lett.* **79**, 2180 (1997).
 - [2] U. Seljak and M. Zaldarriaga, *Phys. Rev. Lett.* **78**, 2054 (1997).
 - [3] M. Bartelmann and P. Schneider, *Phys. Rep.* **340**, 291 (2001).
 - [4] M. Zaldarriaga and R. Scoccimarro, *Astrophys. J.* **584**, 559 (2003).
 - [5] M. Kamionkowski, A. Kosowsky, and A. Stebbins, *Phys. Rev. D* **55**, 7368 (1997).
 - [6] Y. Akrami *et al.* (Planck Collaboration), *Astron. Astrophys.* **641**, A10 (2020).
 - [7] Y. Akrami *et al.* (Planck Collaboration), *Astron. Astrophys.* **641**, A9 (2020).
 - [8] N. Aghanim, Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. J. Banday, R. B. Barreiro, N. Bartolo *et al.* (Planck Collaboration), *Astron. Astrophys.* **641**, A6 (2020).
 - [9] A. Lewis and A. Challinor, *Phys. Rep.* **429**, 1 (2006).
 - [10] H. Hildebrandt *et al.*, *Mon. Not. R. Astron. Soc.* **465**, 1454 (2017).
 - [11] T. M. C. Abbott *et al.* (DES Collaboration), *Phys. Rev. D* **105**, 023520 (2022).
 - [12] R. Dalal *et al.*, [arXiv:2304.00701](https://arxiv.org/abs/2304.00701).
 - [13] P. Motloch, U.-L. Pen, and H.-R. Yu, *Phys. Rev. D* **105**, 083512 (2022).
 - [14] P. J. E. Peebles, *The Large-Scale Structure of the Universe* (Princeton University Press, Princeton, NJ, 1980).
 - [15] N. Aghanim *et al.* (Planck Collaboration), *Astron. Astrophys.* **641**, A8 (2020).
 - [16] P. Creminelli, A. Nicolis, L. Senatore, M. Tegmark, and M. Zaldarriaga, *J. Cosmol. Astropart. Phys.* **05** (2006) 004.
 - [17] J. R. Fergusson, D. M. Regan, and E. P. S. Shellard, [arXiv:1012.6039](https://arxiv.org/abs/1012.6039).
 - [18] J. R. Fergusson, *Phys. Rev. D* **90**, 043533 (2014).
 - [19] J. R. Fergusson and E. P. S. Shellard, *Phys. Rev. D* **80**, 043510 (2009).
 - [20] J. R. Fergusson, M. Liguori, and E. P. S. Shellard, *J. Cosmol. Astropart. Phys.* **12** (2012) 032.
 - [21] J. Smidt, A. Amblard, C. T. Byrnes, A. Cooray, A. Heavens, and D. Munshi, *Phys. Rev. D* **81**, 123007 (2010).
 - [22] F. Bernardeau, C. Bonvin, N. Van de Rijdt, and F. Vernizzi, *Phys. Rev. D* **86**, 023001 (2012).
 - [23] S. Dodelson and P. Zhang, *Phys. Rev. D* **72**, 083001 (2005).
 - [24] M. Rizzato, K. Benabed, F. Bernardeau, and F. Lacasa, *Mon. Not. R. Astron. Soc.* **490**, 4688 (2019).
 - [25] P. Schneider and M. Lombardi, *Astron. Astrophys.* **397**, 809 (2003).
 - [26] R. Scoccimarro, H. A. Feldman, J. N. Fry, and J. A. Frieman, *Astrophys. J.* **546**, 652 (2001).
 - [27] H. Gil-Marín, J. Noreña, L. Verde, W. J. Percival, C. Wagner, M. Manera, and D. P. Schneider, *Mon. Not. R. Astron. Soc.* **451**, 539 (2015).
 - [28] D. W. Pearson and L. Samushia, *Mon. Not. R. Astron. Soc.* **478**, 4500 (2018).
 - [29] G. Cabass, M. M. Ivanov, and O. H. E. Philcox, *Phys. Rev. D* **107**, 023523 (2023).
 - [30] G. Cabass, M. M. Ivanov, O. H. E. Philcox, M. Simonović, and M. Zaldarriaga, *Phys. Rev. Lett.* **129**, 021301 (2022).

- [31] G. Cabass, M. M. Ivanov, O. H. E. Philcox, M. Simonović, and M. Zaldarriaga, *Phys. Rev. D* **106**, 043506 (2022).
- [32] O. H. E. Philcox and M. M. Ivanov, *Phys. Rev. D* **105**, 043517 (2022).
- [33] M. M. Ivanov, O. H. E. Philcox, G. Cabass, T. Nishimichi, M. Simonović, and M. Zaldarriaga, *Phys. Rev. D* **107**, 083515 (2023).
- [34] O. H. E. Philcox, *Phys. Rev. D* **106**, 063501 (2022).
- [35] J. Hou, Z. Slepian, and R. N. Cahn, *Mon. Not. R. Astron. Soc.* **522**, 5701 (2023).
- [36] O. H. E. Philcox, J. Hou, and Z. Slepian, [arXiv:2108.01670](https://arxiv.org/abs/2108.01670).
- [37] G. d'Amico, J. Gleyzes, N. Kokron, K. Markovic, L. Senatore, P. Zhang, F. Beutler, and H. Gil-Marín, *J. Cosmol. Astropart. Phys.* **05** (2020) 005.
- [38] E. Hivon, K. M. Gorski, C. B. Netterfield, B. P. Crill, S. Prunet, and F. Hansen, *Astrophys. J.* **567**, 2 (2002).
- [39] A. Gangui and J. Martin, *Phys. Rev. D* **62**, 103004 (2000).
- [40] D. Munshi, A. Heavens, A. Cooray, J. Smidt, P. Coles, and P. Serra, *Mon. Not. R. Astron. Soc.* **412**, 1993 (2011).
- [41] M. Bucher, B. Racine, and B. van Tent, *J. Cosmol. Astropart. Phys.* **05** (2016) 055.
- [42] A. P. S. Yadav, E. Komatsu, B. D. Wandelt, M. Liguori, F. K. Hansen, and S. Matarrese, *Astrophys. J.* **678**, 578 (2008).
- [43] A. J. Duivendoorn, P. D. Meerburg, and K. Freese, *Phys. Rev. D* **102**, 023521 (2020).
- [44] H. F. Guetjen, J. R. Fergusson, M. Liguori, and E. P. S. Shellard, *Phys. Rev. D* **95**, 043532 (2017).
- [45] D. Munshi, A. Heavens, A. Cooray, J. Smidt, P. Coles, and P. Serra, *Mon. Not. R. Astron. Soc.* **412**, 1993 (2011).
- [46] M. Shiraishi, M. Liguori, and J. R. Fergusson, *J. Cosmol. Astropart. Phys.* **05** (2014) 008.
- [47] K. M. Smith, L. Senatore, and M. Zaldarriaga, [arXiv:1502.00635](https://arxiv.org/abs/1502.00635).
- [48] K. M. Smith and M. Zaldarriaga, *Mon. Not. R. Astron. Soc.* **417**, 2 (2011).
- [49] M. Kamionkowski, T. L. Smith, and A. Heavens, *Phys. Rev. D* **83**, 023007 (2011).
- [50] N. Arkani-Hamed and J. Maldacena, [arXiv:1503.08043](https://arxiv.org/abs/1503.08043).
- [51] W. Sohn, J. R. Fergusson, and E. P. S. Shellard, [arXiv:2305.14646](https://arxiv.org/abs/2305.14646).
- [52] O. H. E. Philcox, *Phys. Rev. D* **107**, 123516 (2023).
- [53] M. Tegmark, *Phys. Rev. D* **55**, 5895 (1997).
- [54] M. Tegmark, A. J. S. Hamilton, M. A. Strauss, M. S. Vogeley, and A. S. Szalay, *Astrophys. J.* **499**, 555 (1998).
- [55] A. J. S. Hamilton, *Lect. Notes Phys.* **665**, 415 (2008).
- [56] A. J. S. Hamilton, *Lect. Notes Phys.* **665**, 433 (2008).
- [57] A. J. S. Hamilton and M. Tegmark, *Mon. Not. R. Astron. Soc.* **312**, 285 (2000).
- [58] O. H. E. Philcox, *Phys. Rev. D* **103**, 103504 (2021).
- [59] O. H. E. Philcox, *Phys. Rev. D* **104**, 123529 (2021).
- [60] O. H. E. Philcox, [arXiv:2303.12106](https://arxiv.org/abs/2303.12106).
- [61] M. Shiraishi, *Phys. Rev. D* **94**, 083503 (2016).
- [62] C. Creque-Sarbinowski, S. Alexander, M. Kamionkowski, and O. Philcox, [arXiv:2303.04815](https://arxiv.org/abs/2303.04815).
- [63] V. Gluscevic and M. Kamionkowski, *Phys. Rev. D* **81**, 123529 (2010).
- [64] N. Bartolo, S. Matarrese, M. Peloso, and M. Shiraishi, *J. Cosmol. Astropart. Phys.* **01** (2015) 027.
- [65] A. Lue, L.-M. Wang, and M. Kamionkowski, *Phys. Rev. Lett.* **83**, 1506 (1999).
- [66] G. Cabass, S. Jazayeri, E. Pajer, and D. Stefanyszyn, *J. High Energy Phys.* **02** (2023) 021.
- [67] O. H. E. Philcox, PolyBin: Binned polyspectrum estimation on the full sky, Astrophysics Source Code Library, record [ascl:2307.020](https://ui.adsabs.org/abs/2307.020) (2023).
- [68] W. R. Coulton, *Phys. Rev. D* **104**, 103527 (2021).
- [69] W. R. Coulton and D. N. Spergel, *J. Cosmol. Astropart. Phys.* **10** (2019) 056.
- [70] K. M. Górski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelman, *Astrophys. J.* **622**, 759 (2005).
- [71] M. Reinecke and D. S. Seljebotn, *Astron. Astrophys.* **554**, A112 (2013).
- [72] D. Blas, J. Lesgourgues, and T. Tram, *J. Cosmol. Astropart. Phys.* **07** (2011) 034.
- [73] W. Hu and T. Okamoto, *Astrophys. J.* **574**, 566 (2002).
- [74] Y. Akrami *et al.* (Planck Collaboration), *Astron. Astrophys.* **641**, A4 (2020).
- [75] A. Kalaja, P. D. Meerburg, G. L. Pimentel, and W. R. Coulton, *J. Cosmol. Astropart. Phys.* **04** (2021) 050.
- [76] P. A. R. Ade *et al.* (Planck Collaboration), *Astron. Astrophys.* **594**, A17 (2016).
- [77] E. Komatsu, D. N. Spergel, and B. D. Wandelt, *Astrophys. J.* **634**, 14 (2005).
- [78] P. A. R. Ade, N. Aghanim, C. Armitage-Caplan, M. Arnaud, M. Ashdown, F. Atrio-Barandela, J. Aumont, C. Baccigalupi, A. J. Banday *et al.* (Planck Collaboration), *Astron. Astrophys.* **571**, A24 (2014).
- [79] J. Bonifacio, H. Goodhew, A. Joyce, E. Pajer, and D. Stefanyszyn, *J. High Energy Phys.* **06** (2023) 212.
- [80] G. Cabass, E. Pajer, D. Stefanyszyn, and J. Supel, *J. High Energy Phys.* **05** (2022) 077.
- [81] E. Pajer, D. Stefanyszyn, and J. Supel, *J. High Energy Phys.* **12** (2020) 198; **04** (2022) 23(E).