

A unified pseudo- C_ℓ framework

David Alonso,^{1,2★} Javier Sanchez³ and Anže Slosar⁴

(The LSST Dark Energy Science Collaboration)

¹*School of Physics and Astronomy, Cardiff University, The Parade, Cardiff CF24 3AA, UK*

²*Department of Physics, University of Oxford, Denys Wilkinson Building, Keble Road, Oxford OX1 3RH, UK*

³*Department of Physics and Astronomy, University of California, Irvine, CA 92697, USA*

⁴*Brookhaven National Laboratory, Physics Department, Upton, NY 11973, USA*

Accepted 2019 January 1. Received 2018 December 24; in original form 2018 September 24

ABSTRACT

The pseudo- C_ℓ is an algorithm for estimating the angular power and cross-power spectra that is very fast and in realistic cases also nearly optimal. The algorithm can be extended to deal with contaminant deprojection and E/B purification, and can therefore be applied in a wide variety of scenarios of interest for current and future cosmological observations. This paper presents NAMASTER, a public, validated, accurate, and easy-to-use software package that, for the first time, provides a unified framework to compute angular cross-power spectra of any pair of spin-0 or spin-2 fields, contaminated by an arbitrary number of linear systematics and requiring B - or E -mode purification, both on the sphere or in the flat-sky approximation. We describe the mathematical background of the estimator, including all the features above, and its software implementation in NAMASTER. We construct a validation suite that aims to resemble the types of observations that next-generation large-scale structure and ground-based cosmic microwave background experiments will face, and use it to show that the code is able to recover the input power spectra in the most complex scenarios with no detectable bias. NAMASTER can be found at <https://github.com/LSSTDESC/NaMaster>, and is provided with comprehensive documentation and a number of code examples.

Key words: methods: data analysis – methods: numerical.

1 INTRODUCTION

Two-point functions have proven to be the most useful summary statistic for cosmology in terms of data compression, due to the near-Gaussian nature of the modes of the cosmic density field that can be easily analysed to extract cosmological information (Gorski 1994; Bond 1995; Tegmark 1997; Tegmark & de Oliveira-Costa 2001). Since the only true observables in any astronomical data set are the intensity and polarization as a function of frequency ν and sky position $\hat{\theta}$, angular two-point correlators play a particularly central role in most cosmological analyses, both in the form of configuration-space two-point correlation functions $\xi(\theta)$ and as Fourier-space¹ power spectra C_ℓ (Asorey et al. 2012; Bonvin & Durrer 2011). Although correlation functions are often easier to estimate from complex observations (Landy & Szalay 1993), a Fourier space analysis is advantageous for many other reasons. The

most important is that Fourier space is a natural space for working with statistics of translation-invariant fields. Therefore, Fourier space analysis has a clearer separation between linear and non-linear scales, as well as Gaussian and non-Gaussian modes (Hamimeche & Lewis 2008). Moreover, individual bandpower measurements are typically significantly less correlated compared to configuration-space analysis which makes covariance matrix estimation easier. This has motivated a large body of literature aimed at designing optimal and efficient power-spectrum estimators and used extensively in cosmic microwavebackground (CMB) two-point measurements.

While maximum-likelihood (Bond, Jaffe & Knox 1998; Wandelt & Hansen 2003) or minimum-variance and quadratic estimators (Tegmark 1997; Hamilton, Tegmark & Padmanabhan 2000; Tegmark & de Oliveira-Costa 2001; Vanneste et al. 2018) can recover the power spectrum with virtually no loss of information (see also the approaches of e.g. Eriksen et al. 2004; Taylor, Ashdown & Hobson 2008; Alsing et al. 2016, based on Bayesian sampling methods), their computational implementation is prohibitively expensive for high-resolution data, since these methods scale as ℓ_{\max}^6 with the maximum multipole ℓ_{\max} (or ℓ_{\max}^4 for optimal algorithms such as that of Wandelt & Hansen 2003). For this reason, pseudo- C_ℓ algorithms (Peebles 1973; Wandelt, Hivon & Górski 2001; Szapudi,

* E-mail: David.Alonso@physics.ox.ac.uk

¹In this paper, we use the term ‘Fourier’ to describe both the flat space Fourier expansion as well as harmonic-space expansion on the sphere. We will also use the term ‘power spectrum’ to describe both the auto-power spectrum of a field and cross-power spectra between two fields.

Prunet & Colombi 2001; Wandelt et al. 2001; Hivon et al. 2002; Hansen, Górski & Hivon 2002; Chon et al. 2004; Tristram et al. 2005) have become a popular alternative that preserves the ℓ_{\max}^3 scaling of spherical harmonic transforms (SHTs, and faster in the flat-sky approximation). This method can be used on fields with arbitrary spin (Kogut et al. 2003) in both curved and flat skies, has been extended to deal with contaminant deprojection in scalar fields (Elsner, Leistedt & Peiris 2017), and can be optimized to measure the CMB E/B -mode power spectrum (Lewis, Challinor & Turok 2002; Bunn et al. 2003; Smith 2006; Grain, Tristram & Stompór 2009; Bunn 2011). Due to the advantages of this method, a number of useful, publicly available codes currently exist to carry out some of these calculations.² Unfortunately, no public, validated, and maintained code exists simultaneously supporting all of the capabilities described above in a consistent manner. This paper presents a public software package, NAMASTER, that unifies all of these features providing a general framework to estimate pseudo- C_ℓ angular power spectra in a wide variety of scenarios. The paper also presents a few novel results that were derived as part of the development of NAMASTER. These include extending the contaminant deprojection formalism to spin-2 fields, combining it with E/B purification and providing a consistent derivation of the same functionality in the flat-sky approximation.

The paper is structured as follows: Section 2 introduces the pseudo- C_ℓ estimator in the curved sky, including mode deprojection and E/B purification. This is then generalized in the flat-sky approximation. In Section 3, we briefly describe the implementation of these methods on NAMASTER, as well as its structure and usage. Section 4 presents the stress tests used to validate the code. We envisage this code to be useful in analysis of both the CMB data and in the tomographic large-scale structure (LSS) from photometric data, so this section illustrates not just its validity but also applicability in typical user-case scenarios in both fields. We conclude in Section 5. The appendices provide further details about the mathematical background used throughout the paper, and present some additional validation tests.

2 THE PSEUDO- C_ℓ ESTIMATOR

Here, we introduce the pseudo- C_ℓ estimator and its relevant extensions. Many of the results presented here can be found in the literature (e.g. Hivon et al. 2002; Kogut et al. 2003; Smith 2006; Grain et al. 2009; Reinecke & Seljebotn 2013; Elsner et al. 2017, and references therein), although some others are new as far as we are aware. In what follows, we will use boldface symbols for vector fields (e.g. \mathbf{a}) and sans-serif symbols for matrices (e.g. \mathbf{C}).

2.1 Spherical harmonic transforms and pseudo- C_ℓ s

Let $\mathbf{a}(\hat{\theta})$ be a spin- s_a quantity defined on the sphere, where $\hat{\theta}$ is the unit vector pointing in a particular direction (θ, φ). We define its spherical harmonic coefficients through a SHT as:

$$\mathbf{a}_{\ell m} \equiv \mathcal{S} \left[\mathbf{a}(\hat{\theta}) \right]_{\ell m}^{s_a} \equiv \int d\hat{\theta} Y_{\ell m}^{s_a \dagger}(\hat{\theta}) \mathbf{a}(\hat{\theta}), \quad (1)$$

$$\mathbf{a}(\hat{\theta}) = \mathcal{S}^{-1} [\mathbf{a}_{\ell m}]_{\hat{\theta}}^{s_a} \equiv \sum_{\ell m} Y_{\ell m}^{s_a}(\hat{\theta}) \mathbf{a}_{\ell m}. \quad (2)$$

Note that, in general we will use vector notation such that, for a complex spin- s_a field a , we form the vector $\mathbf{a} \equiv (\text{Re}(a), \text{Im}(a))$. The

harmonic coefficients above are decomposed in a similar manner into E and B modes: $\mathbf{a}_{\ell m} \equiv (a_{\ell m}^E, a_{\ell m}^B)$ (e.g. see Appendix A and Zaldarriaga & Seljak 1997). The matrices $Y_{\ell m}^s$ are defined in terms of the spin-weighted spherical harmonics, and are described in detail in Appendix A. Some technical details regarding bandlimits and the practicalities of SHTs in a discretized sphere are discussed in Appendix A1. To simplify the notation, we will also often abbreviate the pair (ℓ, m) as \mathbf{l} .

For two isotropic fields \mathbf{a} and \mathbf{b} , the power spectrum is given by their covariance matrix:

$$\langle \mathbf{a}_\mathbf{l} \mathbf{b}_\mathbf{l}^\dagger \rangle \equiv C_\ell^{ab} \delta_{\ell\ell'} \delta_{mm'}. \quad (3)$$

Note that for a general non-zero spin, vectors \mathbf{a} and \mathbf{b} have two-components, each of which is a complex number. Nevertheless the rotational invariance requires that elements of the 2×2 matrix C_ℓ be real.

2.1.1 Power-spectrum estimation

The problem we are attempting to solve is how to estimate C_ℓ^{ab} from a single realization of \mathbf{a} and \mathbf{b} measured on a cut sky. In general, the input data for measuring the power spectrum come pixelized on a sphere. Instead of dealing with an infinitely sampled field, we have a finite number of measurements

$$\mathbf{a}_i = \mathbf{a}_i^{\text{true}} + \mathbf{n}_i, \quad (4)$$

where $\mathbf{a}_i^{\text{true}}$ is the true underlying field whose power spectrum we would like to estimate and \mathbf{n}_i is the noise vector, which in general can be non-white and non-homogeneous. We will denote the pixel covariance of the noise component as $\tilde{\mathbf{N}}$.³ Note that for the sake of power-spectrum estimation, there is no conceptual difference between inhomogeneous noise and a survey mask: masked areas can be simply thought of as areas of infinite noise. Conversely, when we talk about ‘multiplying by the mask’, we really mean a more general process of dividing by the expected variance field.

The underlying power spectrum in equation (3) means that there is a non-trivial true underlying covariance matrix $\tilde{\mathbf{C}}^{\text{true}} = \langle \mathbf{a}^{\text{true}} \mathbf{a}^{\text{true} \dagger} \rangle$, which in addition to the noise component leads to the actual data covariance $\tilde{\mathbf{C}} = \langle \mathbf{a} \mathbf{a}^\dagger \rangle = \tilde{\mathbf{C}}^{\text{true}} + \tilde{\mathbf{N}}$.

In the standard incarnation of the optimal quadratic estimator (Bond et al. 1998), the measurement is produced on inverse covariance-weighted data, namely

$$\mathbf{F}_{ij} \mathbf{B}_j = \frac{1}{2} \mathbf{a}^\dagger \tilde{\mathbf{C}}^{-1} \mathbf{P}_i \tilde{\mathbf{C}}^{-1} \mathbf{a} - \frac{1}{2} \text{Tr} \left[\tilde{\mathbf{C}}^{-1} \mathbf{P}_i \tilde{\mathbf{C}}^{-1} \mathbf{N} \right], \quad (5)$$

where \mathbf{F} is the Fisher matrix, the covariance matrix is modelled as a linear sum over bandpower parameters B_i , i.e.

$$\tilde{\mathbf{C}} = \sum_i \mathbf{P}_i B_i + \mathbf{N}, \quad (6)$$

(which defines the response matrices \mathbf{P}_i) and the last term in equation (5) corresponds to a bias of the convolved estimates due to the presence of noise. Thus, this optimal estimator corresponds to (i) inverse-variance weighting the data, (ii) calculating the spectrum of the weighted data using the response matrices, and (iii) deconvolving these estimates using a mode-coupling matrix which in this particular case happens to be the same as Fisher matrix. The

³Note that the matrices $\tilde{\mathbf{C}}$ and $\tilde{\mathbf{N}}$ used in this section are the data and noise pixel–pixel covariance matrices, and therefore are different from the power-spectrum matrices, C_ℓ and N_ℓ , used in the rest of the paper (e.g. equation 3).

²See, for instance, PolSpice, Xpol, Xpure, and hyperQube.

main computational challenge when implementing this estimator usually lies in storing the covariance matrix and its inverse (or computing it on the fly), particularly in order to estimate the Fisher matrix. This motivates the pseudo- C_ℓ estimator, which replaces these computationally very intensive steps with numerically more efficient ones at the expense of optimality.

The main improvement is gained by replacing the inverse covariance by its diagonal, assuming that the data are uncorrelated between pixels. The response matrices are also replaced by simple SHTs and multiplication of the corresponding harmonic coefficients, followed by an average over bandpowers. Both of these steps reduce to exact optimal estimator in the limit of uncorrelated data. The resulting estimates are still biased and require a mode-decoupling matrix, which can be calculated as we show in Section 2.1.2.

The pseudo- C_ℓ estimates will be optimal when these steps are close to what an optimal quadratic estimator would do. First, this requires that \tilde{C} be close to diagonal, which is true when either the noise is large and uncorrelated, or when the underlying power spectrum is close to white. Multiplication by the mask mixes modes, and replacement of multiplication by the response matrix with bandpower averaging, makes sense only when the mask is ‘well behaved’, i.e. compact with no high-frequency structure in it. If these conditions are not met, the pseudo- C_ℓ estimator will still be unbiased (by construction), but it will get progressively less optimal. In practical surveys, however, the loss of optimality rarely exceeds 10–20 percent and is typically smaller (Leistedt et al. 2013).

In what follows, we will work with two-fields \mathbf{a} and \mathbf{b} and spell out in detail how to calculate the cross-power spectrum. The auto-power spectrum case is the same (with $\mathbf{b} = \mathbf{a}$), but one must additionally subtract the noise contribution to the estimate. So far, the discussion has been completely general, but from now onwards, we will focus on the power estimation of auto- and cross-power spectra for the spin-0 and spin-2 fields only, as these are the most commonly used in cosmology.

2.1.2 Mode coupling

Let $v(\hat{\theta})$ be a sky mask or weights map for \mathbf{a} , and let us define $\mathbf{a}^v \equiv v(\hat{\theta})\mathbf{a}(\hat{\theta})$. As discussed above, we start by considering a naïve estimator for C_ℓ^{ab} from the masked fields:

$$\text{PCL}_\ell(\mathbf{a}^v, \mathbf{b}^w) \equiv \frac{1}{2\ell + 1} \sum_{m=-\ell}^{\ell} \mathbf{a}_1^v \mathbf{b}_1^{w\dagger}. \quad (7)$$

The incomplete sky coverage couples different ℓ modes, and makes this estimator biased. The pseudo- C_ℓ method is based on computing an analytical prediction for this bias and correcting for it. This is straightforward to do using the results in Appendix A. The harmonic coefficients of a masked field are given by

$$\mathbf{a}_1^v = \sum_{l_1 l_2} D_{l_1 l_2}^{s_a} \mathbf{a}_{l_1} v_{l_2}, \quad (8)$$

where D is defined in equation (A10). Using the statistical isotropy of the unmasked field [equation (3)] and the orthogonality relation equation (A11), we obtain:

$$\text{vec} [\langle \text{PCL}_\ell(\mathbf{a}_1^v, \mathbf{b}_1^w) \rangle] = \sum_{\ell'} M_{\ell\ell'}^{s_a s_b} \cdot \text{vec} [C_\ell^{ab}], \quad (9)$$

where we have defined the vectorization operator (Hamimeche & Lewis 2008)

$$\text{vec} \left[\begin{pmatrix} C^{EE} & C^{EB} \\ C^{BE} & C^{BB} \end{pmatrix} \right] \equiv \begin{pmatrix} C^{EE} \\ C^{EB} \\ C^{BE} \\ C^{BB} \end{pmatrix}, \quad (10)$$

$$\text{vec} \left[\begin{pmatrix} C^{TE} \\ C^{TB} \end{pmatrix} \right] \equiv \begin{pmatrix} C^{TE} \\ C^{TB} \end{pmatrix}, \quad \text{vec} [C^{TT}] = C^{TT}, \quad (11)$$

and T here stands for any spin-0 field.

The mode-coupling matrix $M_{\ell\ell'}$ in equation (9) can be computed in terms of the harmonic coefficients of the two masks. We provide analytical expressions for the mode-coupling matrices for different combinations of spin-0 and spin-2 fields in Appendix D.

2.1.3 Bandpowers

Given the loss of information implicit in masking the originally curved-sky field, it is in general not possible to invert the mode-coupling matrix in equation (9) directly. One solution to this problem is to convolve the theory prediction with the mode-coupling matrix at the likelihood level. A more usual approach is instead to bin the coupled pseudo- C_ℓ into bandpowers. A bandpower q is defined as a set of N_q multipoles $\ell_q \equiv (\ell_q^1, \dots, \ell_q^{N_q})$ and a set of weights defined at those multipoles w_q^ℓ , and normalized such that $\sum_{\ell \in \ell_q} w_q^\ell = 1$. The q th bandpower for the coupled pseudo- C_ℓ is then defined as

$$\text{vec} [\tilde{B}_q^{ab}] \equiv \sum_{\ell \in \ell_q} w_q^\ell \text{vec} [\text{PCL}_\ell(\mathbf{a}_1^v, \mathbf{b}_1^w)]$$

$$\langle \text{vec} [\tilde{B}_q^{ab}] \rangle = \sum_{\ell \in \ell_q} w_q^\ell \sum_{\ell'} M_{\ell\ell'}^{s_a s_b} \text{vec} [C_\ell^{ab}], \quad (12)$$

where in the second line, we have taken the expectation value of the first one. One then proceeds by assuming that the true power spectrum is a stepwise function, taking constant values over the multipoles corresponding to each bandpower: $C_\ell^{ab} \equiv \sum_q B_q^{ab} \Theta(\ell \in \ell_q)$ (where Θ is a binary step function). Inserting this in the equation above, it is then possible to find an unbiased estimator for B_q^{ab} in terms of the binned pseudo-bandpowers \tilde{B}_q^{ab} :

$$\text{vec} [B_q^{ab}] = \sum_{q'} (\mathcal{M}^{s_a s_b})_{qq'}^{-1} \text{vec} [\tilde{B}_{q'}^{ab}], \quad (13)$$

where the binned coupling matrix $\mathcal{M}^{s_a s_b}$ is

$$\mathcal{M}_{qq'}^{s_a s_b} \equiv \sum_{\ell \in \ell_q} \sum_{\ell' \in \ell_{q'}} w_q^\ell M_{\ell\ell'}^{s_a s_b}. \quad (14)$$

Since, in general, the true power spectrum is not a stepwise function, the theoretical prediction must be corrected for the bandpower binning before making any comparison with the data:

$$\text{vec} [B_q^{ab, \text{th}}] = \sum_{\ell} \mathcal{F}_{q\ell}^{s_a s_b} \cdot \text{vec} [C_\ell^{ab, \text{th}}], \quad (15)$$

where the filter $\mathcal{F}_{q\ell}^{s_a s_b}$ is given by

$$\mathcal{F}_{q\ell}^{s_a s_b} = \sum_{q'} (\mathcal{M}^{s_a s_b})_{qq'}^{-1} \sum_{\ell' \in \ell_{q'}} w_{q'}^{\ell'} M_{\ell\ell'}^{s_a s_b}. \quad (16)$$

The action of the filter $\mathcal{F}^{s_a s_b}$ is therefore given by a sequence of three operations:

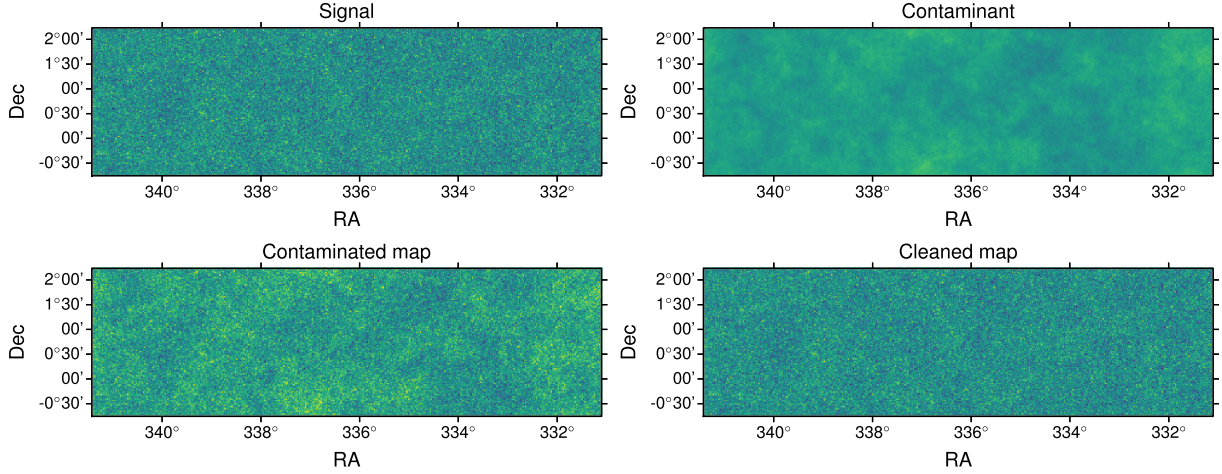


Figure 1. Depiction of the contaminant deprojection process. Top left: underlying true signal map (galaxy overdensity used in Section 4.1.1). Top right: sky contaminant (dust contamination used in Section 4.1.1). Bottom left: contaminated map obtained by adding the previous two maps. Bottom right: cleaned map obtained through equation (21). Since only one contaminant is considered in this case, the cleaning is almost perfect. Residual contamination at the power-spectrum level must still be subtracted in order to avoid a biased estimate.

- (i) Coupling the different multipoles (equation 9).
- (ii) Binning into bandpowers (equation 12).
- (iii) Decoupling the bandpowers (equation 13).

Depending on the structure of the mode-coupling matrix, $\mathcal{F}^{a, sb}$ can be significantly different from a naïve binning operator.

2.2 Contaminant deprojection

In a typical data-analysis case, we deal with contaminants that pollute the measured signal. Usually, we know something about these contaminants. As an example, in case of CMB analysis, we might have a map of a given foreground at a different frequency which can be marginalized out.⁴ But even when we do not, we often have a tracer map of a given effect, for example, the map of dust might tell us something about systematic errors related to imperfect correction to dust reddening, or a map of co-added point spread function (PSF) fluctuations could be used to remove the possible correlations between measured ellipticities. Even if the dust map is non-linearly and perhaps stochastically related to the amount of reddening, as long as that relation is local in space, we get a linear contamination on large scales, very much like complex astrophysics of galaxy formation nevertheless reduces to linear biasing on very large scales (see Appendix B). Thus, we generically deal with a linear model of the kind:

$$\mathbf{a}^{\text{obs}}(\hat{\theta}) = \mathbf{a}^v + \sum_{i=1}^{N_a} \alpha_i \mathbf{f}^i(\hat{\theta}), \quad (17)$$

where the observed map \mathbf{a}^{obs} is composed of a masked signal map that receives contributions from N_a known contaminants \mathbf{f}^i and where the contamination coefficients α_i are unknown (note that we have implicitly applied the same mask to the contaminant templates

\mathbf{f}^i). The deprojection of these contaminants from the data can be done and propagated into the estimated power spectrum in a natural way in the case of quadratic estimators (Rybicki & Press 1992; Slosar, Seljak & Makarov 2004; Elsner, Leistedt & Peiris 2016), and the method can be extended to the pseudo- C_ℓ algorithm (Elsner et al. 2017). The process is illustrated in Fig. 1. A best-fitting value for the coefficients α_i can be found assuming *uncorrelated weights* across the map⁵

$$\hat{\alpha}_i = F_{ij} \int d\hat{\theta} \mathbf{f}^{j\dagger}(\hat{\theta}) \mathbf{a}^{\text{obs}}(\hat{\theta}), \quad (F^{-1})_{ij} \equiv \int d\hat{\theta} \mathbf{f}^{i\dagger}(\hat{\theta}) \mathbf{f}^j(\hat{\theta}), \quad (18)$$

where there is an implicit summation over the repeated index j in equation (18) (which we will omit from now on). The clean map is then given by

$$\mathbf{a}^c(\hat{\theta}) \equiv \mathbf{a}^{\text{obs}}(\hat{\theta}) - \mathbf{f}^i(\hat{\theta}) F_{ij} \int d\hat{\theta}' \mathbf{f}^{j\dagger}(\hat{\theta}') \mathbf{a}^{\text{obs}}(\hat{\theta}') \quad (19)$$

which, when substituting equation (17), yields

$$\mathbf{a}^c(\hat{\theta}) = \mathbf{a}^v(\hat{\theta}) - \mathbf{f}^i(\hat{\theta}) F_{ij} \int d\hat{\theta}' \mathbf{f}^{j\dagger}(\hat{\theta}') \mathbf{a}^v(\hat{\theta}'). \quad (20)$$

The harmonic coefficients of \mathbf{a}^c are therefore given by

$$\mathbf{a}_1^c = \mathbf{a}_1^v - \mathbf{f}_1^i F_{ij} \sum_I \mathbf{f}_I^{j\dagger} \mathbf{a}_I^v \quad (21)$$

Although in general we expect the second term in equation (20) to be very small, since the signal and contaminant maps are uncorrelated, due to this residual statistical contamination, a direct computation of the pseudo- C_ℓ of the cleaned map would yield a biased estimator beyond the mode coupling induced by the mask v . This bias can however be estimated analytically and corrected for. This was done in Elsner et al. (2017) for scalar quantities, and we generalize their result here to fields of arbitrary spin.

⁴It is worth noting that the formalism described here assumes that a given contaminant template has no significant contribution from the signal itself. This might not be the case, for example, if one blindly took a high-frequency sky map to serve as a template of dust, since that may still contain a significant CMB contribution. Care must therefore be exercised when defining the contaminant templates to marginalize over.

⁵The assumption of uncorrelated weights is the key difference between pseudo- C_ℓ and optimal quadratic estimators (Leistedt et al. 2013). Note that, since the templates are already multiplied by the mask, the weights are not equal across the map, but correspond to the mask values in general.

Considering a second map \mathbf{b} with mask w , N_b contaminants \mathbf{g}^i and correlation matrix G_{ij} , the mode-coupled pseudo- C_ℓ of \mathbf{a}^c and \mathbf{b}^c is given by:

$$\begin{aligned} \text{PCL}_\ell(\mathbf{a}^c, \mathbf{b}^c) &= \text{PCL}_\ell(\mathbf{a}^v, \mathbf{b}^w) \\ &- \frac{G_{ij}}{2\ell+1} \sum_m \sum_{\nu'} \mathbf{a}_1^{\nu} \mathbf{b}_1^{\nu\prime\uparrow} \mathbf{g}_1^i \mathbf{g}_1^{i\uparrow} \\ &- \frac{F_{ij}}{2\ell+1} \sum_m \sum_{\nu'} \mathbf{f}_1^i \mathbf{f}_1^{j\uparrow} \mathbf{a}_1^{\nu} \mathbf{b}_1^{\nu\prime\uparrow} \\ &+ \frac{F_{ij} G_{rs}}{2\ell+1} \sum_m \sum_{\nu', \nu''} \mathbf{f}_1^i \mathbf{f}_1^{j\uparrow} \mathbf{a}_1^{\nu} \mathbf{b}_1^{\nu\prime\uparrow} \mathbf{g}_1^r \mathbf{g}_1^{s\uparrow}. \end{aligned} \quad (22)$$

Taking the ensemble average of the equation above, we find that the first term is the usual pseudo- C_ℓ estimator. The combination of the last three terms is what we will refer to as the *deprojection bias*, and can be computed as⁶:

$$\begin{aligned} \Delta C_\ell^{ab} &= -G_{ij} \text{PCL}_\ell(\tilde{\mathbf{g}}_1^i, \mathbf{g}_1^i) - F_{ij} \text{PCL}_\ell(\mathbf{f}_1^i, \tilde{\mathbf{f}}_1^i) \\ &+ F_{ij} G_{rs} \left[\int d\hat{\theta} \mathbf{f}^{j\uparrow}(\hat{\theta}) \tilde{\mathbf{g}}^r(\hat{\theta}) \right] \text{PCL}_\ell(\mathbf{f}^i, \mathbf{g}^s), \end{aligned} \quad (23)$$

where

$$\tilde{\mathbf{g}}_1^i \equiv \mathcal{S} \left[v(\hat{\theta}) \mathcal{S}^{-1} \left[C_{\ell_1}^{ab} \mathcal{S} \left[w \mathbf{g}_1^i \right]_{\hat{\theta}} \right]_{\hat{\theta}} \right]_{\hat{\theta}}^{s_a}, \quad (24)$$

$$\tilde{\mathbf{f}}_1^i \equiv \mathcal{S} \left[w(\hat{\theta}) \mathcal{S}^{-1} \left[C_{\ell_1}^{ab\uparrow} \mathcal{S} \left[v \mathbf{f}_1^i \right]_{\hat{\theta}} \right]_{\hat{\theta}} \right]_{\hat{\theta}}^{s_b}. \quad (25)$$

The procedure to obtain an unbiased estimator of the power spectrum in the presence of contaminants can therefore be summarized as follows:

- (i) Produce clean versions of both maps [i.e. estimate the best-fitting coefficients in equation (18) and subtract the contamination as in equation (20)].
- (ii) Estimate the pseudo- C_ℓ of the clean maps (equation 7).
- (iii) Estimate the deprediction bias in equation (23) and subtract it from the pseudo- C_ℓ above.
- (iv) Use the methods described in Sections 2.1.2 and 2.1.3 to account for the mode-coupling matrix.

Two further points must be addressed before moving on. First, the computation of the deprediction bias requires an estimate of the true power spectrum C_ℓ^{ab} (see equations 24 and 25). In practice, this can be done through an iterative approach, starting from an initial guess of C_ℓ^{ab} and replacing it by the estimated power spectrum in the previous iteration. Alternatively, the ensemble average of equation (22) can be seen as a convolution of C_ℓ^{ab} , which could be treated by computing the convolution matrix as a modified version of the standard mode-coupling matrix that accounts for mode deprediction. Estimating this convolution matrix in this case becomes significantly more computationally intensive, and therefore an iterative approach is preferred.

Secondly, estimating the deprediction coefficients (equation 18) involves inverting the correlation matrix of all the contaminant templates. This will not be in general possible (e.g. if a set of templates are linearly related). The problem can however be solved self-consistently by using the Moore–Penrose pseudo-inverse (Elsner

et al. 2017). This is trivial for a symmetric and positive-definite matrix such as F , and corresponds to setting the inverse of all of its zero eigenvalues (or all those below a given threshold) to zero.⁷

2.3 E/B purification

As described in Lewis et al. (2002), Bunn et al. (2003), Smith (2006), Zhao & Baskaran (2010), Kim & Naselsky (2010), and Bunn (2011), B -mode purification refers to the map-level removal of the contamination from E modes in the B -mode component of a given map caused by an incomplete sky coverage and vice versa.⁸ The procedure is particularly useful in situations in which the E -mode component of the signal is significantly larger than the B modes, as is the case in for the CMB. In this case, removing the leakage at the power-spectrum level (i.e. the standard pseudo- C_ℓ approach) produces a suboptimal estimator in which the variance in the B -mode power spectrum is dominated by the variance of the leaked E modes. This section presents the purification method and the associated modifications to the pseudo- C_ℓ algorithm as described in the references above. In this section, we describe the algorithm in terms of purifying B modes from E -mode contamination, but we note that the reverse, i.e. purifying E modes from B -mode contamination is conceptually exactly the same. Our discussion is also specific for spin-2 fields.

We define a field \mathbf{f} to be a B mode if $(D_s^E)^\dagger \mathbf{f} = 0$, where D_s^E is the differential operator defined in equation (A3). At the same time, and under the definition of the dot product:

$$(\mathbf{f}, \mathbf{g}) \equiv \int d\hat{\theta} \mathbf{f}^\dagger(\hat{\theta}) \mathbf{g}(\hat{\theta}), \quad (26)$$

we define a *pure* B mode as a field that is orthogonal to all E modes.

Since $D_s^{E\dagger} D_s^B = 0$, one can always generate a B mode by applying D_s^B to a scalar field. It is then possible to show that E and B modes thus defined are orthogonal in the full sky:

$$(D_s^E \varphi, D_s^B \psi) = \int d\hat{\theta} (D_s^E \varphi)^\dagger D_s^B \psi = 0, \quad (27)$$

where φ and ψ are two arbitrary scalar fields. This can be done by integrating by parts and noting that the celestial sphere has no boundaries. On a cut sky, however, and for $s = 2$, this is only true if the fields satisfy Neumann and Dirichlet boundary conditions simultaneously (i.e. vanishing value and first derivative on the boundary of the cut sky region).

Let $w(\hat{\theta})$ be a sky window function defining the sky region to be analysed (and the weight to be applied in each pixel). The standard pseudo B mode of a spin-2 field \mathbf{P} is then given by

$$\begin{aligned} \tilde{\mathbf{B}}_1 &\equiv \int d\hat{\theta} w(\hat{\theta}) \left({}_s Y_1^B(\hat{\theta}) \right)^\dagger \mathbf{P} \\ &= \int d\hat{\theta} w(\hat{\theta}) (D_s^B Y_1)^\dagger \mathbf{P}(\hat{\theta}), \end{aligned} \quad (28)$$

Since $D_s^B Y_{\ell m}$ is a B mode, in the absence of w , this expression would correspond to a projection that filters out all the E modes from \mathbf{P} . However, $w(\hat{\theta}) D_s^B Y_{\ell m}$ is not a B mode, and therefore $\tilde{\mathbf{B}}_{\ell m}$ receives contributions from ambiguous E modes (which then propagate into

⁷The choice of eigenvalue threshold in NAMASTER is accessible to the user.

⁸Note that this is a different effect from the leakage between different polarized and unpolarized channels in CMB observations due to instrumental effects (asymmetric beams, pair differencing, polarization angle calibration, etc.).

⁶These equations can be easily derived by expanding \mathbf{a}^v and \mathbf{b}^w in terms of the spherical harmonics of \mathbf{a} , \mathbf{b} , v , and w , making use of equation (3) and repeatedly employing the definitions of the SHT (equations 1 and 2).

the variance of the pseudo- C_ℓ estimator of the power spectrum). The idea behind B -mode purification is to move w to the right of \mathbf{D}_s^B , defining the pure B component:

$$B_1^p = \int d\hat{\theta} (\mathbf{D}_s^B(w Y_1))^\dagger \mathbf{P}(\hat{\theta}). \quad (29)$$

Since $\mathbf{D}_s^B(w Y_1)$ is a B -mode quantity, B_1^p should receive contributions only from B modes.

Expanding $\mathbf{D}_s^B(w Y_1)$, we can write B_1^p as:

$$B_1^p = (\tilde{P}_2)_1^B + 2 \frac{\beta_{\ell,2}}{\beta_{\ell,1}} (\tilde{P}_1)_1^B + \beta_{\ell,2} (\tilde{P}_0)_1^B, \quad (30)$$

where $(a)_1^B$ stands for the B -mode component of field a , and we have defined the fields $\tilde{P}_n = (\partial^{2-n} w)^*(Q + iU)$, where Q and U are the real and imaginary parts of the field P (see Appendix A for the definitions of ∂ and $\beta_{\ell,s}$).

The $B(E)$ -purification of a given field is simply achieved by applying equation (30) to the $B(E)$ -mode component of the field. Note that doing so requires the computation of the first- and second-order derivatives of the weights map w , and therefore purification methods require masks for which this quantities are well defined. This usually involves tapering the mask boundaries to avoid sharp edges. A thorough discussion of these methods can be found in Grain et al. (2009). Note that the derivatives of w can be computed analytically in harmonic space:

$$\partial^n w = S^{-1} \left[\begin{pmatrix} -w_1 \\ \beta_{\ell,n} \end{pmatrix}, 0 \right], \quad (31)$$

At this stage, the pseudo- C_ℓ method proceeds as usual, applying equation (7) to the purified field. The purification process however requires a slight modification of the analytical form of the mode-coupling matrix. Specifically, if either E or B modes of a spin-2 field have been purified equations (D2) and (D3) must be modified by exchanging one factor of

$$\begin{pmatrix} \ell & \ell' & \ell'' \\ 2 & -2 & 0 \end{pmatrix}$$

for

$$\begin{pmatrix} \ell & \ell' & \ell'' \\ 2 & -2 & 0 \end{pmatrix} + 2 \frac{\beta_{\ell,2}}{\beta_{\ell,1}\beta_{\ell'',1}} \begin{pmatrix} \ell & \ell' & \ell'' \\ 1 & -2 & 1 \end{pmatrix} + \frac{\beta_{\ell,2}}{\beta_{\ell'',2}} \begin{pmatrix} \ell & \ell' & \ell'' \\ 0 & -2 & 2 \end{pmatrix}. \quad (32)$$

per purified field.

It is worth noting that, as described in Grain et al. (2009), the performance of E/B purification can be optimized by an appropriate choice of window function. The corresponding optimal spin-0, spin-1, and spin-2 window functions can be estimated using a pre-conditioned conjugate-gradient method. The resulting spin window functions are no longer related by simple covariant derivatives. The performance of this method, compared with simple apodization methods, will depend on the noise properties of the experiment and the footprint geometry. This optimal method is currently not available in NAMASTER, and will be implemented in future releases of the code.

Finally, it is also possible to combine purification and contaminant deprojection. Since E/B purifications involve destroying part of the signal (e.g. projecting out all leaked E modes from a B -mode component), it is more optimal to use the non-pure maps to compute the contamination coefficients [α_i in equation (18)]. Therefore, we will assume here that purification takes place between the steps (i)

and (ii) described at the end of Section 2.2, and not before step (i). In this case, although no additional modifications are needed in the mode-coupling matrix, beyond those we have just described, the expression for the deprojection bias (equation 23) must be slightly altered. In particular, although the fields \tilde{g}^i and \tilde{f}^i in equations (24) and (25) are still computed from the non-pure versions of f^i and g^i , all the maps entering the three instances of PCL_ℓ in equation (23) must be consistently E/B -purified before computing those pseudo- C_ℓ s.

2.4 Beams and noise

In any practical scenario, the observed sky map will contain a mixture of signal s (i.e. the component whose power spectrum we are really interested in) and noise n (i.e. a stochastic contaminant that we do not have a template for). Furthermore, the signal component will often be smoothed on the smallest scales by an instrumental beam (although this latter effect is less relevant for galaxy surveys). We discuss how to deal with both of these complications here.

2.4.1 Beam deconvolution

Including the effect of a spherically symmetric instrumental beam amounts to a simple redefinition of the mode-coupling matrix

$$M_{\ell_1 \ell_2}^{s a s b} \rightarrow M_{\ell_1 \ell_2}^{s a s b} W_{\ell_2}^{ab}, \quad (33)$$

where W_ℓ^{ab} is the product of the harmonic transform of the beams for maps a and b . Note that since any continuous tangent vector field on the sphere must, at some location, be zero (the so-called hairy ball theorem), it is impossible to have an anisotropic beam that is uniform over the entire sphere. However, it is possible that the beam is both anisotropic and spatially varying. In this case, the pseudo- C_ℓ power spectrum estimation will produce biased estimates that will be suppressed by an *effective* beam. Correcting for the beam in such cases can also be done analytically, but we leave this for future work. In practice, if required, the suppression factor can be found by applying the PCL estimator to a set of mock input maps with the same underlying signal but with and without application of the beam smoothing (to both data maps and contaminant maps) and taking the ratio.

2.4.2 Noise bias

In autocorrelation, the two-point correlation of the noise component n will contaminate our estimate of the power spectrum of s , and we must correct for this noise bias. Depending on the statistical properties of n , different strategies can be used to do this, which we describe in Appendix F.

2.5 Flat-sky pseudo- C_ℓ s

Using the curved-sky expression presented above and in Appendix A for the analysis of small sky patches, where the curvature of the sphere can be neglected, is numerically inefficient for two main reasons:

(i) Curved-sky data formats usually store information about the full sphere, even if the data being stored is limited to a small sky patch. This can lead to an inefficient use of memory and storage space, because the standard numerical implementations of spherical transforms operate on full-sky data.

(ii) SHTs are notoriously slower than discrete Fourier transforms (DFTs), both due to the scaling of both algorithms ($\mathcal{O}(N_{\text{pix}}^{3/2})$ for SHTs versus $\mathcal{O}(N_{\text{pix}} \log(N_{\text{pix}}))$ for DFTs) and because by SHTs involve operations over the full sphere.

(iii) Fourier transforms over a flat patch offer exact quadrature, which makes them inherently more numerically stable.⁹

For this reason, NAMASTER also supports the computation of power spectra from flat-sky maps. Appendix C describes the dictionary between spin- s fields and their Fourier/harmonic coefficients defined in flat and curved skies, and we present the flat-sky versions of the pseudo- C_ℓ methods described in the previous sections here.

2.5.1 Standard pseudo- C_ℓ

Using the definitions presented in Appendix C, it is easy to show that the Fourier coefficients of a masked field are given by (Louis et al. 2013)

$$\mathbf{a}_\mathbf{l}^v \equiv \mathcal{D}(v \mathbf{a})_1^{s_a} = \sum_{\mathbf{k}} \frac{\Delta \mathbf{k}^2}{2\pi} R(s_a(\varphi_1 - \varphi_{\mathbf{k}})) \mathbf{a}_{\mathbf{k}} v_{1-\mathbf{k}}. \quad (34)$$

Here, \mathcal{D} stands for a standard 2D DFT, and R is a rotation matrix in two dimensions. Note that, in the context of flat-sky fields, the vector \mathbf{l} is a 2D wave vector, and must not be confused with the abbreviation $\mathbf{l} \equiv (\ell, m)$ used in previous sections. The quantity $\mathbf{a}_\mathbf{l}$ is therefore the Fourier transform of the real-space flat-sky field $\mathbf{a}(\mathbf{x})$, where \mathbf{x} is the angular displacement with respect to a given reference point. Using equation (34), it is then straightforward to estimate the covariance of two masked fields at the same wavenumber:

$$\langle \mathbf{a}_\mathbf{l}^v \mathbf{b}_\mathbf{l}^{w\dagger} \rangle = \sum_{\mathbf{k}} \frac{\Delta \mathbf{k}^2}{(2\pi)^2} R(s_a \Delta \varphi) C_{\mathbf{k}}^{ab} R^\dagger(s_b \Delta \varphi) v_{1-\mathbf{k}} w_{1-\mathbf{k}}^*, \quad (35)$$

where $\Delta \varphi \equiv \varphi_1 - \varphi_{\mathbf{k}}$.

At this stage, it is natural to connect directly with the final bandpowers since, unlike in the curved-sky case, there is no natural minimal binning ($\Delta \ell = 1$) for the unmasked power spectrum. In this case, we will define a bandpower B_q , indexed by an integer q as the average of the covariance above over a set of \mathbf{l} values S_q :

$$\text{vec} [B_q^{ab}] \equiv \frac{\Delta \mathbf{k}^2}{N_q} \sum_{\mathbf{l} \in S_q} \text{vec} [\mathbf{a}_\mathbf{l}^v \mathbf{b}_\mathbf{l}^{w\dagger}], \quad (36)$$

where N_q is the number of Fourier-space pixels covered by S_q . Typically, S_q will be the set of pixels within an annulus defined by an interval ($|\mathbf{k}|_{\text{min}}, |\mathbf{k}|_{\text{max}}$). Taking the expectation value of the equation above, we can relate this estimator to the true power spectrum as:

$$\langle \text{vec} [B_q^{ab}] \rangle = \sum_{\mathbf{l} \in S_b} N_q^{-1} \sum_{\mathbf{k}} M_{\mathbf{l}\mathbf{k}}^{s_a s_b} \cdot \text{vec} [\hat{C}_{\mathbf{k}}^{ab}]. \quad (37)$$

In this case, the unbinned mode-coupling matrices are a function of \mathbf{l} and \mathbf{k} as vectors. We provide analytical formulae for them in Appendix D. The mode-coupling matrix for the bandpowers is then given in terms of the unbinned one as:

$$\mathcal{M}_{qq'}^{s_a s_b} = \sum_{\mathbf{l} \in S_q} N_q^{-1} \sum_{\mathbf{k} \in S_{q'}} M_{\mathbf{l}\mathbf{k}}^{s_a s_b}. \quad (38)$$

At this point, it is important to note that, with this procedure, computing the coupling matrix becomes an $\mathcal{O}(N_{\text{pix}}^2)$ problem (or

at best $\mathcal{O}(\ell_{\text{max}}^2 N_{\text{pix}})$), which is worse than the $\mathcal{O}(\ell_{\text{max}}^3)$ pseudo- C_ℓ algorithm we described in the curved-sky case. The key in that case is the orthogonality relation equation (A11), which significantly simplifies the expressions (D1)–(D3). In a flat sky, and in the presence of a discrete and finite Cartesian grid, it is not possible however to average over azimuthal Fourier degrees of freedom at constant l in order to obtain a numerically stable version of the mask power spectrum for arbitrary weight maps v and w . This would not be a problem in the continuum limit ($(\Delta x, \Delta y) \rightarrow 0$), where similar simplifying relations can be found to vastly improve the computational efficiency of the method. Taking the continuum approximation is unfortunately not accurate enough in practice, as described in Asgari et al. (2016), and a fully discrete approach is usually necessary. This slower performance is normally compensated by the smaller pixel numbers and faster DFTs that can be used in the flat sky, so that the flat-sky approximation is still preferable over a curved-sky treatment for small sky patches. Appendix E discusses these issues in detail, providing the continuum-limit expressions mentioned above.

2.5.2 Contaminant cleaning

Using the same notation as in Section 2.2, the contaminant-cleaned version of \mathbf{a} is simply given by [cf. equations (20) and (21)]:

$$\mathbf{a}^c(\mathbf{x}) = \mathbf{a}^v(\mathbf{x}) - \mathbf{f}^i(\mathbf{x}) F_{ij} \sum_{\mathbf{x}'} \Delta \mathbf{x}^2 \mathbf{f}^{j\dagger}(\mathbf{x}') \mathbf{a}^v(\mathbf{x}'), \quad (39)$$

$$\mathbf{a}_\mathbf{l}^c = \mathbf{a}_\mathbf{l}^v - \mathbf{f}_\mathbf{l}^i F_{ij} \sum_{\mathbf{k}} \Delta \mathbf{k}^2 \mathbf{f}_\mathbf{k}^{j\dagger} \mathbf{a}_\mathbf{k}^v, \quad (40)$$

where $(F^{-1})_{ij} = \sum_{\mathbf{x}} \Delta \mathbf{x}^2 \mathbf{f}^{i\dagger}(\mathbf{x}) \mathbf{f}^j(\mathbf{x})$.

The deprojection bias also takes a similar form [cf. equation (23)]:

$$\Delta C_{\mathbf{l}}^{ab} = -G_{ij} \tilde{\mathbf{g}}_i^j \mathbf{g}_1^{i\dagger} - F_{ij} \mathbf{f}_i^j \tilde{\mathbf{f}}_1^{j\dagger} + \dot{F}_{ij} G_{rs} \left[\sum_{\mathbf{x}} \Delta \mathbf{x}^2 \mathbf{f}^{j\dagger}(\mathbf{x}) \tilde{\mathbf{g}}^r(\mathbf{x}) \right] \mathbf{f}_i^s \mathbf{g}_1^{s\dagger}, \quad (41)$$

where

$$\tilde{\mathbf{g}}_i^j \equiv \mathcal{D} \left[v(\mathbf{x}) \mathcal{D}^{-1} \left[C_{\mathbf{l}}^{ab} \mathcal{D} \left[w \mathbf{g}_i^{s_b} \right]_{\mathbf{l}_1} \right]_{\mathbf{x}} \right]_{\mathbf{l}_1}^{s_a}, \quad (42)$$

$$\tilde{\mathbf{f}}_1^j \equiv \mathcal{D} \left[w(\mathbf{x}) \mathcal{D}^{-1} \left[C_{\mathbf{l}_1}^{ab\dagger} \mathcal{D} \left[v \mathbf{f}_i^{s_a} \right]_{\mathbf{l}_1} \right]_{\mathbf{x}} \right]_{\mathbf{l}_1}^{s_b}. \quad (43)$$

2.5.3 E and B purification

The logic behind E/B purification is the same in flat and curved skies, and we will not repeat it here. The analogue of a pure B mode (equation 30) in the flat-sky approximation is:

$$B_1^p = (\tilde{P}_2)_1^B + 2\ell^{-1} (\tilde{P}_1)_1^B + \ell^{-2} (\tilde{P}_0)_1^B, \quad (44)$$

(and a similar relation for the pure E component), where $P_n = (\partial^{2-n} w) * (Q + iU)$, and w is the sky mask. The derivatives of w can be taken by using the following relation:

$$\partial^n w = \mathcal{D}^{-1} [(-\ell^n w_{\mathbf{l}}, 0)], \quad (45)$$

Note that, in the flat sky, a mathematically (but not computationally) simpler relation for the pure component can be found, given

⁹Note that it is also possible to achieve exact quadrature with certain spherical pixelization schemes [e.g. Doroshkevich et al. (2011)].

by:

$$B_1^p = \int \frac{dk^2}{2\pi} B_k w_{1-k} \frac{k^2}{\ell^2}, \quad (46)$$

and similarly for E modes. Comparing with equation (35) (where R is defined in equation C9), it is easy to see that the key to work out the expressions for the pure- E and B coupling matrices is simply to replace all factors of c and s with k^2/ℓ^2 and 0 respectively in equations (D5) and (D6). Again, under the assumption that purification takes place after contaminant deprojection, the only modification to the expression for the deprojection bias is to make sure that all functions of l in equation (41) are consistently purified [but not any of the other fields appearing there or in equations (42) and (43)].

3 CODE STRUCTURE

All of the methods described above have been implemented in a software package called NAMASTER.¹⁰ We briefly describe the code structure here.

The code is written in C and wrapped into PYTHON to facilitate its use and its combination with other software libraries for astronomy. NAMASTER itself is based on a number of these libraries:

(i) HEALPIX (Górski et al. 2005) is currently the only pixelization scheme supported for curved-sky calculations. Through HEALPIX, NAMASTER also makes use of cfitsio (Pence 1999).

(ii) The basic curved-sky operation carried out by NAMASTER is the SHT. For this, the code makes use of libsharp (Reinecke & Seljebotn 2013).

(iii) In flat sky, SHTs are replaced by 2D DFTs, for which NAMASTER uses the FFTW library (Frigo & Johnson 2005).

(iv) The GNU Scientific Library (Galassi et al. 2009) is also used for some numerical calculations.

As described in the previous sections, some of the operations (e.g. SHTs, deprojection, and purification) take place in each field individually, while others depend on the correlation of a pair of fields (e.g. PCL, mode-coupling matrix, and bias deprojection). It is therefore inefficient to carry out the per-field operations on every field every time a power spectrum is computed, especially if a large number of fields are being cross-correlated. For this reason, at the PYTHON level, NAMASTER is structured around two main classes that incorporate the two different types of operations we just described. Each of these classes are associated to a counterpart structure in C.

3.1 NmtField

This class stores all the necessary information about one individual observed spin-0 or spin-2 field. A `NmtField` object is defined by a sky mask or weights map [e.g. v in equation (17)], 1 or 2 sky maps (depending on the spin), corresponding to an observation of the field on the sky [a^{obs} in equation (17)], a set of contaminant templates [f_i in equation (17)], and the choice to purify or not the E - or B -mode component of the field. Once initialized, this class carries out the following operations:

- (i) SHT or DFT of the field (equations 1 or C10).
- (ii) Contaminant deprojection (equation 20).
- (iii) E - and/or B -mode purification (equation 30).

¹⁰The code source is hosted in <https://github.com/LSSTDESC/NaMaster>, and its documentation can be found at <http://namaster.readthedocs.io/>.

3.2 NmtWorkspace

This class stores the information necessary to compute an unbiased estimate of the power spectrum of two `NmtFields`. The main objective of these objects is to compute mode-coupling matrices. `NmtWorkspaces` are equipped with read/write methods to avoid redoing these calculations when estimating the power spectra of several fields with a common set of masks (e.g. for a large number of tomographic bin or simulations).

NAMASTER also includes other convenience classes to handle bandpowers, covariance matrices, etc., as well as routines to carry out useful operations (e.g. mask apodization and Gaussian simulations) and to wrap up commonly used sequences of operations (e.g. pseudo- C_ℓ computation followed by binning and deconvolution).

The workflow for a typical pseudo- C_ℓ run would be:

```

1 from pymaster import *
2
3 ...
4
5 f2 = NmtField(mask, [map_q, map_u],
6                 templates = [[dust_q, dust_u]],
7                 purify_b = True, beam=beam_ell)
8 f0 = NmtField(mask, [map_t])
9
10 b = NmtBin(nside, nlb = 10)
11
12 wsp = NmtWorkspace()
13 wsp.compute_coupling_matrix(f0, f2, b)
14
15 cl_bias = deprojection_bias(f0, f2, cl_theory)
16
17 cl_coupled = compute_coupled_cell(f0, f2)
18
19 cl_decoupled = wsp.decouple_cell(cl_coupled,
                                cl_bias)

```

Briefly:

- (i) In line 1, we import NAMASTER's PYTHON module.
- (ii) All preliminary I/O operations (reading maps, masks, etc.) take place implicitly in line 3.
- (iii) In line 5, we create a field with spin 2. The field is defined in terms of a mask and maps of the (Q, U) Stokes parameters. We also provide a set of contaminants to deproject, an instrumental beam and request for the field's B modes to be purified. Deprojection and purification will take place at this stage in that order, as explained in Section 2.3. The field's SHT is also estimated upon initialization.
- (iv) We define a second field in line 8. This one is a simpler spin-0 field with no contaminants or beam.
- (v) We define the output bandpower structure in line 10. In this case, we will use bins of 10 multipoles, although more general schemes are supported.
- (vi) In lines 12 and 13, we use an `NmtWorkspace` object to compute the mode-coupling matrix.
- (vii) The deprojection bias is computed in line 15 using a best guess for the true underlying power spectrum.
- (viii) The pseudo- C_ℓ is computed in line 17.
- (ix) The deconvolved bandpowers are finally estimated using the workspace in line 19. Note that the mode-coupling matrix must have been pre-computed in line 13 for this to be successful.

Further details regarding the different features of NAMASTER can be found in the online documentation <http://namaster.readthedocs.io/>.

4 CODE PERFORMANCE AND VALIDATION

The core of our code-validation suite is based on two science examples. These examples are not meant to be realistic mocks of actual data, however, they do contain the dynamic range and noise levels of a typical experimental set-up that we expect in the next decade, while providing us with an exact, known true power spectrum to compare against. This section presents these two examples and the results of this validation in terms of accuracy and computational performance. Results will be presented for both curved- and flat-sky realizations. Note that, in all cases, we use the true signal power spectra as input to compute the bias from contaminant deprojection [see equations (24) and (25)]. This allows us to verify that the code works as expected, and to isolate any residual bias associated with software bugs, rather than an imperfect guess of the true power spectrum.

The following sections present the suite of simulations used to validate the code and the results of this validation for the curved-sky calculation. The flat-sky validation follows a similar procedure, and is presented in Appendix G.

4.1 Validation suite

We base our validation suite on two science cases of relevance for the most relevant next-generation cosmology experiments. We describe these here.

4.1.1 Galaxy clustering and weak lensing

The first validation suite exemplifies the use of pseudo- C_ℓ algorithms in the analysis of LSS data. Current and next-generation imaging surveys such as DES (Dark Energy Survey Collaboration et al. 2016), Kilo-Degree Survey (de Jong et al. 2017), Hyper Suprime Camera (HSC) (Aihara et al. 2018), Large Synoptic Survey Telescope (LSST) (Ivezić et al. 2008), and *Euclid* (Laureijs et al. 2011) will constrain the growth and geometry of structure through cosmic time by using the overdensity of galaxies as a proxy for the matter inhomogeneities (a spin-0 field) and their correlated shape distortions (a spin-2 field at first order), mostly caused by weak lensing. These measurements will mostly follow a tomographic approach, using the angular auto- and cross-power spectra between these observables across different bins of photometric redshift (Krause et al. 2017; Joudaki et al. 2018; DES Collaboration et al. 2018).

LSS data are often characterized by inhomogeneous sky coverages and complex mask structures, associated with the presence of bright stars and observational artefacts. They are also affected by many sky contaminants, associated to Galactic sources (dust absorption and star density) and observing conditions (seeing, airmass, etc.) as discussed in Leistedt et al. (2016). The typically flatter power spectra for both galaxy clustering and lensing (e.g. when compared with the damping tail of CMB temperature fluctuations), makes the pseudo- C_ℓ approach ideal for this kind of analysis (Leistedt et al. 2013, 2016). This science case therefore allows us to validate several aspects of the code simultaneously:

(i) The unbiased reconstruction of the six possible correlations between spin-0 and spin-2 fields.

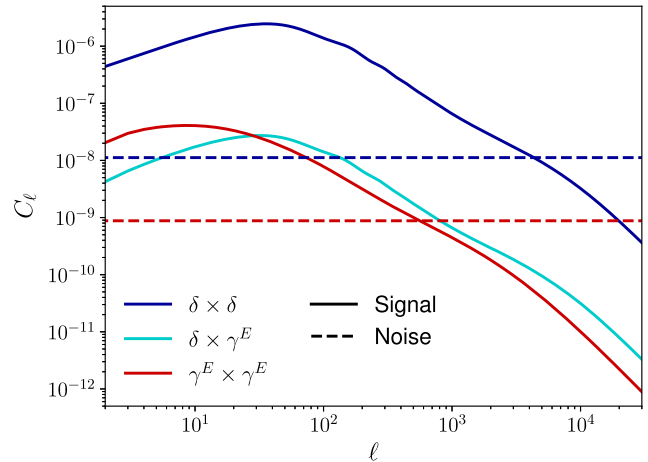


Figure 2. Signal and noise power spectra used in the LSS validation set. Dark blue, cyan, and red curves show the spectra associated with the δ - δ , δ - γ^E , and γ^E - γ^E correlations, respectively, where γ^E is the lensing E mode. All other cross-correlations are zero. The dashed lines show the noise power spectra, associated with shot noise (red) and intrinsic shape scatter (blue).

- (ii) The inclusion of inhomogeneous noise and the correction of the associated noise bias.
- (iii) The effects of deprojection for several contaminant templates in different-spin fields.
- (iv) The code's ability to handle complex masks and high-resolution pixelization.

For this validation suite, we therefore use a mock data set mimicking a typical observation of an LSST-like survey in a particular redshift bin. Slightly different set-ups are used for the curved- and flat-sky validation suites. The characteristics of these data are as follows:

(i) *Signal*: we generate maps of the galaxy overdensity δ and of the shear field (γ_1, γ_2) for a redshift bin centred at $z \simeq 1$ with a width $\Delta z \sim 0.1$ and Gaussian photo- z tails with width $\sigma_z = 0.06$. For δ_g , we assume unit bias $b(z) = 1$ and no contributions from redshift-space distortions or magnification. For γ , we assume no contribution from intrinsic alignments. The maps are generated as Gaussian realizations of the corresponding angular power spectra computed with the Core Cosmology Library¹¹ for cosmological parameters compatible with the 2015 *Planck* measurements (Planck Collaboration XIII 2016). These power spectra are shown in Fig. 2. Sample realizations of δ and γ can be seen in Figs 3 and 4 for the curved- and flat-sky cases, respectively.

(ii) *Mask*: for the curved-sky data, we use a mask built as a combination of the sky coverage output by the LSST `OpSim` database (Delgado et al. 2014) and a more conservative galactic cut based on the reddening map of Schlegel, Finkbeiner & Davis (1998). To add an extra layer of complexity, we drill 100 additional 1-deg holes on the resulting combined mask. We also explore the improvement stemming from the downweighting of higher noise in an inverse-variance way in the presence of inhomogeneous noise (see description below), which is the reason why the default version of this mask is not binary, but traces the inverse of the noise variance (see top left panel of Fig. 3). For the flat-sky data, we use a high-resolution mask constructed from the bright-object mask used for the first public data release from the HSC collaboration (Aihara et al.

¹¹<https://github.com/LSSTDESC/CCL>

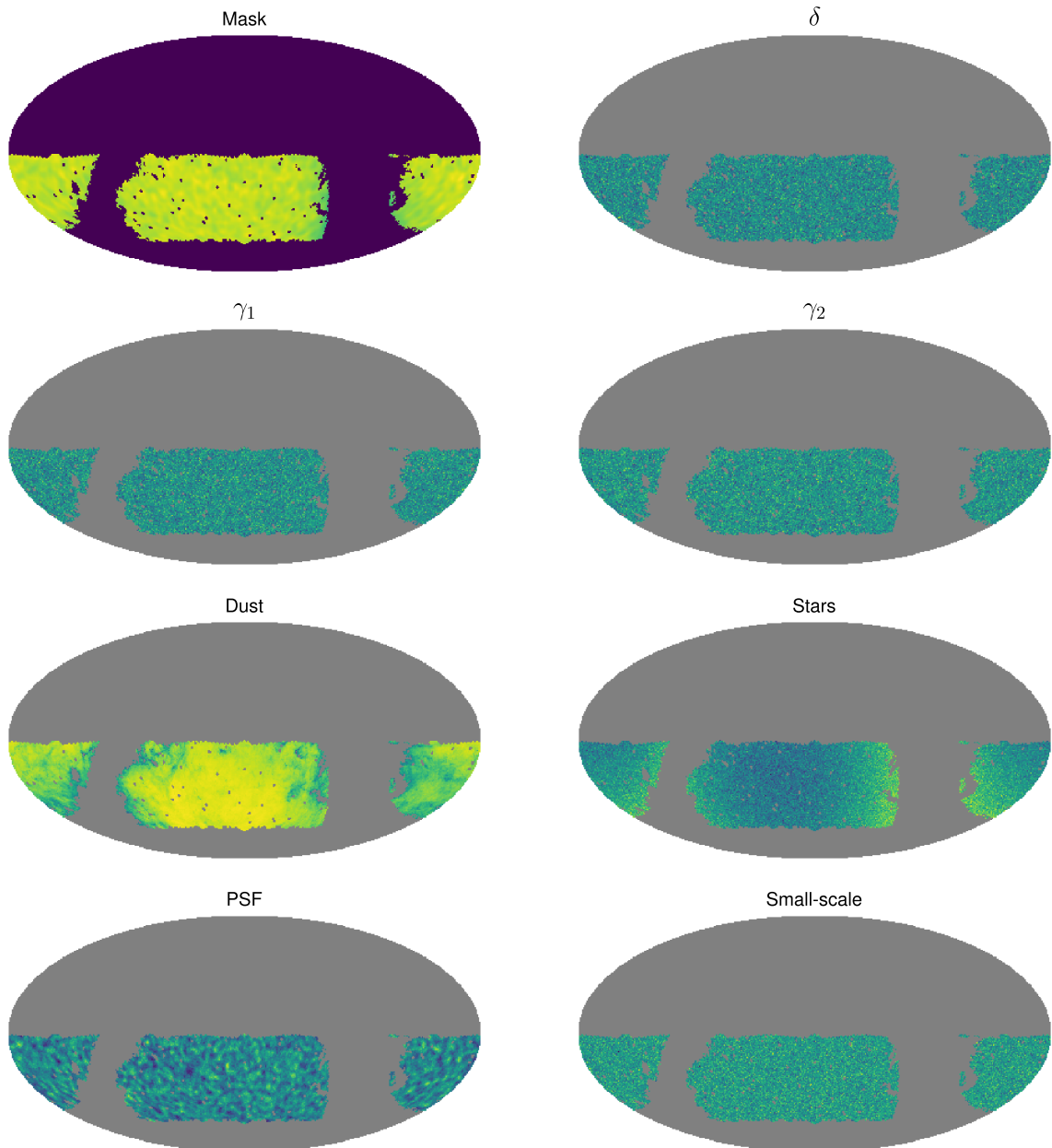


Figure 3. Inputs for the LSS, curved-sky validation set. From top to bottom and left to right: (1) sky mask, (2) Gaussian realization of the galaxy overdensity δ , (3) and (4) cosmic shear γ_i , (5) dust contaminant for δ , (6) star contaminant for δ , (7, Q component) PSF contaminant for γ , and (8, Q component) small-scale contaminant for γ .

2018). The chosen footprint corresponds to the VVDS field, covers approximately 20 deg^2 and contains structure on a wide range of scales. As can be seen in the top left panel of Fig. 4, the mask is not binary, and its value at each pixel correspond to the pixel's fractional masked area.

(iii) *Contaminants*: we contaminate the signal maps described above with several different types of residuals. In all cases, we add contaminants linearly, to avoid deviating from the model described in Section 2.2. The contaminant amplitudes are chosen so that any residual contamination in the relevant range of multipoles will be detected statistically by the validation suite. Maps of these contaminants are shown in the four bottom panels of Figs 3 and 4, and their impact on the angular power spectrum is presented in Fig. 5 for the curved-sky case.

(a) *Dust*: for the curved-sky suite, we include a linear contaminant for δ_g proportional to the fluctuations around the mean of the dust reddening map of Schlegel et al. (1998). This contamination affects mostly the largest scales, and simulates the effect of an imperfect correction of the dust absorption on the galaxy number density. The amplitude of the dust fluctuations is chosen so that their power spectrum within the unmasked region be ~ 20 per cent of the galaxy power spectrum at $\ell \sim 120$. For the flat-sky suite, we generate a flat-sky Gaussian realization of a field with a power spectrum $C_\ell \propto \ell^{-2.4}$, corresponding to the large- ℓ behaviour of the curved-sky dust map. The amplitude was chosen such that the contaminant power spectrum be ~ 10 per cent of the signal at $\ell \sim 400$.

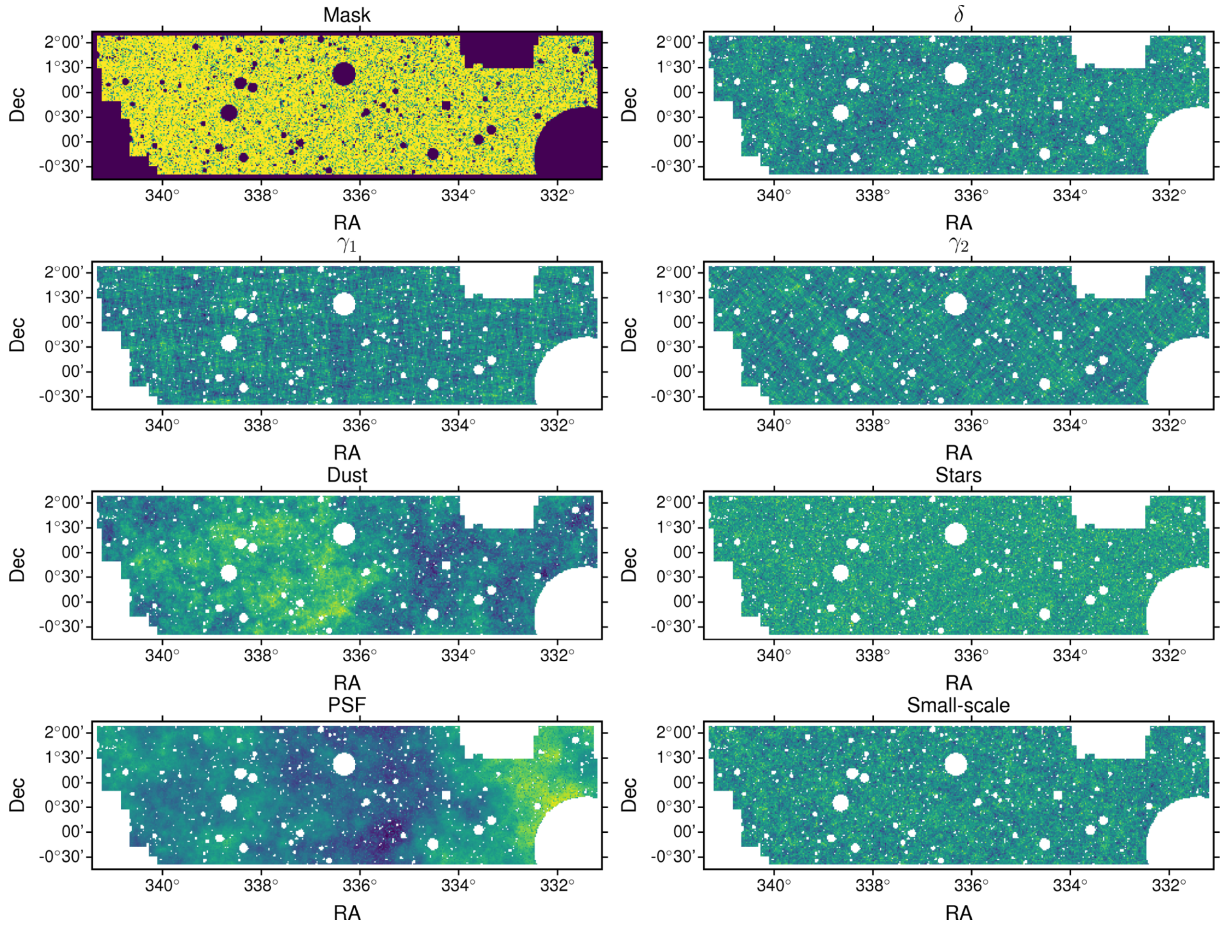


Figure 4. Same as Fig. 3 for the LSS, flat-sky validation set.

(b) *Stars*: we add a second contaminant to δ_g associated with star contamination. For the curved-sky realizations, this contaminant is proportional to the fluctuation around the mean of a star density map generated using the LSST CatSim catalogue (Connolly et al. 2014). We modify this contaminant slightly by adding a white-noise component that dominates the small-scale spectrum beyond $\ell \simeq 400$. We do this in order to simulate the noise-like distribution of stars on small scales as well as to include a contaminant source that, unlike dust, can surpass the signal on small scales. The amplitude of this contaminant is chosen such that its power spectrum is about 3 per cent of the signal at $\ell \sim 400$.

(c) *PSF fluctuations*: we add a large-scale contaminant for weak lensing in the form of fluctuations in the image PSF with a characteristic scale corresponding to the telescope field of view. We generate this contaminant as a spin-2 field with equal E - and B -mode amplitude and a power spectrum

$$C_\ell^{EE} = C_\ell^{BB} \propto \left[2 \frac{J_1(\ell\theta_{\text{FoV}})}{\ell\theta_{\text{FoV}}} \right]^2, \quad (47)$$

where $J_1(x)$ is the order-1 Bessel function, and $2J_1(\ell\theta)/\ell\theta$ is the Fourier transform of a circular aperture with radius θ . For this, we assume $\theta_{\text{FoV}} = 1:75$, and therefore we try to mimic the signature of PSF fluctuations that vary between different pointings with a $3:5$ diameter. The amplitude of the power spectrum is chosen so that the contaminant amounts to 30 per cent of the signal power

spectrum at $\ell \sim 50$. We use the same prescription for both the curved- and the flat-sky cases.

(d) *Small-scale contamination*: finally, we include an additional small-scale contaminant to weak lensing with a flat power spectrum, unit E - B ratio, and an amplitude fixed so the contaminant's power spectrum is 20 per cent of the signal at $\ell = 500$. Although we do not identify this contaminant with any specific source, it could arise from e.g. the impact of stars or blending on shape measurement.

(iv) *Noise*: for the curved-sky case, we generate inhomogeneous noise realizations. To do this, we first generate maps with Gaussian, white and homogeneous noise with a pixel variance given by $1/n_{\text{pix}}$ for δ and $\sigma_y^2/n_{\text{pix}}$ for each shear component, where n_{pix} is the average number of sources in a pixel's area, and $\sigma_y = 0.28$ is the intrinsic ellipticity scatter per component. These noise realizations are then locally rescaled by a factor $f(\hat{\theta}) = f_{\text{star}}(\hat{\theta}) f_{\text{depth}}(\hat{\theta})$ made up of two components:

(a) $f_{\text{star}}(\hat{\theta})$ is a function of the local star density modelled with the same star map used to generate the galaxy clustering contaminant described above. The function interpolates between $f_{\text{star}} = 1$ for low star densities and $f_{\text{star}} = 1.5$ for high star densities. The specific functional form used is

$$f_{\text{star}} = \frac{r_{\text{star}} - 4}{2r_{\text{star}} - 4}, \quad (48)$$

where $r_{\text{star}} = N_{\text{star}}/\max(N_{\text{star}})$ and N_{star} is the star density map.

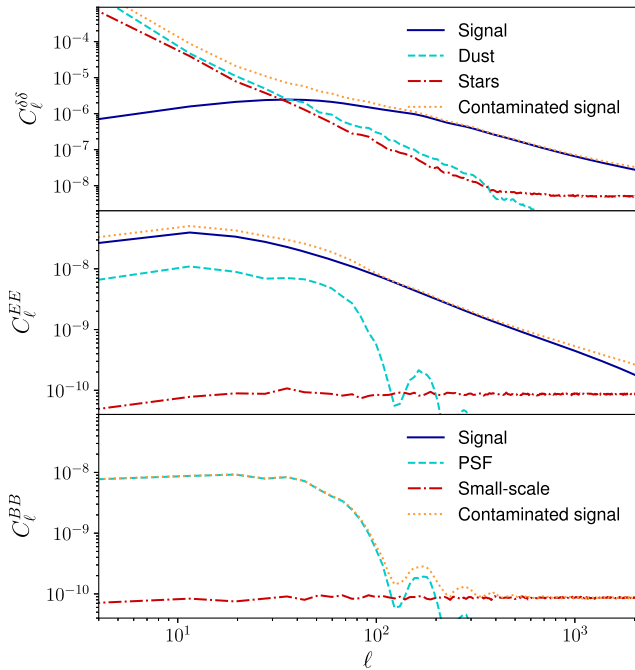


Figure 5. Top: signal power spectrum (solid dark blue) for the autocorrelation of the galaxy overdensity, as well as the contaminants associated to dust (dashed cyan) and stars (dotted–dashed red). Middle: E -mode lensing power spectrum (solid dark blue) and the spectra of the PSF and small-scale contaminants used in our validation set (dashed cyan and dotted–dashed red, respectively). Bottom: same as middle panel for the B -mode power spectrum. Note that in this case the input signal power spectrum is 0. In all cases, the orange dotted line shows the combination of signal and contaminants.

(b) $f_{\text{depth}}(\hat{\theta})$ is a map of stochastic depth variations generated as a Gaussian random field with mean 1 and fluctuations with a typical amplitude of 3–4 per cent and a characteristic scale of $\theta_{\text{FoV}} = 1^\circ.75$.

The logic behind these simulated noise maps is that both the presence of stars and fluctuations in the survey depth will modulate the number density of observed objects, and with it the noise in both δ and γ . A more optimal estimate of the power spectrum could therefore be achieved by using an inverse-variance weighting scheme in which the mask is inversely proportional to $f(\hat{\theta})$.

For simplicity, the flat-sky simulations use only the homogeneous part of these noise realizations.

(v) *Pixelization*: the main suite of curved-sky simulations used to validate the code use HEALPIX resolution parameter $N_{\text{side}} = 1024$, corresponding to a pixel size of ~ 3.4 arcmin. This resolution allows us to validate the code on a large number of realizations without facing strong computational challenges on the main scales of interest ($\ell \lesssim 2000$). As noted in Appendix H, however, the code has also been validated with a smaller suite of high-resolution simulations with $N_{\text{side}} = 4096$, and we have been able to run it successfully up to a maximum $N_{\text{side}} = 8192$ ($\theta_{\text{pix}} \sim 0.4$ arcmin). The flat-sky simulations use square pixels of 0.6 arcmin a side.

We should note that, unlike the simulated maps used here, the true galaxy distribution is markedly non-Gaussian. However, none of the expressions presented in Section 2 for the pseudo- C_ℓ method assume Gaussianity, and therefore the method will deliver unbiased estimates of the harmonic-space two-point function for non-Gaussian fields too.

4.1.2 CMB B modes from the ground

Our second example corresponds to the measurement of B modes in the polarized CMB emission, closely associated with the presence of primordial gravitational waves. This is one of the main science cases targeted by current ground-based Stage-3 experiments such as Background Imaging of Cosmic Extragalactic Polarization (BICEP)/Keck (BICEP2 Collaboration et al. 2016), Advanced Atacama Cosmology Telescope (Pol) (De Bernardis et al. 2016; Louis et al. 2017) or the Simons Array (Suzuki et al. 2016; POLARBEAR Collaboration et al. 2017), as well as next-generation facilities such as the Simons Observatory (The Simons Observatory Collaboration et al. 2018) or CMB Stage-4 (Abazajian et al. 2016). Potential future space missions, such as LiteBIRD (Suzuki et al. 2018) will be able to make low-resolution, curved-sky maps of the polarized CMB to put constraints on the primordial B -mode signal from its reionization bump, at $\ell < 30$, and will likely make use of pixel-based likelihoods and optimal quadratic methods. Ground-based experiments, on the other hand, will be systematics limited at low ℓ , and will be able to make higher resolution maps, for which pseudo- C_ℓ methods are better suited.

The measurement of CMB B modes is affected by two challenges: the presence of foreground residuals left in the data by any component separation method, and the much larger amplitude of the E -mode signal, which can dominate the uncertainties in the B -mode power spectrum due to E -to- B leakage. This science case therefore allows us to validate the code when both contaminant deprojection and B -mode purification are required, as well as in the presence of a non-negligible instrumental beam.

Our first validation suite is therefore inspired on a Southern hemisphere, high-sensitivity polarization experiment with the following characteristics:

(i) *Signal*: we generate Q and U maps of the CMB as Gaussian realization of a lensed power spectrum corresponding to the best-fitting parameters from *Planck* (Planck Collaboration XIII 2016). For simplicity, we assume no delensing and zero primordial B -modes, so that the B -mode power spectrum is entirely due to the effects of CMB lensing. The signal power spectra are shown in Fig. 6.

(ii) *Mask*: for the curved-sky case, we assume a survey targeting a 4000 deg² footprint. The corresponding mask is analytically apodized using the ‘ C^2 ’ apodization method of Grain et al. (2009) with an apodization scale of 10 deg. This apodization method is based on smoothing the transition at the mask edge with a sinusoidal function that interpolates between 0 and 1 in a differentiable manner. The observed area encompasses one of the lowest foreground regions observable from the Southern hemisphere. For the flat-sky case, we assume a smaller sky fraction of 500 deg² with a similar apodization. The assumed masks are shown in the top panels of Fig. 7.

(iii) *Noise*: for the curved- and flat-sky cases, we assume white noise levels of 1 and 0.5 μK arcmin in intensity, respectively. We also include a $1/f$ -like component associated with atmospheric noise with a characteristic scale ℓ_{knee} and tilt α_{knee} such that the noise power spectrum can be written as

$$N_\ell = \sigma_N^2 \left[1 + \left(\frac{\ell_{\text{knee}}}{\ell} \right)^{\alpha_{\text{knee}}} \right], \quad (49)$$

where σ_N^2 is the white-noise variance in one steradian. We use $\alpha_{\text{knee}} = 2.4$ and $\ell_{\text{knee}} = 10$ and 300 for the curved- and flat-sky cases, respectively. The beam-deconvolved noise power spectra in both cases are shown in Fig. 6.

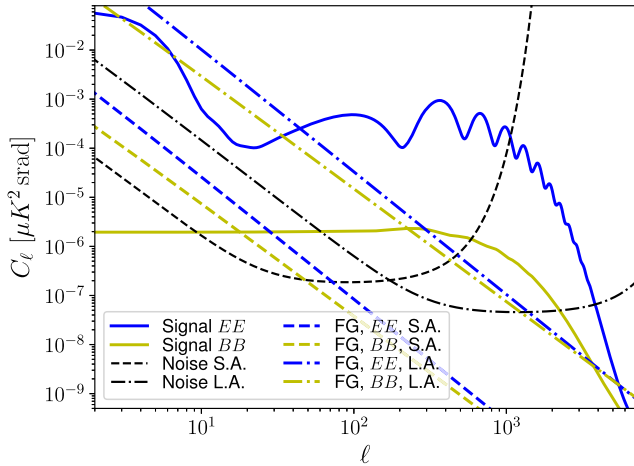


Figure 6. E - and B -mode power spectra (blue and yellow, respectively) for the cosmological CMB signal (solid lines), for the curved-sky foreground contaminants (dashed lines), and for the flat-sky foreground contaminants (dotted–dashed lines). The noise power spectra used for the full- and flat-sky simulations are shown as black dashed and dotted–dashed lines respectively. These noise curves are intended to depict two different types of experiments: a small-aperture telescope (S.A., dashed lines) targeting primordial tensor modes, and a large-aperture telescope (L.A., dotted–dashed lines) targeting the small-scale lensing B -modes.

(iv) *Beam*: we assume a Gaussian beam with full width at half-maximum of $\theta_{\text{FWHM}} = 20$ and 1.4 arcmin for the curved- and flat-sky cases, respectively. This resolution corresponds to what could be achieved at 150 GHz with a small aperture experiment (0.4 m, S.A.) and a large aperture experiment (6 m, L.A.) in either case. The choice of ℓ_{knee} described above for the noise curves is also consistent with the atmospheric noise levels achievable with those types of telescopes (The Simons Observatory Collaboration et al. 2018). As shown in Fig. 6, the combination of beam size and ℓ_{knee} makes the two experiments sensitive to the B -mode power spectrum in different ranges of scales, corresponding to the primordial B -mode signal around the recombination bump (S.A.) and to the high- ℓ lensing signal (L.A.).

(v) *Contaminants*: we assume a dust-like residual for both the curved- and flat-sky cases. For the curved-sky set-up, we use maps of the dust emission at 150 GHz generated by PySM (Thorne et al. 2017) and scaled so that the amplitude of the residual power spectrum is about 10 per cent of the lensing B -mode signal at $\ell = 50$. For the flat-sky signal, we generate a Gaussian realization of a dust-like component with power spectra $C_\ell^{EE} \propto \ell^{-2.5}$ and $C_\ell^{BB} \propto \ell^{-2.3}$, where the spectral tilts were estimated from the high- ℓ behaviour of the PySM maps in the low-foreground footprint. The amplitudes of these power spectra were fixed by requiring that the B -mode power be about 20 per cent of the lensing B -mode power spectrum at $\ell = 550$ and so that $C_\ell^{EE}/C_\ell^{BB} \simeq 2$, as found in Planck Collaboration XIII (2016). The foreground residual maps we use are shown in the bottom panels of Fig. 7, and their power spectra are compared with the signal and noise spectra in Fig. 6.

(vi) *Pixelization*: the curved-sky case, targeting the B -mode recombination bump at $\ell \sim 100$, uses a HEALPIX resolution parameter $N_{\text{side}} = 256$, corresponding to 14-arcmin pixels and sufficient to measure power spectra to multipoles $\ell \sim 500$. The flat-sky case, targeting the lensing B modes, uses square 2-arcmin pixels. These correspond to scales $\ell \lesssim 5000$, well within the signal-dominated regime given the noise level and beam sizes quoted above.

4.2 Curved-sky large-scale structure validation

We generate 1000 realizations of contaminated maps of the galaxy overdensity δ and cosmic shear γ using the models described in Section 4.1.1 with resolution $N_{\text{side}} = 1024$. In our fiducial run, we included all contaminants, and used a non-binary mask that traces the shot-noise fluctuations due to the presence of stars and depth variations. For each simulation, we use NAMASTER to compute all six possible cross-correlations between the δ and the E and B modes of γ (which we will label $\gamma_{E,B}$), including map-level contaminant deprojection and removal of the noise and deprojection bias in the power spectrum. We use bandpowers defined by averaging over groups of $\Delta\ell = 3$ multipoles. The typical running time for one realization is ~ 200 s using 32 OpenMP threads. Subsequent runs are typically faster since we cache the coupling matrices, which are the slowest step in the pipeline. More details about timing and scalability of NAMASTER can be found in Appendix I.

The upper plot in the left-hand panel of Fig. 8 shows the six deconvolved bandpowers averaged over simulations (solid lines) compared with the analytical prediction for the input signal power spectra (darker dashed lines). The middle panel of the same figure shows the average residual with respect to the analytical prediction normalized by the error on the mean (given by the standard deviation of all $N_{\text{sim}} = 1000$ simulations divided by $\sqrt{N_{\text{sim}}}$). We see that the latter are accurately recovered by NAMASTER in all cases, including all null correlations ($\delta \times \gamma_B$, $\gamma_E \times \gamma_B$, and $\gamma_B \times \gamma_B$), and that the normalized residuals fluctuate around zero with typical variations of $\Delta C_\ell/\sigma_\ell \lesssim 3$, corresponding to expected $\lesssim 3\sigma$ fluctuations. For comparison, the bottom panel of the figure shows the mean relative difference for power spectra of contaminated maps without deprojection, where we can appreciate that the residuals, $\Delta C'_\ell/\sigma'_\ell$, appear to not be centred around zero in some cases, i.e. the estimated power spectra are biased.

To quantify the presence of any residual bias in the estimator we start by computing, for each simulation and cross-power spectrum in the validation suite, the quantity:

$$\chi_i^2 = (\mathbf{s}_i - \mathbf{t})^T \text{Cov}^{-1} (\mathbf{s}_i - \mathbf{t}), \quad (50)$$

where \mathbf{s}_i is the vector of N_{dof} power-spectrum values estimated in the i th simulation, \mathbf{t} is the analytical prediction for the input power spectrum, and Cov is the covariance matrix of the estimator, which we compute by averaging over all simulations:

$$\text{Cov} = \frac{1}{N_{\text{sim}}} \sum_{i=1}^{N_{\text{sim}}} (\mathbf{s}_i - \bar{\mathbf{s}})(\mathbf{s}_i - \bar{\mathbf{s}})^T, \quad \bar{\mathbf{s}} \equiv \frac{1}{N_{\text{sim}}} \sum_{i=1}^{N_{\text{sim}}} \mathbf{s}_i. \quad (51)$$

The full covariance matrix is very close to diagonal, as shown in the right-hand panel of Fig. 8. Under the assumption that the estimator is unbiased ($\langle \mathbf{s} \rangle = \mathbf{t}$), the values of χ_i^2 should follow a χ^2 distribution with an N_{dof} degrees of freedom. Fig. 9 shows the distribution of χ_i^2 across simulations approximating the covariance matrix as diagonal (blue histograms), compared with the predicted χ^2 distribution for $N_{\text{dof}} = 681$ (marked by the vertical dashed lines) for the six different power-spectrum combinations. Both distributions are found to be compatible. Moreover, the dotted–dashed lines show the values of χ^2 found for the mean residual with respect to the input power spectra (middle panel of Fig. 8). The probability to exceed (PTE) associated with these χ^2 values are higher than 10 per cent in all cases, and therefore our validation suite does not show any statistically significant bias in the estimator. This bias, if any, must be at least $\sqrt{1000} \sim 30$ times smaller than the statistical uncertainties associated with our validation suite, which are representative of the statistical power of next-generation galaxy surveys.

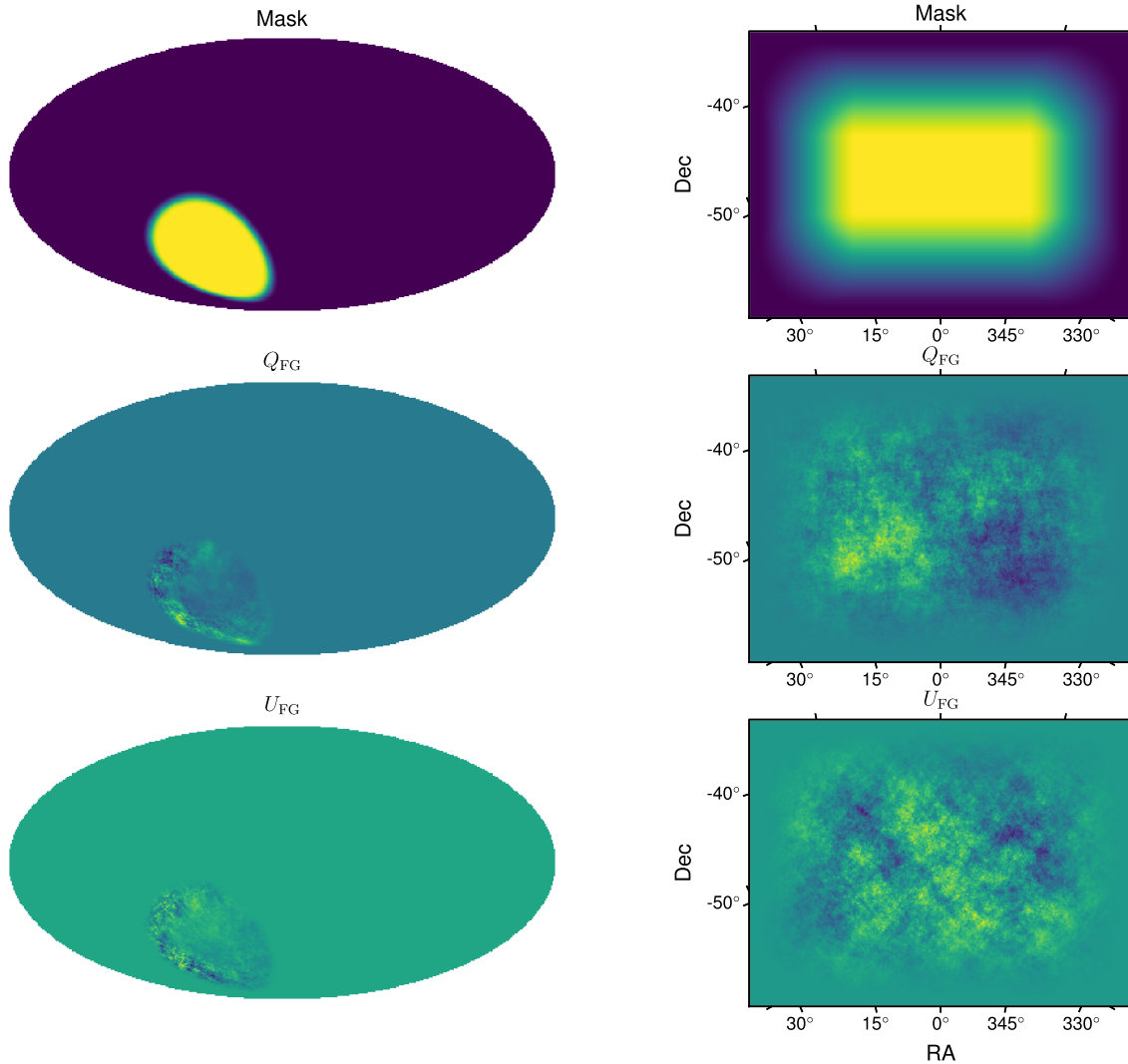


Figure 7. Left-hand panels: mask (top), and Q and U components (middle, bottom) of the foreground contaminant used in the CMB curved-sky validation set. Right-hand panels: same as left-hand panels for the flat-sky validation set. All maps have units of μK_{CMB} .

Our validation suite also allows us to explore the impact of the suboptimal inverse-variance weighting of the pseudo- C_ℓ estimator. The comparison with the uncertainties associated with optimal quadratic estimators has been made before in the literature (Leistedt et al. 2013), and we will not explore this here. Instead, we can study the impact on the final uncertainties of using a mask that implements exact inverse-noise weighting (i.e. the mask shown in the top left corner of Fig. 3, inversely proportional to the number density variations induced by stars and depth fluctuations) against a top-hat mask that weights all pixels equally. Repeating our analysis for the latter mask, we observe a negligible increase of ~ 0.5 -1 per cent in the diagonal error bars. Thus, except in the case of large variations in noise properties (as will be the case for inhomogeneous scanning strategies in CMB experiments, for instance), the exact weighting scheme used with the pseudo- C_ℓ estimator is unlikely to produce significantly better uncertainties.

4.3 Curved-sky CMB B -mode validation

We have carried out a similar set of tests on an ensemble of 1000 simulations from the CMB validation suite described in Section 4.1.2. In

our fiducial run, we both deprojected the foreground contaminants and subsequently used B -mode purification to decouple E and B modes at the map level, since this is the main feature of the code we want to validate. The left-hand panel of Fig. 10 shows the mean of the EE , EB , and BB power spectra computed for all simulations (solid lines) compared with the expectation from the input spectra for EE and BB (dashed lines, mostly superposed on the solid ones). The average residuals normalized over the error on the mean, shown in the lower panel of the same figure, fluctuate around zero with an amplitude of up to 3σ , as would be expected in the absence of a bias in the estimator.

The above figure was created using B -mode purification. Disabling this option has no impact on the bias of the estimates, but increases its variance as we show in the Fig. 11. We plot diagonal error bars with (red) and without (blue) B -mode purification for the BB (solid), EB (dashed), and EE (dotted-dashed) power spectra. Although there is no practical improvement in the error bars for the EE spectrum (there is even a slight increase due to the loss of the ambiguous E modes to purification), and the improvement is only mild for the EB case, very significant gains can be obtained for the BB power spectrum, reducing uncertainties by up to a factor of ~ 10 on large scales.

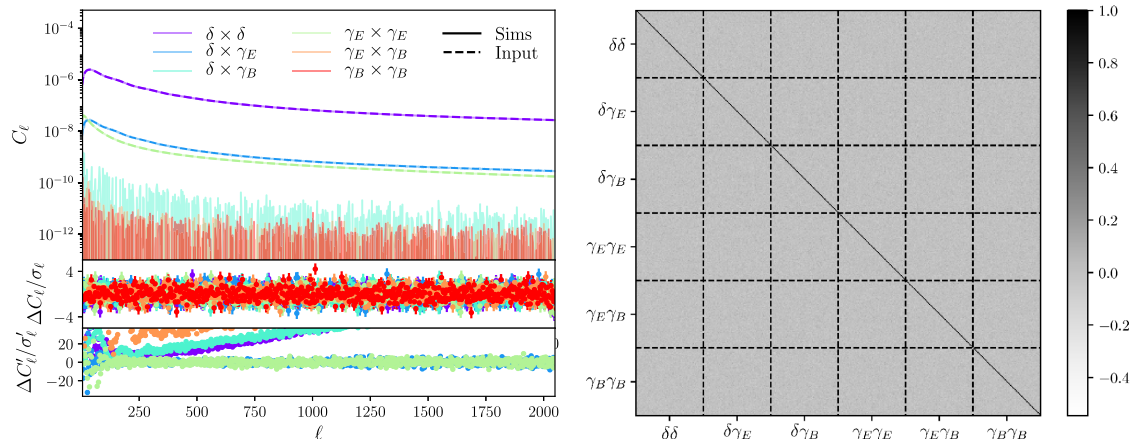


Figure 8. Left-hand panel: average over all simulations in our curved-sky LSS validation suite of the six power spectra between δ , γ_E , and γ_B (solid lines). The analytical prediction for the input signal power spectra are shown as darker dashed lines, and lie mostly superposed on the solid line (for the non-zero cases $\delta \times \delta$, $\delta \times \gamma_E$, and $\gamma_E \times \gamma_E$). The null power spectra ($\delta \times \gamma_B$, $\gamma_E \times \gamma_B$, and $\gamma_B \times \gamma_B$) are compatible with zero. The middle subplot shows the mean residuals with respect to the input power spectra normalized by the 1σ error over the mean (i.e. the scatter over simulations divided by $\sqrt{N_{\text{sim}}}$). The residuals are scattered around zero and their fluctuations are well within those expected under the assumption of Gaussian statistics given 1000 independent realizations (see Fig. 9). For comparison, the lower panel shows the mean relative difference for the power spectra of contaminated maps without contaminant deprojection. Right-hand panel: correlation matrix ($C_{ij}/\sqrt{C_{ii}C_{jj}}$) of all measured power spectra estimated from 1000 realizations.

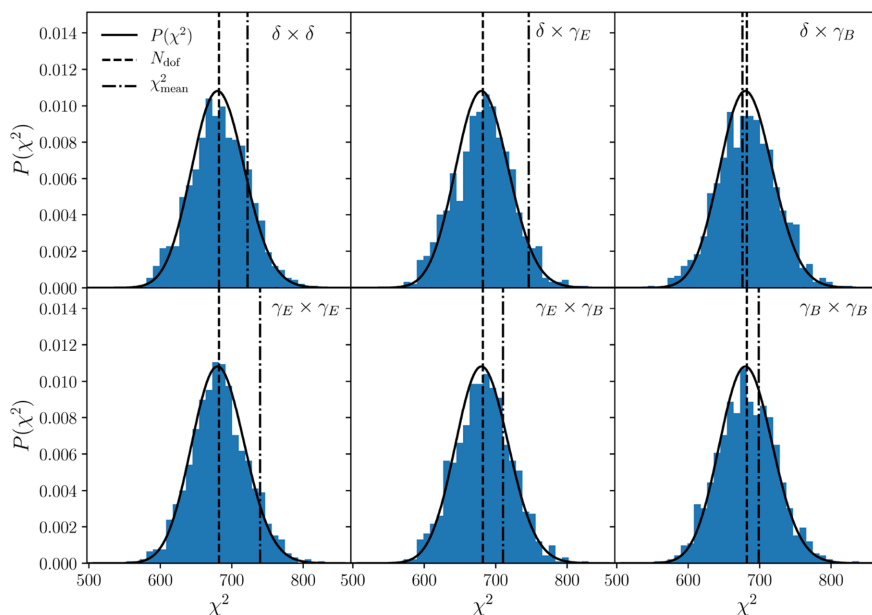


Figure 9. Distribution of the χ^2 values found for each of the 1000 simulations in our curved-sky LSS validation suite (blue histograms). Results are shown for all possible cross-correlations, labelled in the top right corner of each subplot. The simulated distributions are compared with the expected χ^2 distribution (solid lines) assuming Gaussian statistics for $N_{\text{dof}} = 681$ degrees of freedom (denoted by the dashed vertical lines). The dotted–dashed vertical lines mark the χ^2 value found for the mean residual with respect to the input power spectrum (see middle panel of the left-hand panel in Fig. 8). The associated PTE values are $\gtrsim 0.1$, and thus no statistically significant bias is found by our validation suite.

At this point, we can use the validation suite to illustrate the relation between the input power spectrum and the theoretical prediction for the de-convolved bandpowers. As described by equations (15) and (16), the mode-coupling associated with the sky mask and the subsequent decoupling into bandpowers performed by inverting the binned mode-coupling matrix implies that the window function $\mathcal{F}_{q\ell}$ in equation (16) differs in general from a simple binning operator, even if the bandpower weights in equation (12) are chosen to be top-hat functions with a fixed width $\Delta\ell = 12$. This is illustrated in the

right-hand panel of Fig. 10, which shows, in black, the exact window function for different bandpowers (solid black), in comparison with the input top-hat bandpower weights (solid red). Results are shown for the BB power spectrum. Depending on the specific problem, and on the range of scales under study using the simpler top-hat windows or even simply evaluating the theory power spectra at the bandpower centres may lead to e.g. significant biases in subsequent parameter inference stages. NAMASTER provides simple and fast functions to apply the exact window functions to theory power spectra.

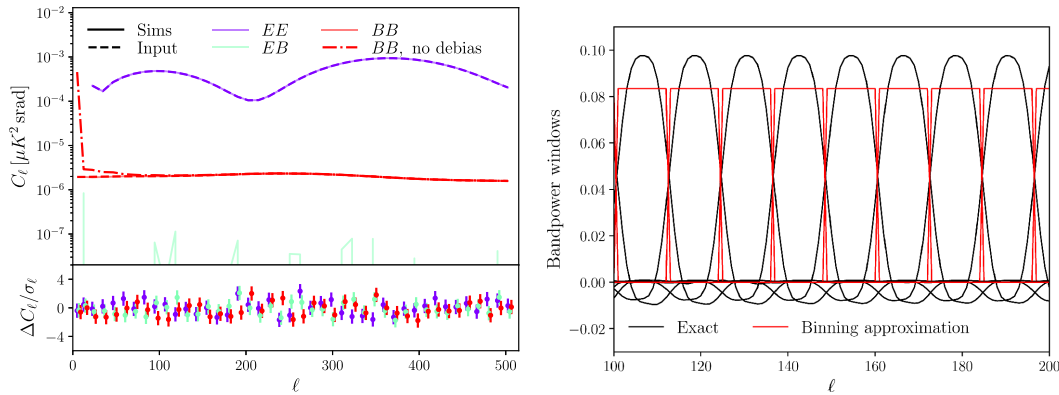


Figure 10. Left: average over all simulations in our curved-sky CMB validation suite of the three spin-2 power spectra (EE , EB , and BB , solid lines) using mode deprojection and B -mode purification. The analytical prediction for the input signal power spectra are shown as dashed lines, and line mostly under the solid lines for the non-zero cases (EE and BB). The null power spectrum EB is compatible with zero. For comparison, the red dotted–dashed line shows the result of not correcting for the mode-deprojection bias when estimating the B -mode power spectrum, which leads to a strong bias on small ℓ . The bottom panel of this figure shows the mean residuals with respect to the input power spectra normalized by the 1σ error on the mean (i.e. the scatter over simulations divided by $\sqrt{N_{\text{sim}}}$). The residuals are scattered around zero with random fluctuations of up to $\sim 3\sigma$. Right: exact bandpower window functions (solid black), which fully account for the effects of mode coupling, compared with the input top-hat bandpower weights (red). Using the correct window functions is in general important in order to avoid parameter biases. Results are shown for the BB power spectrum.

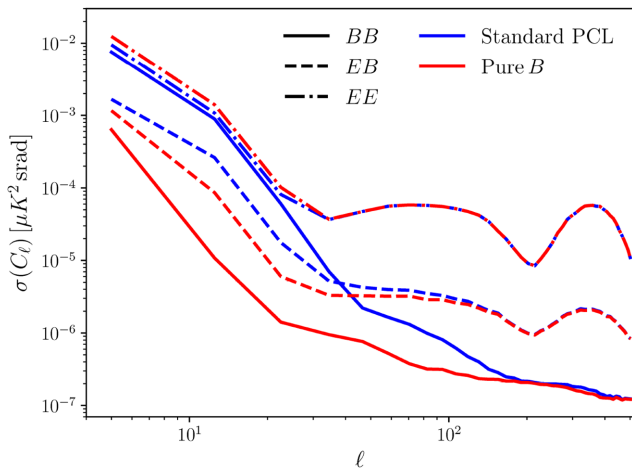


Figure 11. 1σ errors on the BB , EB , and EE power spectra (solid, dashed, and dotted–dashed lines respectively) for the standard pseudo- C_ℓ estimator (blue) and for the estimator with purified B modes (red). B -mode purification is able to produce significant gains in sensitivity for the B -mode power spectrum, particularly at low ℓ .

As in the case of the LSS validation set, we also study the distribution of χ^2 values from each simulation in order to quantify the presence of any possible bias. Unlike in the LSS suite, we find that in this case the covariance matrix is noticeably non-diagonal, and cannot be replaced by the diagonal errors when computing the χ^2 . The non-zero off-diagonal elements are caused by two effects. First, the limited footprint area, map apodization, and choice of bandpowers lead to a noticeable level of anticorrelation between neighbouring bandpowers at low ℓ . This can be seen in the left-hand panel of Fig. 12, which shows the covariance matrix of simulations without foreground contaminants or deprojection. Second, introducing contaminant deprojection produces additional correlations that affect the lowest bandpowers and also involves a level of cross-talk between the three different cross-spectra.

The distribution of χ^2 values found from the simulations is shown in Fig. 13 for the three power spectra. As before, the vertical dashed lines show the number of degrees of freedom ($N_{\text{dof}} = 43$), while the dotted–dashed line marks the χ^2 value of the mean residuals (bottom panel of Fig. 6). The associated PTE values are all above 40 per cent and therefore, as in the case of the LSS validation set, we do not detect any significant bias in the estimator when including both contaminant deprojection and B -mode purification.

5 DISCUSSION

The pseudo- C_ℓ power-spectrum estimator is one particular example of quadratic minimum-variance estimators as presented by Tegmark (1997) in which the proposal covariance matrix of the data is assumed to be diagonal in real space. This allows us to significantly reduce the computational complexity of the problem, from $\mathcal{O}(N_{\text{pix}}^3)$ to $\mathcal{O}(N_{\text{pix}}^{3/2})$ through both analytical and numerical simplifications, as described in Section 2. A number of techniques that are naturally included in the standard quadratic estimator, such as mode deprojection and B -mode purification, can also be implemented in the standard pseudo- C_ℓ algorithm without altering the computational complexity of the method. Although this speed boost is achieved at the expense of optimality (in terms of the estimator’s variance), the degradation in sensitivity will be small or negligible as long as the data do not deviate significantly from the diagonal assumption. This implies that the pseudo- C_ℓ approach is a good choice for non-steep power spectra for which the typical correlation length is smaller than the mask structure, and in general on small scales, where the impact of the power-spectrum shape drops as the number of available modes increases (Efstathiou 2004, 2006). As we have argued, this will be the case for a wide range of relevant science cases pursued by next-generation cosmological observations.

This paper presents NAMASTER, the first public, validated and easy-to-use power-spectrum estimation software package that implements the pseudo- C_ℓ method to compute angular cross-power spectra for any pair of spin-0 or spin-2 fields, including all their generalizations (flat-sky, deprojection and purification). We have

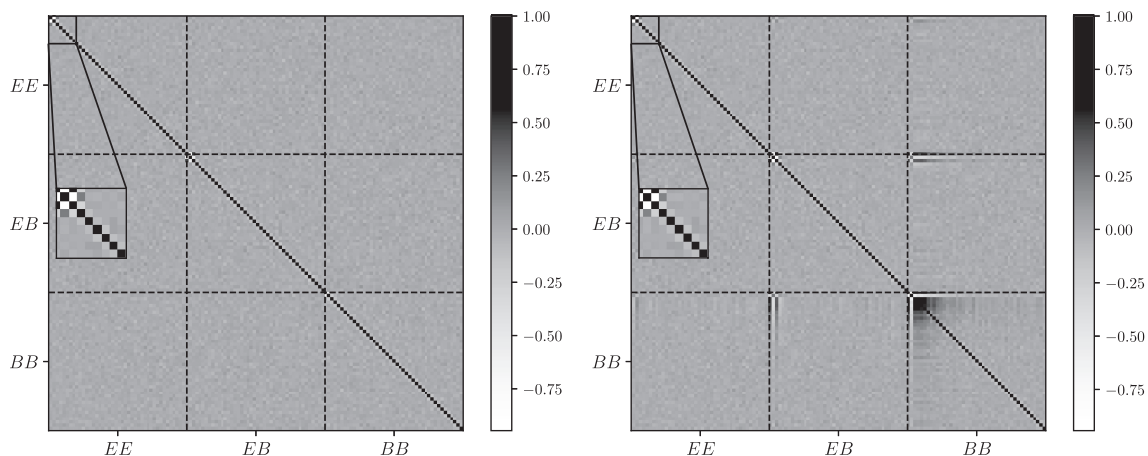


Figure 12. Left: covariance matrix for power spectra computed in the absence of foreground contaminants computed from the 1000 simulations of our curved-sky CMB validation suite. The incomplete sky coverage causes a noticeable anticorrelation of neighbouring bandpowers at low ℓ for the three power spectra (EE , EB , and BB , see inset for the EE case), but beyond that, the covariance matrix is close to diagonal. Right: the same as the left-hand panel but this time for simulations containing foreground contaminants. The combination of B -mode purification and contaminant deprojection produces noticeable correlations at low ℓ , especially in the BB spectrum, that must be taken into account.

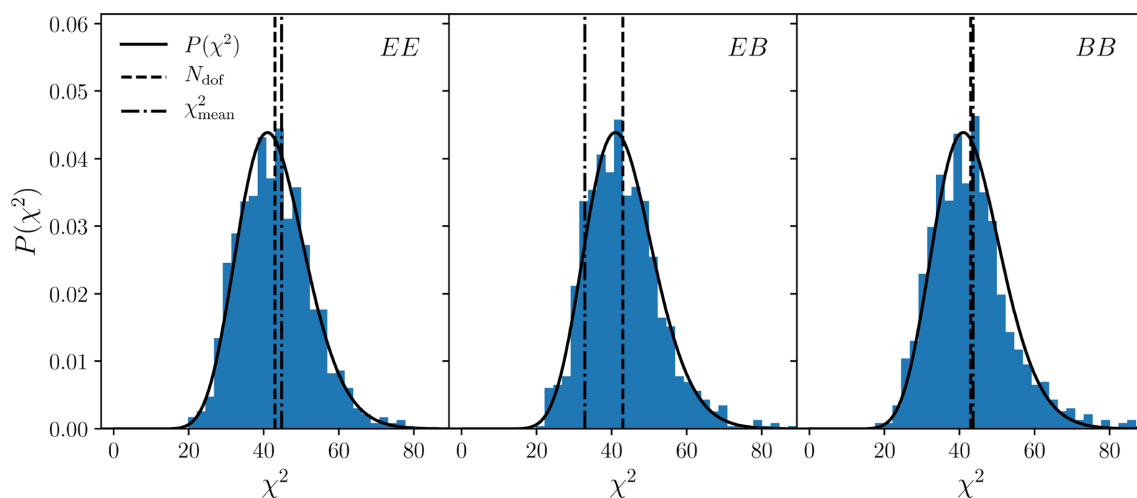


Figure 13. Distribution of χ^2 values found for the 1000 simulations in our curved-sky CMB validation suite (blue histograms). Results are shown for the three polarized spectra (EE , EB , and BB), labelled in the top right corner of each panel. The simulated distributions agree well with the expected χ^2 distribution (solid line) assuming Gaussian statistics for $N_{\text{dof}} = 43$ degrees of freedom (denoted by the dashed vertical lines). The dotted–dashed vertical lines mark the χ^2 value found for the mean residual with respect to the input power spectrum (see lower half of the left-hand panel in Fig. 10). The associated PTE values are $\gtrsim 0.4$, and thus no statistically significant bias is found by our validation suite.

described the steps taken to validate the code, showing that it is able to provide unbiased results in a wide variety of scenarios. In particular, we have constructed a realistic validation suite that resembles the type of data, masks, and contaminants that will affect future wide-area optical galaxy surveys and ground-based CMB polarization experiments. As shown in Section 4, we are able to recover the input power spectrum in all cases (which include varying degrees of contamination, purification, different spins, etc.) with no detectable bias down to $\lesssim 1/30$ th of the statistical uncertainties.

Besides describing and validating the code, and in order to generalize the pseudo- C_ℓ estimator to all the cases supported by NAMASTER, this paper has also presented a number of results that are, to the best of our knowledge, new. These include:

- (i) Extending the analytical treatment of contaminant deprojection presented in Elsner et al. (2016) to fields of arbitrary spin.
- (ii) Combining contaminant deprojection and E/B purification.
- (iii) Consistently deriving all of this functionality (standard pseudo- C_ℓ , deprojection, and purification for arbitrary spins) in the flat-sky approximation.

Although NAMASTER is intended to make power-spectrum estimation as easy as possible, some tasks are too dependent upon the properties of the data to be analysed to automatize them in a general way, and users are therefore in charge of them. We list the most important of these here:

(i) NAMASTER operates directly on *maps*. Galaxy survey data, however, are normally provided in the form of object catalogues, and the conversion those into the desired maps (e.g. galaxy overdensity or shear) must be done externally.

(ii) *E/B* purification can only be carried out in the context of pseudo- C_ℓ estimators when the mask is differentiable up to its second derivatives, which can usually be achieved with an appropriate apodization. Although NAMASTER provides tools to carry out this type of operations, some of the properties of the input mask (e.g. sharp corners) will in general spoil differentiability even after apodization. Users must therefore take care to inspect the mask derivatives before making use of NAMASTER's purification feature.

(iii) A crucial part of any quadratic power-spectrum estimation is the removal of noise bias in the autocorrelation of a given data set. Although we have provided analytic formulas to estimate this bias in certain simple scenarios (see Section 2.4), these are often not applicable. Users must therefore make sure that a sufficiently accurate estimate of the noise bias can be achieved (e.g. through simulations) or to use only cross-correlation data.

(iv) Contaminant deprojection will be able to provide an accurate estimate of the power spectrum as long as two conditions are met. First, users must be able to provide a comprehensive list of possible contaminants as would be observed in their data (e.g. beam-smoothed, mean-subtracted, etc.). Second, a good estimate of the underlying signal power spectra must be provided in order to calculate and subtract the deprojection bias. We have outlined a possible iterative process to do so in Section 2.2.

NAMASTER does not currently produce a covariance matrix estimate. Although computationally efficient analytical estimates of the Gaussian covariance matrix can be derived under certain approximations (Efstathiou 2004; Brown, Castro & Taylor 2005), this is not widely supported by the code for several reasons. First, for many problems, such as LSS power-spectrum estimation, the covariance matrix contains important contributions from the connected four-point function that depend on the precise nature of the problem and go beyond the estimation step that NAMASTER strives to do well. Second, with all additional options included, especially if a large number of contaminant templates are deprojected, the exact analytical estimates are either computational unfeasible or overly approximate. Since a majority of the forthcoming experiments will rely on a large number of mock data sets available for testing systematic effects, these maps also offer a natural path towards estimating the data covariance.

NAMASTER is open-source software and publicly available at <http://github.com/LSSTDESC/NaMaster>. To maximize the usefulness of this software for the community, it is accompanied by an extensive documentation and example code. NAMASTER is written in C and OpenMP-parallelized to maximize its performance, although we encourage its use through the companion PYTHON wrapper, the fastest growing language of choice in the analysis of astronomical data. The code is in constant development, and we will strive to implement further functionality to e.g. estimate covariance matrices, extend the range of supported sky pixelizations and enhance its computational performance.

ACKNOWLEDGEMENTS

This paper has undergone internal review in the LSST Dark Energy Science Collaboration. The internal reviewers were Scott Dodelson, Francois Lanusse, and Boris Leistedt. Besides them, we would like

to thank Rupert Allison, Jo Dunkley, Franz Elsner, Daniel Lenz, Thibaut Louis, Mathew Madhavacheril, Sigurd Næss, Hiranya Peiris, Naomi Robertson, and Ben Thorne for useful comments and discussions.

Author contributions are as follows. DA: Co-led project; wrote the code; contributed to validation. JS: Co-led project; wrote code; contributed to validation. AS: Co-led project; contributed to validation.

DA acknowledges support from the Beecroft trust and from the Science and Technology Facilities Council (STFC) through an Ernest Rutherford Fellowship, grant reference ST/P004474/1. JS acknowledges support from the United States Department of Energy. The DESC acknowledges ongoing support from the Institut National de Physique Nucléaire et de Physique des Particules in France; the Science and Technology Facilities Council in the United Kingdom; and the Department of Energy, the National Science Foundation, and the LSST Corporation in the United States. DESC uses resources of the IN2P3 Computing Center (CC-IN2P3–Lyon/Villeurbanne – France) funded by the Centre National de la Recherche Scientifique; the National Energy Research Scientific Computing Center (NERSC), a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231; STFC DiRAC HPC Facilities, funded by UK BIS National E-infrastructure capital grants; and the UK particle physics grid, supported by the GridPP Collaboration. This work was performed in part under DOE Contract DE-AC02-76SF00515.

REFERENCES

- Abazajian K. N. et al., 2016, preprint ([arXiv:1610.02743](https://arxiv.org/abs/1610.02743))
- Aihara H. et al., 2018, *PASJ*, 70, S4
- Alsing J., Heavens A., Jaffe A. H., Kiessling A., Wandelt B., Hoffmann T., 2016, *MNRAS*, 455, 4452
- Asgari M., Taylor A., Joachimi B., Kitching T. D., 2016, *MNRAS*, 479, 454
- Asorey J., Crocce M., Gaztañaga E., Lewis A., 2012, *MNRAS*, 427, 1891
- BICEP2 Collaboration et al., 2016, *Phys. Rev. Lett.*, 116, 031302
- Bond J. R., 1995, *Phys. Rev. Lett.*, 74, 4369
- Bond J. R., Jaffe A. H., Knox L., 1998, *Phys. Rev. D*, 57, 2117
- Bonvin C., Durrer R., 2011, *Phys. Rev. D*, 84, 063505
- Brown M. L., Castro P. G., Taylor A. N., 2005, *MNRAS*, 360, 1262
- Bunn E. F., 2011, *Phys. Rev. D*, 83, 083003
- Bunn E. F., Zaldarriaga M., Tegmark M., de Oliveira-Costa A., 2003, *Phys. Rev. D*, 67, 023501
- Chon G., Challinor A., Prunet S., Hivon E., Szapudi I., 2004, *MNRAS*, 350, 914
- Connolly A. J. et al., 2014, *Proc. SPIE*, 9150, 915014
- Dark Energy Survey Collaboration et al., 2016, *MNRAS*, 460, 1270
- De Bernardis F. et al., 2016, *Proc. SPIE*, 9910, 991014
- de Jong J. T. A. et al., 2017, *A&A*, 604, A134
- Delgado F., Saha A., Chandrasekharan S., Cook K., Petry C., Ridgway S., 2014, *Proc. SPIE*, 9150, 915015
- DES Collaboration et al., 2018, *Phys. Rev. D*, 98, 043526
- Doroshkevich A. G. et al., 2011, *Int. J. Mod. Phys. D*, 20, 1053
- Driscoll J. R., Healy D. M., 1994, *Adv. Appl. Math.*, 15, 202
- Efstathiou G., 2004, *MNRAS*, 349, 603
- Efstathiou G., 2006, *MNRAS*, 370, 343
- Elsner F., Leistedt B., Peiris H. V., 2016, *MNRAS*, 456, 2095
- Elsner F., Leistedt B., Peiris H. V., 2017, *MNRAS*, 465, 1847
- Eriksen H. K. et al., 2004, *ApJS*, 155, 227
- Frigo M., Johnson S. G., 2005, *Proc. IEEE*, 93, 216
- Galassi M. et al. 2009, GNU Scientific Library Reference Manual, 3 edn. Available at: <https://www.gnu.org/software/gsl/manual/gsl-ref.pdf>
- Gorski K. M., 1994, *ApJ*, 430, L85

- Górski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M., Bartelmann M., 2005, *ApJ*, 622, 759
- Grain J., Tristram M., Stompor R., 2009, *Phys. Rev. D*, 79, 123515
- Hamilton A. J. S., Tegmark M., Padmanabhan N., 2000, *MNRAS*, 317, L23
- Hamimeche S., Lewis A., 2008, *Phys. Rev. D*, 77, 103013
- Hansen F. K., Górski K. M., Hivon E., 2002, *MNRAS*, 336, 1304
- Hivon E., Górski K. M., Netterfield C. B., Crill B. P., Prunet S., Hansen F., 2002, *ApJ*, 567, 2
- Ivezić Ž. et al., 2008, preprint (arXiv:0805.2366)
- Joudaki S. et al., 2018, *MNRAS*, 474, 4894
- Kim J., Naselsky P., 2010, *A&A*, 519, A104
- Kogut A. et al., 2003, *ApJS*, 148, 161
- Krause E. et al., 2017, preprint (arXiv:1706.09359)
- Landy S. D., Szalay A. S., 1993, *ApJ*, 412, 64
- Laureijs R. et al., 2011, preprint (arXiv:1110.3193)
- Leistedt B. et al., 2016, *ApJS*, 226, 24
- Leistedt B., Peiris H. V., Mortlock D. J., Benoit-Lévy A., Pontzen A., 2013, *MNRAS*, 435, 1857
- Lewis A., Challinor A., Turok N., 2002, *Phys. Rev. D*, 65, 023505
- Louis T. et al., 2017, *J. Cosmol. Astropart. Phys.*, 6, 031
- Louis T., Naess S., Das S., Dunkley J., Sherwin B., 2013, *MNRAS*, 435, 2040
- McEwen J. D., Puy G., Thiran J., Vanderghyest P., Ville D. V. D., Wiaux Y., 2011, *Proc. SPIE*, 8138, 81381F
- Peebles P. J. E., 1973, *ApJ*, 185, 413
- Pence W., 1999, in Mehringer D. M., Plante R. L., Roberts D. A., eds, ASP Conf. Ser. Vol. 172, *Astronomical Data Analysis Software and Systems VIII*. Astron. Soc. Pac., San Francisco, p. 487
- Planck Collaboration XIII, 2016, *A&A*, 594, A13
- POLARBEAR Collaboration et al., 2017, *ApJ*, 848, 121
- Reinecke M., Seljebotn D. S., 2013, *A&A*, 554, A112
- Rybicki G. B., Press W. H., 1992, *ApJ*, 398, 169
- Schlegel D., Finkbeiner D., Davis M., 1998, in Colombi S., Mellier Y., Raban B., eds, *Wide Field Surveys in Cosmology*. Editions Frontieres, Paris, p. 297
- Slosar A., Seljak U., Makarov A., 2004, *Phys. Rev. D*, 69, 123003
- Smith K. M., 2006, *Phys. Rev. D*, 74, 083002
- Suzuki A. et al., 2016, *J. Low Temp. Phys.*, 184, 805
- Suzuki A. et al., 2018, *J. Low Temp. Phys.*, 193, 1048
- Szapudi I., Prunet S., Colombi S., 2001, *ApJ*, 561, L11
- Taylor J. F., Ashdown M. A. J., Hobson M. P., 2008, *MNRAS*, 389, 1284
- Tegmark M., 1997, *Phys. Rev. D*, 55, 5895
- Tegmark M., de Oliveira-Costa A., 2001, *Phys. Rev. D*, 64, 063001
- The Simons Observatory Collaboration et al., 2018, preprint (arXiv:1808.07445)
- Thorne B., Dunkley J., Alonso D., Naess S., 2017, *MNRAS*, 469, 2821
- Tristram M., Macías-Pérez J. F., Renault C., Santos D., 2005, *MNRAS*, 358, 833
- Vanneste S., Henrot-Versillé S., Louis T., Tristram M., 2018, *Phys. Rev. D*, 98, 103526
- Wandelt B. D., Hansen F. K., 2003, *Phys. Rev. D*, 67, 023001
- Wandelt B. D., Hivon E., Górski K. M., 2001, *Phys. Rev. D*, 64, 083003
- Zaldarriaga M., Seljak U., 1997, *Phys. Rev. D*, 55, 1830
- Zhao W., Baskaran D., 2010, *Phys. Rev. D*, 82, 023001

SUPPORTING INFORMATION

Supplementary data are available at [MNRAS](https://www.mnras.org) online.

Supplementary material.

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

APPENDIX A: GENERALITIES AND SPHERICAL HARMONIC TRANSFORMS

Let ∂ and $\bar{\partial}$ be the following complex differential operators defined on the sphere when acting on a spin- s quantity f_s :

$$\partial f_s \equiv -(\sin \theta)^s \left(\partial_\theta + i \frac{\partial_\varphi}{\sin \theta} \right) (\sin \theta)^{-s} f_s(\theta, \varphi), \quad (\text{A1})$$

$$\bar{\partial} f_s \equiv -(\sin \theta)^{-s} \left(\partial_\theta - i \frac{\partial_\varphi}{\sin \theta} \right) (\sin \theta)^s f_s(\theta, \varphi).$$

The following properties can be easily derived for the action of these operators, and are useful to derive some of the formulas presented here:

- (i) If f_s is a spin- s quantity, $(f_s)^*$ is a spin- $(-s)$ quantity.
- (ii) ∂f_s is a spin- $(s+1)$ quantity, and $\bar{\partial} f_s$ is a spin- $(s-1)$ quantity.
- (iii) $(\partial^n f_s)^* = \bar{\partial}^n (f_s)^*$
- (iv) $\partial(fg) = f\partial g + g\partial f$
- (v) $\partial^2(fg) = f\partial^2 g + g\partial^2 f + \partial f\partial g$

We start by defining the spin-weighted spherical harmonics with spin $s \geq 0$:

$${}_s Y_{\ell m} \equiv \beta_{\ell, s} \partial^s Y_{\ell m}, \quad -{}_s Y_{\ell m} \equiv \beta_{\ell, s} (-1)^s \bar{\partial}^s Y_{\ell m}, \quad (\text{A2})$$

where $\beta_{\ell, s} \equiv \sqrt{(\ell-s)!/(\ell+s)!}$ and $Y_{\ell m}$ are the standard spherical harmonics. These functions satisfy the property: $({}_s Y_{\ell m})^* = (-1)^{s+m} {}_{-s} Y_{\ell -m}$.

We can then define the E - and B -mode spherical harmonic vectors as:

$$\begin{aligned} {}_s \mathbf{Y}_{\ell m}^E &\equiv \mathbf{D}_s^E Y_{\ell m} \equiv -\frac{\beta_{\ell, s}}{2} \begin{pmatrix} \partial^s + \bar{\partial}^s \\ -i(\partial^s - \bar{\partial}^s) \end{pmatrix} Y_{\ell m} \\ &= -\frac{1}{2} \begin{pmatrix} {}_s Y_{\ell m} + (-1)^s {}_{-s} Y_{\ell m} \\ -i({}_s Y_{\ell m} - (-1)^s {}_{-s} Y_{\ell m}) \end{pmatrix} \end{aligned} \quad (\text{A3})$$

$$\begin{aligned} {}_s \mathbf{Y}_{\ell m}^B &\equiv \mathbf{D}_s^B Y_{\ell m} \equiv -\frac{\beta_{\ell, s}}{2} \begin{pmatrix} i(\partial^s - \bar{\partial}^s) \\ \partial^s + \bar{\partial}^s \end{pmatrix} Y_{\ell m} \\ &= -\frac{1}{2} \begin{pmatrix} i({}_s Y_{\ell m} - (-1)^s {}_{-s} Y_{\ell m}) \\ {}_s Y_{\ell m} + (-1)^s {}_{-s} Y_{\ell m} \end{pmatrix}, \end{aligned} \quad (\text{A4})$$

which also defines the differential operators $\mathbf{D}_s^{E,B}$. For $s=0$, these functions are simply ${}_0 \mathbf{Y}_{\ell m}^E = (Y_{\ell m}, 0)$ and ${}_0 \mathbf{Y}_{\ell m}^B = (0, Y_{\ell m})$.

Finally, we define the matrix operator $\hat{\mathbf{Y}}_{\ell m}^s$ to have ${}_s \mathbf{Y}_{\ell m}^{E,B}$ as columns:

$$\hat{\mathbf{Y}}_{\ell m}^s \equiv ({}_s \mathbf{Y}_{\ell m}^E, {}_s \mathbf{Y}_{\ell m}^B). \quad (\text{A5})$$

These matrices satisfy the following properties:

$$\hat{\mathbf{Y}}_{\ell m}^{s\dagger} = (-1)^{m+s} \hat{\mathbf{Y}}_{\ell -m}^{-s} \quad (\text{A6})$$

$$\int d\hat{\theta} \hat{\mathbf{Y}}_{\ell m}^{s\dagger} \hat{\mathbf{Y}}_{\ell' m'}^s = \hat{1} \delta_{\ell\ell'} \delta_{mm'}, \quad (\text{A7})$$

$$\int d\hat{\theta} \left(\hat{\mathbf{Y}}_1^{s\dagger}(\hat{\theta}) \hat{\mathbf{Y}}_1^s(\hat{\theta}) \right) \hat{\mathbf{Y}}_2^0(\hat{\theta}) \equiv \hat{\mathbf{D}}_{11_2}, \quad (\text{A8})$$

where we have abbreviated the pair (ℓ, m) as l , and

$$\begin{aligned} \hat{\mathbf{D}}_{11_2}^s &= (-1)^{s+m} \sqrt{\frac{(2\ell+1)(2\ell_1+1)(2\ell_2+1)}{4\pi}} \\ &\times \begin{pmatrix} \ell & \ell_1 & \ell_2 \\ -m & m_1 & m_2 \end{pmatrix} \begin{pmatrix} \ell & \ell_1 & \ell_2 \\ s & -s & 0 \end{pmatrix} \hat{a}_{\ell+\ell_1+\ell_2}. \end{aligned} \quad (\text{A9})$$

Here

$$\hat{d}_n = \frac{1}{2} \begin{pmatrix} 1 + (-1)^n & -i[1 - (-1)^n] \\ i[1 - (-1)^n] & 1 + (-1)^n \end{pmatrix}, \quad (\text{A10})$$

and the Wigner $3j$ symbols satisfy the orthogonality relation

$$\sum_{mm_1} \begin{pmatrix} \ell & \ell_1 & \ell_2 \\ m & m_1 & m_2 \end{pmatrix} \begin{pmatrix} \ell & \ell_1 & \ell_3 \\ m & m_1 & m_3 \end{pmatrix} = \frac{\delta_{\ell_2 \ell_3} \delta_{m_2 m_3}}{2\ell_2 + 1} \quad (\text{A11})$$

A1 Pixelization and signal bandlimits

SHT transforms as defined in equation (1) are, strictly speaking, defined for continuous fields on the sky and results in an infinite number of spherical harmonic coefficients. In practice, we have to always deal with discretized sky maps with a finite number of pixels at which the field is sampled, and consequently use a finite number of spherical coefficients. If the sampling of the underlying signal is too sparse, we will get aliasing of higher frequency modes. In flat spaces, Nyquist's theorem provides a clear prescription for uniform sampling on a grid. Since there is no such natural sampling on a sphere, the situation is more complex and depends on the pixelization scheme used (Driscoll & Healy 1994; McEwen et al. 2011). Nevertheless, clear heuristic prescriptions exist. Most importantly, the typical pixel separation needs to be smaller than the wavelength of the highest frequency mode present in the underlying signal. As a natural consequence, the smoothing of the map must always precede any downgrade in resolution (or equivalently downsampling) if it is necessary to do so for numerical expediency. Similarly, any mask should be applied at the highest possible resolution and the resulting map smoothed before being downgraded.

For the same reasons, even if the signal is correctly sampled in a bandlimited sense, one must always estimate power over the entire power band, even if one is not interested in certain region of ℓ space. Otherwise, the power-spectrum estimator will try to 'explain' the variance associated with ℓ modes not considered by artificially inflating the power in the estimated modes. For the HEALPIX pixelization scheme (Górski et al. 2005), commonly used in cosmology and also in this paper, Leistedt et al. (2013) suggest the following prescription: if one wants to measure power up to ℓ_{\max} , one should employ pixelization of $N_{\text{side}} = \ell_{\max}/2$ and estimate bandpowers up to $2\ell_{\max}$, but then discard measurements above ℓ_{\max} after correcting for pixel-response suppression and accounting for their contribution to the covariance matrix of measurements.

In this paper, we do not consider these issues any further as NAMASTER is a generic tool and it is up to the user to ensure bandlimit and pixelization constraints are properly accounted for.

APPENDIX B: LINEAR CONTAMINATION FROM NON-LINEAR CONTAMINATION

Consider a contaminant field c (e.g. a PSF size) that affects the observable field o (e.g. the number count densities) through a non-linear, but *local* on some scale local functional F :

$$\delta_o(\hat{\theta}) \rightarrow \delta_o + F[\delta_c] \quad (\text{B1})$$

where we can, without loss of generality, take out the mean effects (i.e. there is some mean PSF size and fluctuations around this mean size produce additional fluctuations in the number density of fluctuations).

In general, for sufficiently non-linear functionals F , the Taylor expansion around $\delta_c(\hat{\theta})$ is not valid. However, one can look at the

effect of smoothing by a kernel of size $S(R)$, which in Fourier space is equivalent by multiplication by $e^{-k^2 R^2}$. We get

$$\delta_o^R(\hat{\theta}) \rightarrow \delta_o^R + S(R)[F[\delta_c]] = \delta_o^R + F'_R[\delta_c^R], \quad (\text{B2})$$

where we used a shorthand $\delta_x^R = S(R)[\delta_x]$ and F'_R is now a *different* functional that takes a smoothed δ_c and returns out a smoothed δ_o . The crucial point is that F'_R must still be local, since one cannot turn a local functional into a non-local one by application of smoothing.

This means, that for a well behaved δ_c , such that variances of both δ_c^R and δ_o^R are small for a large enough R , the Taylor expansion is valid for F'_R (even if it is not for F), giving

$$\delta_o^R(\hat{\theta}) \rightarrow \delta_o^R + b\delta_c^R + c\nabla^2\delta_c^R \quad (\text{B3})$$

or in Fourier space

$$\delta_o(\hat{\theta}) \rightarrow \delta_o + b\delta_c + ck^2\delta_c \quad (\text{B4})$$

for $k \ll R^{-1}$. In other words, a 'local' contaminant becomes 'linearly biased' in the $k \rightarrow 0$ regime, very much like very non-linear physics of galaxy formations results in linearly biased galaxy field on large scales.

APPENDIX C: FLAT SKIES AND FOURIER TRANSFORMS

C1 Spin- s fields in flat sky

In the flat sky, we will label a position in the plane by two coordinates $(x, y) \equiv \mathbf{x}$. These coordinates can be directly related to increments in the spherical coordinates (θ, φ) by considering a sufficiently small map centred around the equator (i.e. $\sin\theta \sim 1$). In this case, we identify the x coordinate with a latitude shift ($\delta\theta = -\delta x$) and the y coordinate with longitude shifts ($\delta\varphi = \delta y$). The differential operator ∂ now takes the form:

$$\partial = (\partial_x - i\partial_y), \quad \bar{\partial} = (\partial_x + i\partial_y). \quad (\text{C1})$$

The analogue of the standard spherical harmonics for flat skies are plane waves $e^{i\mathbf{k}\mathbf{x}}$, and the action of ∂ on these is:

$$\partial^s e^{i\mathbf{k}\mathbf{x}} = (ik)^s e^{-is\varphi_k} e^{i\mathbf{k}\mathbf{x}}, \quad \bar{\partial}^s e^{i\mathbf{k}\mathbf{x}} = (ik)^s e^{is\varphi_k} e^{i\mathbf{k}\mathbf{x}} \quad (\text{C2})$$

In analogy with the full sky case, we start by defining the basis functions:

$${}_s\mathcal{Y}_{\mathbf{k}}(\mathbf{x}) \equiv k^{-s} \partial^s e^{i\mathbf{k}\mathbf{x}} = i^s e^{-is\varphi_k} e^{i\mathbf{k}\mathbf{x}}, \quad (\text{C3})$$

$${}_{-s}\mathcal{Y}_{\mathbf{k}}(\mathbf{x}) \equiv (-k)^{-s} \bar{\partial}^s e^{i\mathbf{k}\mathbf{x}} = (-i)^s e^{is\varphi_k} e^{i\mathbf{k}\mathbf{x}}, \quad (\text{C4})$$

where φ_k is the polar angle of \mathbf{k} . We then define the Fourier coefficients of a spin- s complex field $a(\mathbf{x})$ as:

$${}_s a_1 \equiv \int \frac{d\mathbf{x}^2}{2\pi} {}_s\mathcal{Y}_1^*(\mathbf{x}) a(\mathbf{x}), \quad {}_{-s} a_1 \equiv \int \frac{d\mathbf{x}^2}{2\pi} {}_{-s}\mathcal{Y}_1^*(\mathbf{x}) a^*(\mathbf{x}).$$

$$a(\mathbf{x}) = \int \frac{d\ell^2}{2\pi} {}_s\mathcal{Y}_1(\mathbf{x}) {}_s a_1, \quad a^*(\mathbf{x}) = \int \frac{d\ell^2}{2\pi} {}_{-s}\mathcal{Y}_1(\mathbf{x}) {}_{-s} a_1. \quad (\text{C5})$$

These are then related to the E - and B -mode coefficients as:

$${}_s E_1 \equiv -\frac{1}{2} [{}_s a_1 + (-1)^s {}_{-s} a_1], \quad (\text{C6})$$

$$i {}_s B_1 \equiv -\frac{1}{2} [{}_s a_1 - (-1)^s {}_{-s} a_1]. \quad (\text{C7})$$

Note the preceding $(-)$ sign. For scalar fields ($s \equiv 0$) the E and B modes are defined omitting that sign.

In analogy with our curved-sky nomenclature, let us now write a as a vector such that in real space $\mathbf{a}(\mathbf{x}) \equiv (\text{Re}(a), \text{Im}(a))$, and in Fourier space $\mathbf{a}_1 \equiv ({}_s E_1, {}_s B_1)$. We can rewrite the equation (C5) in vectorial form:

$$\mathbf{a}(\mathbf{x}) \equiv \int \frac{dL^2}{2\pi} {}_s E_1(\mathbf{x}) \mathbf{a}_1, \quad \mathbf{a}_1 \equiv \int \frac{d\mathbf{x}^2}{2\pi} {}_s E_1^\dagger(\mathbf{x}) \mathbf{a}(\mathbf{x}), \quad (\text{C8})$$

where, in analogy with equations (A3) and A4, we have defined the matrix basis functions:

$$\begin{aligned} {}_s E_1 &\equiv -\frac{1}{2} \begin{bmatrix} {}_s \mathcal{Y}_1 + (-1)^s {}_{-s} \mathcal{Y}_1 & i({}_s \mathcal{Y}_1 - (-1)^s {}_{-s} \mathcal{Y}_1) \\ -i({}_s \mathcal{Y}_1 - (-1)^s {}_{-s} \mathcal{Y}_1) & {}_s \mathcal{Y}_1 + (-1)^s {}_{-s} \mathcal{Y}_1 \end{bmatrix} \\ &= -\frac{1}{2l^s} \begin{pmatrix} \bar{\partial}^s + \bar{\partial}^s & i(\bar{\partial}^s - \bar{\partial}^s) \\ -i(\bar{\partial}^s - \bar{\partial}^s) & \bar{\partial}^s + \bar{\partial}^s \end{pmatrix} e^{i\mathbf{l}\mathbf{x}} \\ &= -i^s \begin{pmatrix} \cos(s\varphi_l) & \sin(s\varphi_l) \\ -\sin(s\varphi_l) & \cos(s\varphi_l) \end{pmatrix} e^{i\mathbf{l}\mathbf{x}} \\ &= -i^s \mathbf{R}^\dagger(s\varphi_l) e^{i\mathbf{l}\mathbf{x}}. \end{aligned} \quad (\text{C9})$$

Here, $\mathbf{R}(\varphi)$ is a rotation matrix.

C2 Discrete description and DFTs

As we discuss in Section 2.5.1, it is more convenient to describe power-spectrum estimation methods in a discretized flat sky. Let the patch of the sky under inspection be contained by a rectangle of sides L_x and L_y (in units of radians), and let us discretize this rectangle by dividing it into an $N_x \times N_y$ grid with pixels of area $\Delta\mathbf{x}^2 \equiv \Delta x \Delta y = (L_x/N_x)(L_y/N_y)$. Each pixel in this grid is then labelled by a pair of integers $\mathbf{n} \equiv (n_x, n_y)$, and is assigned coordinates $\mathbf{x}_\mathbf{n} \equiv (n_x \Delta x, n_y \Delta y)$. Each field in the pixelized map $\mathbf{a}(\mathbf{x})$ is therefore defined for $n_x \in [0, N_x - 1]$, $n_y \in [0, N_y - 1]$.

In this case, the spin Fourier transform (equation C9) of the pixelized field can be computed as its DFT:

$$\mathbf{a}_\mathbf{k} \equiv \mathcal{D}(\mathbf{a})_\mathbf{k}^{s_a} \equiv \sum_{\mathbf{x}} \frac{\Delta\mathbf{x}^2}{2\pi} {}_s E_\mathbf{k}^\dagger(\mathbf{x}) \mathbf{a}_\mathbf{x}, \quad (\text{C10})$$

where the wavenumber \mathbf{k} is now discretized as $\mathbf{k} = (j_x \Delta k_x, j_y \Delta k_y)$, the integers $j_{(x,y)}$ run from $-N_{(x,y)}/2$ to $N_{(x,y)}/2 - 1$ ¹² and the pixel size is $\Delta k_{(x,y)} \equiv 2\pi/L_{(x,y)}$.

The following properties of the DFT are worth recalling:

(i) Periodicity:

$$\begin{aligned} \mathbf{a}_{(k_x, k_y)} &= \mathbf{a}_{(k_x + N_x \Delta k_x, k_y)} \\ &= \mathbf{a}_{(k_x, k_y + N_y \Delta k_y)} \\ &= \mathbf{a}_{(k_x + N_x \Delta k_x, k_y + N_y \Delta k_y)}. \end{aligned} \quad (\text{C11})$$

(ii) For a real-valued scalar a , its DFT satisfies $a_{\mathbf{k}_j}^* = a_{\mathbf{k}_{N-j}}$ (where $\mathbf{N} \equiv (N_x, N_y)$).

(iii) The orthogonality relation of the basis functions now takes the form:

$$\sum_{\mathbf{x}} {}_s E_1^\dagger(\mathbf{x}) {}_s E_\mathbf{k}(\mathbf{x}) = N_x N_y \delta_{\mathbf{l}, \mathbf{k}} \hat{\mathbf{l}}. \quad (\text{C12})$$

¹²This is the valid domain when $N_{(x,y)}$ is even. For odd $N_{(x,y)}$, the interval becomes $[-(N_{(x,y)} - 1)/2, (N_{(x,y)} - 1)/2]$.

(iv) The power spectrum of a pixelized field is defined as:

$$\langle \mathbf{a}_1 \mathbf{b}_1^\dagger \rangle \equiv \frac{\delta_{\mathbf{l}, \mathbf{k}}}{\Delta k^2} C_1^{ab}. \quad (\text{C13})$$

APPENDIX D: MODE-COUPLING MATRICES

For curved skies, and different combinations of spin-0 and spin-2 fields, the mode-coupling matrices introduced in equation (9) are given by (Hivon et al. 2002; Kogut et al. 2003):

$$M_{\ell\ell'}^{00} = \frac{2\ell' + 1}{4\pi} \sum_{\ell''} P_{\ell''}^{vw} \begin{pmatrix} \ell & \ell' & \ell'' \\ 0 & 0 & 0 \end{pmatrix}^2 \mathbf{1}, \quad (\text{D1})$$

$$M_{\ell\ell'}^{02} = M_{\ell\ell'}^{0+} \mathbf{1}, \quad (\text{D2})$$

$$M_{\ell\ell'}^{0+} = \frac{2\ell' + 1}{4\pi} \sum_{\ell''} P_{\ell''}^{vw} \begin{pmatrix} \ell & \ell' & \ell'' \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \ell & \ell' & \ell'' \\ 2 & -2 & 0 \end{pmatrix},$$

$$M_{\ell\ell'}^{22} = \begin{pmatrix} M_{\ell\ell'}^+ & 0 & 0 & M_{\ell\ell'}^- \\ 0 & M_{\ell\ell'}^+ & -M_{\ell\ell'}^- & 0 \\ 0 & -M_{\ell\ell'}^- & M_{\ell\ell'}^+ & 0 \\ M_{\ell\ell'}^- & 0 & 0 & M_{\ell\ell'}^+ \end{pmatrix},$$

$$M_{\ell\ell'}^\pm = \frac{2\ell' + 1}{4\pi} \sum_{\ell''} P_{\ell''}^{vw} \begin{pmatrix} \ell & \ell' & \ell'' \\ 2 & -2 & 0 \end{pmatrix}^2 \frac{1 \pm (-1)^{\ell+\ell'+\ell''}}{2}, \quad (\text{D3})$$

where $\mathbf{1}$ is the unit matrix, and $P_\ell^{vw} \equiv (2\ell + 1) \text{PCL}_\ell(v, w)$.

In the flat sky, the unbinned mode-coupling matrices introduced in equation (37) are given by:

$$M_{\mathbf{l}\mathbf{k}}^{00} \equiv \frac{(2\pi)^2}{L_x^2 L_y^2} v_{\mathbf{l}-\mathbf{k}} w_{\mathbf{l}-\mathbf{k}}^*, \quad (\text{D4})$$

$$M_{\mathbf{l}\mathbf{k}}^{02} \equiv \frac{(2\pi)^2}{L_x^2 L_y^2} v_{\mathbf{l}-\mathbf{k}} w_{\mathbf{l}-\mathbf{k}}^* \begin{pmatrix} c & -s \\ s & c \end{pmatrix} \quad (\text{D5})$$

$$M_{\mathbf{l}\mathbf{k}}^{22} \equiv \frac{(2\pi)^2}{L_x^2 L_y^2} v_{\mathbf{l}-\mathbf{k}} w_{\mathbf{l}-\mathbf{k}}^* \begin{pmatrix} c^2 & -cs & -cs & s^2 \\ cs & c^2 & -s^2 & -cs \\ cs & -s^2 & c^2 & -cs \\ s^2 & cs & cs & c^2 \end{pmatrix}, \quad (\text{D6})$$

where $c \equiv \cos 2\Delta\varphi$ and $s \equiv \sin 2\Delta\varphi$.

APPENDIX E: FLAT-SKY PSEUDO- C_ℓ S IN THE CONTINUUM LIMIT

In the continuum limit, the Fourier coefficients of the masked field (see equation 34) are:

$$\begin{aligned} \mathbf{a}_1^v &= \int \int \frac{d\mathbf{k}^2 d\mathbf{q}^2}{2\pi} \left[\int \frac{d\mathbf{x}^2}{(2\pi)^2} {}_s E_1^\dagger(\mathbf{x}) {}_s E_\mathbf{k}(\mathbf{x}) {}_0 E_\mathbf{q}(\mathbf{x}) \right] \mathbf{a}_\mathbf{k} v_\mathbf{q} \\ &= \int \frac{d\mathbf{k}^2}{2\pi} \mathbf{R}(s_a \Delta\varphi) \mathbf{a}_\mathbf{k} v_{\mathbf{l}-\mathbf{k}}, \end{aligned} \quad (\text{E1})$$

The covariance of the Fourier coefficients of two masked fields is then given by:

$$\langle \mathbf{a}_1^v \mathbf{b}_1^{w\dagger} \rangle = \int \frac{d\mathbf{k}^2}{(2\pi)^2} \mathbf{R}(s_a \Delta\varphi) C_\mathbf{k}^{ab} \mathbf{R}^\dagger(s_b \Delta\varphi) v_{\mathbf{l}-\mathbf{k}} w_{\mathbf{l}-\mathbf{k}}^*. \quad (\text{E2})$$

Defining the pseudo- C_ℓ as the normalized angular average of $\mathbf{a}_1 \mathbf{b}_1^\dagger$, we obtain:

$$\begin{aligned} \langle \text{PCL}_\ell(\mathbf{a}, \mathbf{b}) \rangle &\equiv \frac{(2\pi)^2}{S} \int \frac{d\varphi_l}{2\pi} \langle \mathbf{a}_1^v \mathbf{b}_1^{w\dagger} \rangle \\ &= \int \frac{k dk q dq}{(2\pi)^2} \left[\frac{(2\pi)^2}{S} \int \frac{d\varphi_q d\varphi_k d\varphi_l}{2\pi} v_{\mathbf{q}} w_{\mathbf{q}}^* \right. \\ &\quad \left. \times R(s_a \Delta\varphi) C_k^{ab} R^\dagger(s_b \Delta\varphi) \delta(\mathbf{q} - \mathbf{l} + \mathbf{k}) \right], \quad (\text{E3}) \end{aligned}$$

where S is the observed sky area (in steradian), and where we have eliminated the dependence on $\mathbf{l} - \mathbf{k}$ by introducing an additional integral over $d\mathbf{q}^2 \delta(\mathbf{q} - \mathbf{l} + \mathbf{k})$.

As shown in Asgari et al. (2016), these expressions can be simplified through the following steps:

(i) Substitute

$$\delta(\mathbf{q} - \mathbf{l} + \mathbf{k}) \rightarrow \int \frac{d\mathbf{r}^2}{(2\pi)^2} e^{i(\mathbf{q}-\mathbf{l}+\mathbf{k})\mathbf{r}} \quad (\text{E4})$$

(ii) Integrate over the angular parts of \mathbf{r} , \mathbf{l} , and \mathbf{k} using the following relation:

$$\int_0^{2\pi} d\varphi e^{i x \cos \varphi} e^{i n \varphi} = 2\pi i^n J_n(x), \quad (\text{E5})$$

where J_n is the cylindrical Bessel function of order n .

(iii) Integrate over the angular part of \mathbf{q} , defining the power spectrum of the masks as:

$$\tilde{C}_{\mathbf{q}}^{vw} \equiv \frac{(2\pi)^2}{S} \int d\varphi_{\mathbf{q}} v_{\mathbf{q}} w_{\mathbf{q}}^* \quad (\text{E6})$$

(iv) Solve the last isolated integral over the radial part of \mathbf{r} by using the following relation:

$$\int_0^\infty dr r J_0(qr) J_n(kr) J_n(lr) = \frac{\cos n\theta}{\pi k l \sin \theta}, \quad (\text{E7})$$

where θ is the angle between sides ℓ and k of the triangle formed by three sides of length q , ℓ , and k .

(v) Make the change of variables $q^2(\ell, k, \theta) \equiv \ell^2 + k^2 - 2k\ell \cos \theta$ to simplify the integral over the radial part of \mathbf{q} .

This yields the following relation analogous to equation (9):

$$\text{vec} [C_l^{ab}] = \int_0^\infty dk M_{\ell k}^{s_a s_b} \cdot \text{vec} [C_k^{ab}], \quad (\text{E8})$$

where:

$$M_{\ell k}^{00} = \frac{k}{2\pi} \int_0^\pi \frac{d\theta}{\pi} \tilde{C}_{q(\ell, k, \theta)}^{vw} \quad (\text{E9})$$

$$M_{\ell \ell'}^{02} = M_{\ell \ell'}^{0+} \hat{\mathbf{1}}, \quad M_{\ell k}^{0+} = \frac{k}{2\pi} \int_0^\pi \frac{d\theta}{\pi} \tilde{C}_{q(\ell, k, \theta)}^{vw} \cos 2\theta \quad (\text{E10})$$

$$M_{\ell \ell'}^{22} = \begin{pmatrix} M_{\ell \ell'}^{++} & 0 & 0 & M_{\ell \ell'}^{--} \\ 0 & M_{\ell \ell'}^{++} & -M_{\ell \ell'}^{--} & 0 \\ 0 & -M_{\ell \ell'}^{--} & M_{\ell \ell'}^{++} & 0 \\ M_{\ell \ell'}^{--} & 0 & 0 & M_{\ell \ell'}^{++} \end{pmatrix} \quad (\text{E11})$$

$$M_{\ell \ell'}^{\pm\pm} = \frac{k}{2\pi} \int_0^\pi \frac{d\theta}{\pi} \tilde{C}_{q(\ell, k, \theta)}^{vw} \frac{1 \pm \cos 4\theta}{2}. \quad (\text{E12})$$

These can also be expressed as integrals over q :

$$M_{\ell k}^{XY} = \frac{k}{2\pi} \int \frac{q dq}{2\pi} \tilde{C}_{\mathbf{q}}^{vw} F(\ell, k, q) G^{XY}(\ell, k, q), \quad (\text{E13})$$

where $F(\ell, k, q) = 0$ if ℓ , k , and q do not form a triangle, and

$$F(\ell, k, q) \equiv \frac{4}{\sqrt{2\ell^2 k^2 + 2k^2 q^2 + 2q^2 \ell^2 - \ell^4 - k^4 - q^4}}$$

otherwise. The functions G^{XY} are given by

$$G^{00} \equiv 1, \quad (\text{E14})$$

$$G^{0+}(\ell, k, q) \equiv \frac{\ell^4 + k^4 + q^4 - 2k^2 q^2 - 2\ell^2 q^2}{2k^2 \ell^2}, \quad (\text{E15})$$

$$G^{++}(\ell, k, q) \equiv \left[\frac{\ell^4 + k^4 + q^4 - 2k^2 q^2 - 2\ell^2 q^2}{2k^2 \ell^2} \right]^2, \quad (\text{E16})$$

$$G^{--}(\ell, k, q) \equiv 1 - G^{++}(\ell, k, q). \quad (\text{E17})$$

Since in the flat-sky limit ℓ is a continuous variable, bandpowers $B_{\mathbf{q}}^{ab}$ can be defined as averages over a given interval in ℓ , $[\ell_{\mathbf{q}}^{\min}, \ell_{\mathbf{q}}^{\max}]$:

$$B_{\mathbf{q}}^{ab} \equiv \int_{\ell_{\mathbf{q}}^{\min}}^{\ell_{\mathbf{q}}^{\max}} \frac{d\ell}{\ell_{\mathbf{q}}^{\max} - \ell_{\mathbf{q}}^{\min}} C_{\ell}^{ab}. \quad (\text{E18})$$

The binned coupling matrix is therefore given by:

$$\mathcal{M}_{\mathbf{q}\mathbf{q}'}^{s_a s_b} \equiv \int_{\ell_{\mathbf{q}}^{\min}}^{\ell_{\mathbf{q}}^{\max}} \frac{d\ell}{\ell_{\mathbf{q}}^{\max} - \ell_{\mathbf{q}}^{\min}} \int_{\ell_{\mathbf{q}'}^{\min}}^{\ell_{\mathbf{q}'}^{\max}} d\ell' M_{\ell \ell'}^{s_a s_b}. \quad (\text{E19})$$

The main complication in using the pseudo- C_ℓ formalism in the continuum limit is the need to compute the angle-averaged mask pseudo power spectrum C_ℓ^{vw} . Flat fields are most easily analysed when pixelized in a Cartesian grid, and under this set-up the mask power spectrum is not a well-defined quantity for infinitesimally small intervals of ℓ . This leads to non-negligible biases and poor performance (Asgari et al. 2016) due to the need to use highly resolved finite intervals to compute the integrals presented in the previous section. This motivates the discrete formalism used throughout this paper, which connects directly with the storage format of flat-sky maps.

APPENDIX F: NOISE BIAS CALCULATION

Different strategies can be followed to estimate the pseudo- C_ℓ noise bias:

Homogeneous noise: if the noise is homogeneous, and can be described by a known noise power spectrum N_ℓ , the procedure is simple: first, convolve the known noise power spectrum with the pseudo- C_ℓ mode-coupling matrix. Then, estimate the contribution from contaminant deprojection by using N_ℓ in lieu of C_ℓ^{ab} in equations (23)–(25). The total noise bias is then the sum of both contributions, corrected for the pseudo- C_ℓ convolution using equation (15).

Uncorrelated noise: if the noise is inhomogeneous but uncorrelated (i.e. $\langle \mathbf{n}(\hat{\theta}) \mathbf{n}^\dagger(\hat{\theta}') \rangle = \mathbb{1} \sigma_n^2(\hat{\theta}) \delta(\hat{\theta}, \hat{\theta}')^{13}$), it is also possible to find an analytical estimate of the noise bias.

¹³Here, $\delta(\hat{\theta}, \hat{\theta}')$ is the Dirac δ -function on the sphere, and σ_n^2 is the local noise variance in one steradian.

(i) It is not difficult to prove that the pseudo- C_ℓ of the noise contribution is given by

$$\text{PCL}_\ell(\mathbf{n}, \mathbf{n}) = \mathbb{1} \int \frac{d\hat{\theta}}{4\pi} v^2(\hat{\theta}) \sigma_n^2(\hat{\theta}) \quad (\text{F1})$$

$$= \mathbb{1} A^2 \int \frac{d\hat{\theta}}{4\pi} \sigma_n^{-2}(\hat{\theta}), \quad (\text{F2})$$

where, in the second line, we have assumed that we use an inverse-variance noise weighting scheme, in which the weights map is $v(\hat{\theta}) = A \sigma_n^{-2}(\hat{\theta})$.

(ii) The additional bias from contaminant deprojection can also be computed analytically, starting from equation (22), as:

$$\begin{aligned} \Delta C_\ell^{mn} = & -2F_{ij} \text{PCL}_\ell(f^i, v^2 \sigma_n^2 f^j) \\ & + F_{ij} F_{rs} \left[\int d\hat{\theta} v^2 \sigma_n^2 f^{j\dagger} f^r \right] \text{PCL}_\ell(f^i, f^s). \end{aligned} \quad (\text{F3})$$

F1 General case

In general, if the noise is both non-white and inhomogeneous, the total noise bias can be estimated by averaging the result of applying the pseudo- C_ℓ estimator to a large number of noise realizations with the same noise properties as the data, assuming that those properties are sufficiently well characterized to produce those simulations. If that is not the case, a common approach to avoiding the noise bias altogether is to use only cross-correlations of data splits with independent noise contributions.

APPENDIX G: FLAT-SKY VALIDATION

In order to validate the flat-sky features of NAMASTER, we follow the same procedure used in Sections 4.2 and 4.3 for the LSS and CMB validation suites, respectively.

Before we discuss these results, a technical point, specific to the flat-sky case, must first be addressed. The theory prediction for the pseudo- C_ℓ method is given by equation (37) [before multiplying by the inverse binned mode-coupling matrix, equation (38)]. This involves a convolution of the power spectra interpolated into the

discretized Fourier plane with the unbinned mode-coupling matrix. This is an $\mathcal{O}(N_{\text{pix}}^2)$ operation that would be too expensive to perform at every step of a Monte Carlo chain, and therefore NAMASTER uses a different approximate approach. This consists of first interpolating the theory power spectrum into the discrete Fourier plane, followed by an averaging of the interpolated power spectrum into a set of radial rings with a width given by the minimum wavenumber probed (given by $\min(2\pi/L_x, 2\pi/L_y)$), and spanning all possible values of $|l|$. This binned power spectrum, an object of size $\sim \sqrt{N_{\text{pix}}}$, is then convolved with a high-resolution version of the mode-coupling matrix. This high-resolution matrix is given by a variation of equation (38) in which S_q represents all the values of the wavenumber \mathbf{k} in one of the narrow rings described above (instead of the final wide bandpowers). The convolved power spectrum is then multiplied by the inverse binned coupling matrix [given by equation (38) exactly] to produce the final theoretical prediction. As we show below, this procedure, which is $\mathcal{O}(N_{\text{pix}})$, is able to reproduce the exact theoretical prediction with a negligible error in all cases explored here.

Fig. G1 shows, with solid lines, the mean over all simulations of the six LSS power spectra and the three polarized CMB power spectra in the left- and right-hand panels, respectively. The analytical predictions for the non-zero input power spectra are shown as dashed lines in both cases, and agree with the simulation mean almost perfectly. The mean residuals normalized by the 1σ error on the mean are shown in the lower half of both figures, and feature fluctuations with amplitudes smaller than $\sim 3\sigma$ – 4σ .

Again, we quantify the presence of a bias in the estimator by studying the distribution of the χ^2 values across simulations. In analogy with Figs 9 and 13, these are shown in the upper and lower panels of Fig. G2 for the LSS and CMB suites, respectively. In all cases, the distributions (blue histograms) agree well with the expected χ^2 distribution (solid lines) for the appropriate number of degrees of freedom (vertical dashed lines). The χ^2 values of the mean residuals shown in the lower panels of Fig. G1 and marked by vertical dotted–dashed lines in Fig. G2 have PTE values above 10 per cent. We therefore do not observe any significant bias in the estimator, proving NAMASTER’s usefulness in the analysis of current and next-generation experiments in situations requiring the use of the flat-sky approximation

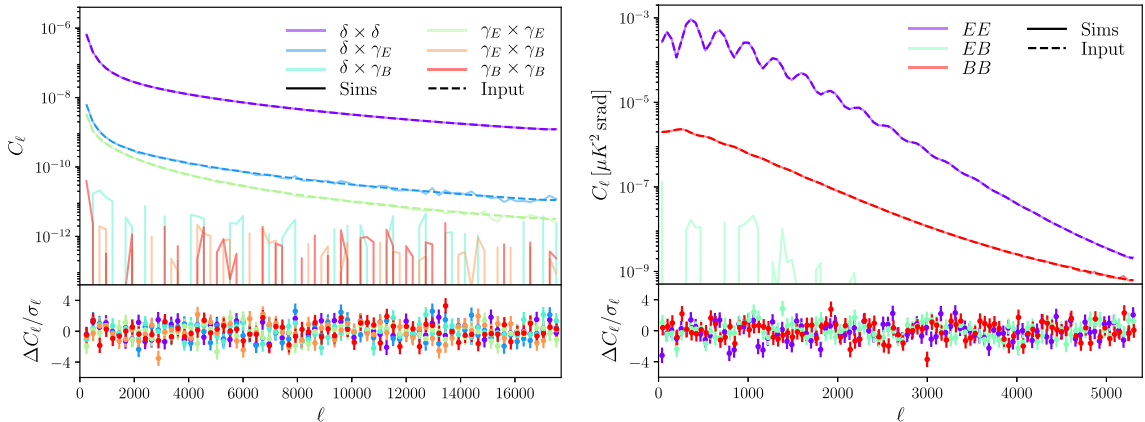


Figure G1. Left and right: the analogues of the left-hand panels of Figs 8 and 10 for the flat-sky LSS and CMB validation sets, respectively.

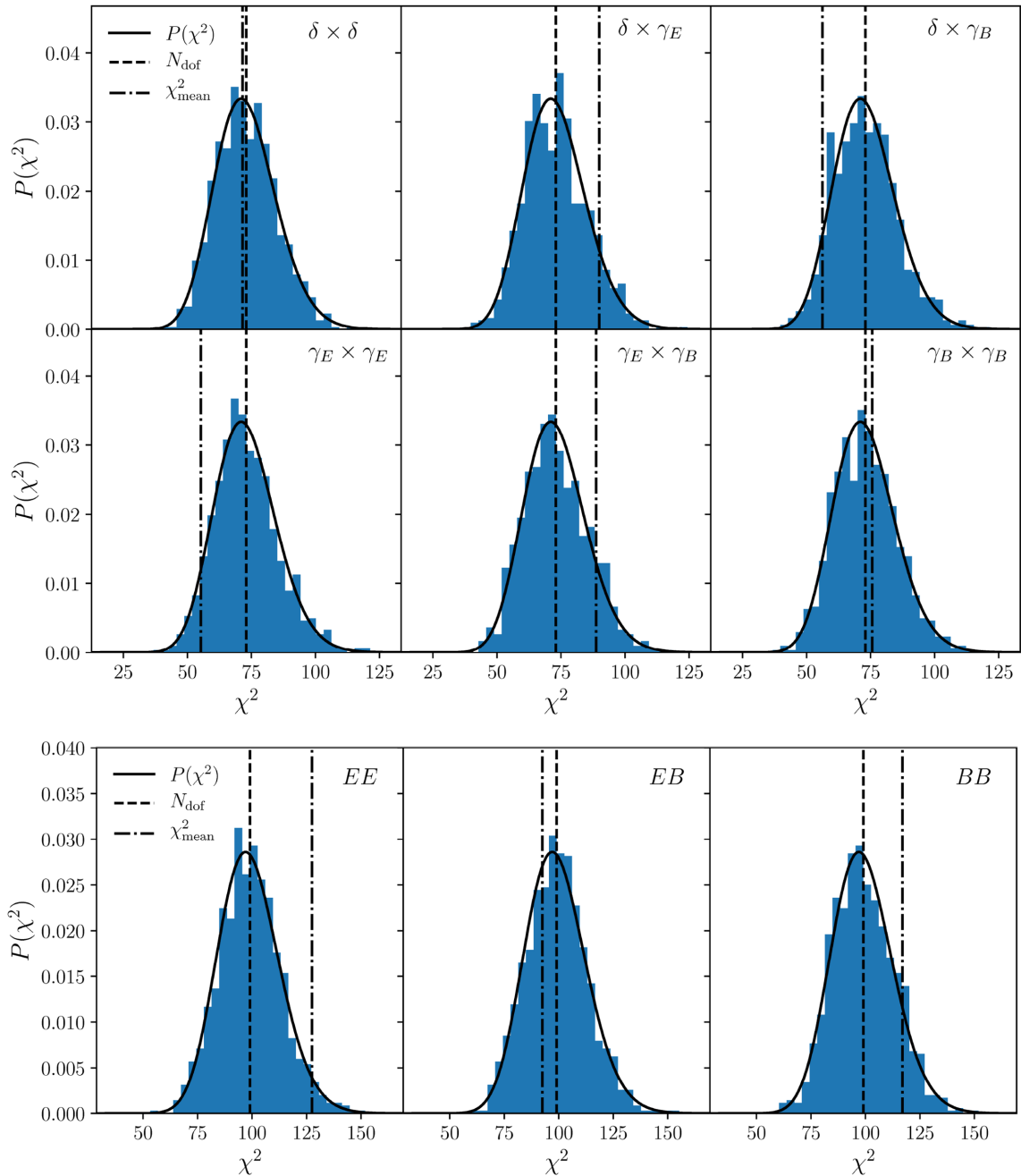


Figure G2. Top and bottom: the analogues of Figs 9 and 13 for the flat-sky LSS and CMB validation sets, respectively.

APPENDIX H: CURVED-SKY VALIDATION FOR DIFFERENT PIXEL RESOLUTIONS

We checked that NaMaster returns unbiased power spectra for all possible values of the HEALPIX resolution parameter N_{side} . In this appendix, we show that we recover the correct power spectra using the inputs from our curved-sky LSS validation suite (see

Section 4.1.1). In particular, we tested N_{side} ranging from 8 to 8192. For the latter, the computational resources needed make it very challenging to generate and analyse a large enough number of simulations to have a good statistical power. However, we generate 100 simulations with $N_{\text{side}} = 4096$ and show the results in Fig. H1. We can see that the power spectra are recovered without any noticeable bias in the expected range of validity $\ell \lesssim 2N_{\text{side}}$.

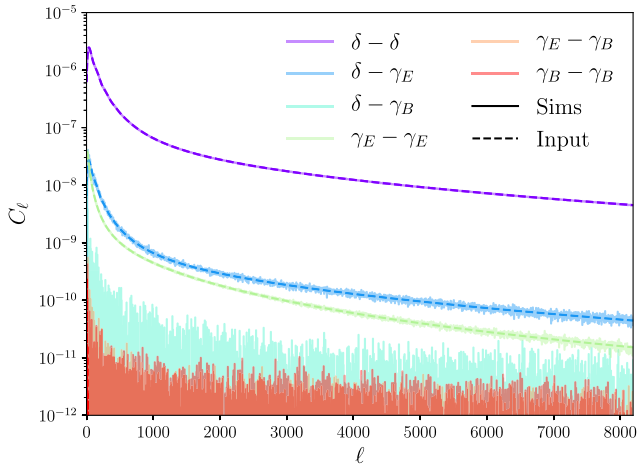


Figure H1. Input (dashed darker lines) and mean measured (solid semi-transparent lines) power spectra for simulations HEALPIX $N_{\text{side}} = 4096$ using the curved-sky LSS validation suite (see Section 4.1.1).

APPENDIX I: PERFORMANCE AND SCALABILITY

One of the known advantages of the pseudo- C_ℓ estimator is that it is usually much faster than other known estimators (Elsner et al. 2017) since the complexity scales as $\mathcal{O}(\ell_{\text{max}}^3)$, where ℓ_{max} is the maximum multipole at which the power spectra are calculated. In order to check how NAMASTER execution time changes with ℓ_{max} , we compute the curved-sky angular auto- and cross-power spectra of a spin-0 and a spin-2 fields, with no contaminants. We change the resolution of the HEALPIX maps by varying the N_{side} parameter in the range [16, 2048], and we calculate the power spectra up to $\ell_{\text{max}} = 3N_{\text{side}}$. We use one interactive Cori-Haswell node at the NERSC. This kind of nodes have two 16-core Intel Xeon (E5-2698 v3) processors at 2.3 GHz, and 128 GB DDR4 of RAM. We executed the code using 32 OpenMP threads. In the top panel of Fig. 11, we can see the results for this test. We show the results for the auto-power spectra for the spin-0 and spin-2 fields (Data 00, and Data 22, respectively), as well as the results for the cross-power spectra (Data 02). In all cases, we find that the code scales as expected (the solid red line shows the expected $\propto \ell_{\text{max}}^3$ behaviour). We can also see that, for very small N_{side} , data handling time starts to be comparable to the power-spectrum computation time, increasing the execution time with respect to the expectation.

Not all parts of the power-spectrum estimation take the same amount of time. For example, even though the scaling with ℓ_{max} is the same, the computation of the mode-coupling matrix typically takes longer than a performing single SHT. Computation times can be significantly increased if deprojection over a large number of maps is required or if E/B purification is required. The bottom panel of Fig. 11 shows the execution times associated to the most relevant types of operations. These are (using the labels provided in the figure):

- (i) *Field* $s = 0$ and $s = 2$: generating a `NmtField` object. This entails taking the SHT of a given spin-0 or spin-2 map as well as that of its associated weights map.
- (ii) *Deproj.*: generating a `NmtField` object with contaminant deprojection. This is the same as the above in addition to computing the SHT of a set of contaminant templates. In this case, we include

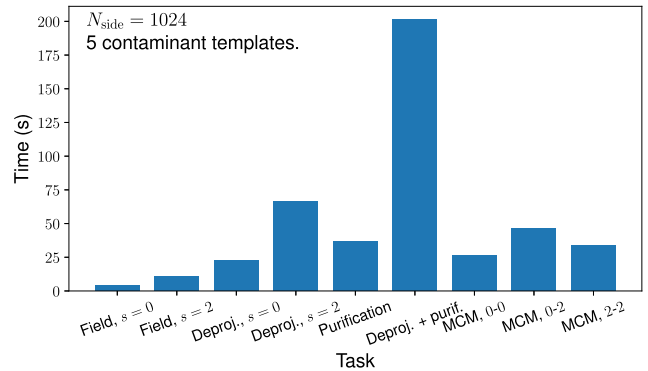
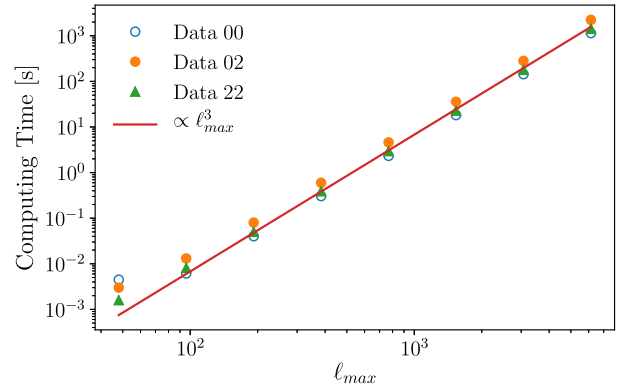


Figure 11. Top panel: computing time for the auto-power spectra of curved-sky spin-0 (data 00, hollow circles), and spin-2 (data 22, solid triangles) contamination-free maps, as well as their cross-power (data 02, solid circles) as a function of ℓ_{max} . We include the expected scaling $\propto \ell_{\text{max}}^3$ for the pseudo- C_ℓ estimator (solid red line) for comparison. Bottom panel: execution time for different tasks associated with the pseudo- C_ℓ estimator. We show times for fields with resolution $N_{\text{side}} = 1024$, which can be extrapolated to other resolutions using a $\propto N_{\text{side}}^3$ scaling, as described above. The description of the different tasks shown is given in the main text.

five templates, and the additional computation time will increase linearly with the number of templates.

(iii) *Purification*: generating a `NmtField` object with B -mode purification. This is the same as ‘Field $s = 2$ ’ in addition to computing the pure B modes. This entails estimating a more accurate SHT of the apodized mask as well as a combination of harmonic- and real-space operations described in Section 2.3.

(iv) *Deproj. + purif.*: generating a `NmtField` object with both B -mode purification and contaminant deprojection. This entails all of the operations implied by ‘Deproj.’ and ‘Purification’ in addition to the computation of the B -mode-purified templates [see equations (24) and (25)].

(v) *MCM* s_1-s_2 : generating the pseudo- C_ℓ mode-coupling matrix for a spin- s_1 and a spin- s_2 field.

The times reported in this figure correspond to the time required to perform a single instance of the corresponding operation for fields with resolution $N_{\text{side}} = 1024$ using 16 OpenMP threads. Note that, even though the computation of the mode-coupling matrix dominates the total execution time in the simplest scenarios, in the most complex situation, involving both purification and deprojection, setting up a single field can take significantly longer.

This paper has been typeset from a \LaTeX file prepared by the author.