arXiv:astro-ph/0012120 v2 7 Dec 2000

How to measure CMB polarization power spectra without losing information

Max Tegmark & Angelica de Oliveira-Costa
Dept. of Physics, Univ. of Pennsylvania, Philadelphia, PA 19104; max@physics.upenn.edu
(Submitted to Phys. Rev. D. December 7, 2000)

We present a method for measuring CMB polarization power spectra given incomplete sky coverage and test it with simulated examples such as Boomerang 2001 and MAP. By augmenting the quadratic estimator method with an additional step, we find that the E and B power spectra can be effectively disentangled on angular scales substantially smaller than the width of the sky patch in the narrowest direction. We find that the basic quadratic and maximum-likelihood methods display a unneccesary sensitivity to systematic errors when $T-E$ cross-correlation is involved, and show how this problem can be eliminated at negligible cost in increased error bars. We also test numerically the widely used approximation that sample variance scales inversely with sky coverage, and find it to be an excellent approximation on scales substantially smaller than the sky patch.

98.62.Py, 98.65.Dx, 98.70.Vc, 98.80.Es

I. INTRODUCTION

As experimental groups roar ahead to map the CMB intensity with increasing resolution and sensitivity, a second parallel front is being opened up: CMB polarization. Theoretical advances now allow model predictions for CMB polarization to be computed with exquisite accuracy [1–3], and it is known that polarization measurements can substantially improve the accuracy with which parameters can be measured over the no-polarization case by breaking the degeneracy between certain parameter combinations [4–6]. CMB polarization can also provide crucial cross-checks and tests of the underlying theory. After placing ever tighter upper limits [7,8], a number of experimental teams are likely to make the first detections of CMB polarization in the next year or two, about a decade after unpolarized CMB fluctuations were first detected [9]. It is therefore timely to address real-world issues related to extracting polarized power spectra from experiments with incomplete sky coverage and complicated noise. This is the purpose of the present paper, with emphasis on the problem of separating the two polarization signals known as E and B [1,2]. Important steps in this direction were first taken in [10,11], for the special case of experiments mapping circles in the sky.

The E/B problem goes back to the fact that the polarization tensor field on the sky can be separated into a curl-free and a divergence free component, and is most naturally expressed in terms of two scalar fields, denoted E and B by analogy with electromagnetism [1,2]. Not only does this separation eliminate the coordinate system dependence that plagues the familiar Stokes parameters, but E and B also probe distinct physical effects, making them the natural meeting point for theory and observation. Unfortunately, the correspondence between E and B and the measured Stokes parameters is not spatially local, involving a partial differential equation, which means that it is not possible to uniquely recover E and B from a map with merely partial sky coverage. This issue is particularly important since the B-signal from inflation-produced gravity waves potentially offers a unique probe of ultra-high-energy physics, but may be swamped by a leakage from a larger E-signal [12,13].

A key goal of CMB analysis is to constrain cosmological models, and information-theoretical methods have been frequently employed in the literature to study how accurately this can be done in principle with a given data set, using the Fisher information matrix formalism [14,15]. For CMB polarization experiments, this has been useful both for optimizing experimental design [16,17] and for accuracy forecasting in general [4]. These information-theoretical tools are equally useful for data analysis, since they provide a simple way of checking whether cosmological information is being lost in the data analysis pipeline. Each step in such a pipeline typically compresses the input data into a smaller set of numbers, and if the output can be shown to retain all the cosmological information of the input, the method is said to be lossless. Lossless methods have been developed and extensively tested for the unpolarized case, for both mapmaking [18,19] and power spectrum estimation [20,21]. As we will see, it is possible to draw heavily on these methods for the polarized case as well, although a number of adaptations make them simpler to interpret and help improve the E/B-separation and robustness towards systematic errors.

The rest of this paper is organized as follows. In Section II, we discuss basic methods involved in polarization data analysis. To keep the presentation from becoming too abstract, many method details and extensions are postponed to Section III, where they are illustrated with plots from actual applications. This section also assesses the effectiveness of the various methods numerically, with applications to five experimental examples. A step-by-step summary of how to compute the signal covariance matrix is given in Appendix A for the reader wishing to do so in practice. Our conclusions are summarized in Section IV.

II. METHOD

In this section, we first establish some basic notation, then discuss the extraction of maps from raw time-ordered data, and finally cover the extraction of power spectra from maps.

A. Notation

The linear polarization pattern of the CMB sky is characterized by the two Stokes parameters Q and U in each sky direction. Q and U are the components of a rank 2 tensor (spinor), loosely speaking a vector without an arrow on it, so Q and U maps are defined only relative to a convention providing a reference direction at each point in the sky. Let us discretize the T, Q and U maps into N pixels centered at unit vectors $\widehat{\mathbf{r}}_1, ..., \widehat{\mathbf{r}}_N$, and write $T_i = T(\widehat{\mathbf{r}}_i)$, $Q_i = Q(\widehat{\mathbf{r}}_i)$, $U_i = U(\widehat{\mathbf{r}}_i)$ (we let T denote the unpolarized intensity, often called I). We group these numbers into three N-dimensional vectors $\mathbf{T}$, $\mathbf{Q}$ and $\mathbf{U}$ and group these in turn into a single $3N$-dimensional vector

$$\mathbf{x} \equiv \begin{pmatrix} \mathbf{T} \\ \mathbf{Q} \\ \mathbf{U} \end{pmatrix}. \tag{1}$$

The statistical properties of $\mathbf{x}$ have been computed in full detail in the literature [1,2], and are characterized by four separate power spectra: C_ℓ^T for the unpolarized signal T, C_ℓ^E for the E-polarization, C_ℓ^B for the B-polarization and C_ℓ^X for the cross-correlation between T and E. We will find it useful for data analysis purposes to recast the polarization problem in exactly the same mathematical form as the simpler unpolarized case, encoding all complications in a set of matrices. The vector $\mathbf{x}$ has a vanishing expectation value ($\langle\mathbf{x}\rangle = \mathbf{0}$), and we can write its covariance matrix as

$$\mathbf{C} \equiv \langle\mathbf{x}\mathbf{x}^t\rangle = \sum_i p_i \mathbf{P}_i \tag{2}$$

for a set of parameters p_i and known matrices $\mathbf{P}_i$. If the four observed power spectra are negligibly small for all multipoles ℓ above some value $\ell_{\rm max}$ (which is always the case because of the smoothing caused by the finite angular resolution of an experiment), then we define these parameters to be

$$\begin{cases} p_1, ..., p_{\ell_{\rm max}-1} & = \delta T^2{}_2^T, ..\delta T^2{}_{\ell_{\rm max}}^T, \\ p_{\ell_{\rm max}}, ..., p_{2\ell_{\rm max}-2} & = \delta T^2{}_2^X, ..\delta T^2{}_{\ell_{\rm max}}^X, \\ p_{2\ell_{\rm max}-1}, ..., p_{3\ell_{\rm max}-3} & = \delta T^2{}_2^E, ..\delta T^2{}_{\ell_{\rm max}}^E, \\ p_{3\ell_{\rm max}-2}, ..., p_{4\ell_{\rm max}-4} & = \delta T^2{}_2^B, ..\delta T^2{}_{\ell_{\rm max}}^B, \\ p_{4\ell_{\rm max}-3} & = \eta = 1. \end{cases} \tag{3}$$

Here

$$\delta T^2{}_\ell^P \equiv \frac{\ell(\ell+1)}{2\pi} C_\ell^P \tag{4}$$

is the familiar rescaled power that is normally used in power spectrum plots in place of C_ℓ^P. The index P denotes any of the four power spectrum types, *i.e.*, $P = T, X, E, B$. The parameter η is the normalization of the detector noise in the maps relative to the predicted value, and is normally equal to unity. Since $\mathbf{C}$ depends linearly on the parameters p_i, we can define the $\mathbf{P}$-matrices as

$$\mathbf{P}_i \equiv \frac{\partial\mathbf{C}}{\partial p_i}. \tag{5}$$

In other words, the first $4(\ell_{\rm max}-1)$ $\mathbf{P}$-matrices give the contributions from the T, X, E and B power spectra, and the last one is simply the noise covariance matrix, giving the contribution from experimental noise.

Appendix A summarizes how to compute the $\mathbf{C}$-matrix, and is intended for the reader who wishes to write software to do this in practice. The $\mathbf{P}$-matrix corresponding to $\delta T^2{}_\ell^P$ is obtained from these formulas by simply setting all power spectra to zero, with the single exception $\delta T^2{}_\ell^P = 1$, *i.e.*, $C_\ell^P = 2\pi/\ell(\ell+1)$. These $\mathbf{P}$-matrices are therefore independent of the actual power spectra, and depend merely on the relative orientations of the map pixels.

B. Background: from timestream to T, Q & U maps

To place our problem in context, this section briefly reviews how to reduce experimental data to maps in the form of equation (1). Although it is widely known how to do this, we will see that there are some subtle issues related to unmeasured modes.

1. The basic inversion

Suppose we have observed the sky a large number of times with (perhaps) polarized detectors in a variety of different orientations. Let y_i denote the number measured in the i^{th} observation, and group this time-ordered data set (TOD) into an M-dimensional vector $\mathbf{y}$. The observed temperature fluctuation y_i seen through a linear polarizer takes the form

$$y_i = \frac{1}{2}\left[T_{k_i} + Q_{k_i}\cos(2\alpha_i) + U_{k_i}\sin(2\alpha_i)\right] + n_i, \tag{6}$$

where α_i gives the clockwise angle between the polarizer and the reference direction of the coordinate system, n_i denotes the detector noise, and k_i is the number of the pixel pointed to during the i^{th} observation. If the experiment instead measures the difference between two perpendicular polarizations, y_i takes the simpler form

$$y_i = Q_{k_i}\cos(2\alpha_i) + U_{k_i}\sin(2\alpha_i) + n_i, \tag{7}$$

An unpolarized experiment is described by $y_i = T_{k_i} + n_i$. Grouping the numbers into vectors, we can write any of these three expressions as a simple matrix equation

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n}, \tag{8}$$

where the matrix $\mathbf{A}$ encompasses all the relevant details of the observations. For a pure polarization experiment as in equation (7), $\mathbf{A}$ will contain only zeroes except for a single sine and cosine entry in each row, in columns corresponding to the pixel observed. More general observing strategies such as the beam-differencing of the MAP Satellite or modulated beams clearly retain the simple form of equation (8), merely with a slightly more complicated (but known) $\mathbf{A}$-matrix. For a well-designed experiment such as MAP, the system of equations (8) is highly overdetermined, and the estimate $\tilde{\mathbf{x}}$ of the map triplet $\mathbf{x}$ given by the familiar equation [18]

$$\tilde{\mathbf{x}} = \mathbf{W}\mathbf{y}, \quad \mathbf{W} \equiv [\mathbf{A}^t\mathbf{M}\mathbf{A}]^{-1}\mathbf{A}^t\mathbf{M} \tag{9}$$

is unbiased ($\langle\tilde{\mathbf{x}}\rangle = \mathbf{W}\mathbf{A}\mathbf{x} + \mathbf{W}\langle\mathbf{n}\rangle = \mathbf{x}$ since $\mathbf{W}\mathbf{A} = \mathbf{I}$). If $\mathbf{M}$ a reasonable approximation to $\mathbf{N}^{-1}$, then the map noise $\mathbf{W}\mathbf{n}$ will have minimum variance to first order, with covariance matrix $\mathbf{\Sigma} \equiv \mathbf{W}\mathbf{N}\mathbf{W}^t \approx [\mathbf{A}^t\mathbf{N}^{-1}\mathbf{A}]^{-1}$. Both the map triplet $\tilde{\mathbf{x}}$ and its exact noise covariance matrix $\mathbf{\Sigma}$ can be computed in $\sim N^3$ time, and even faster in many important cases [19,22–30]. Note that the last $\mathbf{P}$-matrix is this noise covariance matrix, *i.e.*, $\mathbf{P}_{4\ell_{\max}-3} = \mathbf{\Sigma}$.

2. The problem of missing modes

In many cases, the inversion in equation (9) fails because the matrix to be inverted is singular. Although there are typically much more measurements y_i than unknowns x_i ($M \gg 3N$), symmetries or other properties of the observing strategy often imply that $\mathbf{A}^t\mathbf{x} = \mathbf{0}$ for certain vectors $\mathbf{x}$, *i.e.*, that $\mathbf{A}$ is singular. A ubiquitous example is lack of sensitivity to the mean (monopole) in the map. Experiments measuring linear polarization without cross-linking may be sensitive to only a certain linear combination of Q and U, unable to recover the two separately. As another example, the PIQUE experiment [7] measures only sums of Q-values $90°$ apart on a circle in the sky, thereby losing information about modes taking values $(+1,-1,+1,-1)$ at four corners of a square.

All such problems can in principle be dealt with by regularizing the inversion (see, *e.g.*, the Appendix of [20]), which sets the unmeasurable modes to zero in the final maps, and keeping track of which modes are missing during subsequent analysis of $\tilde{\mathbf{x}}$. In practice, however, it is often more convenient to eliminate this extra bookkeeping requirement by encoding the corresponding information in the noise covariance matrix $\mathbf{\Sigma}$ as near-infinite noise for the missing modes. This ensures that the missing modes are given essentially zero weight in any subsequent analysis. This "deconvolution" technique is described in detail and tested numerically in [31], and in practice corresponds to adding a matrix $\sigma^{-2}\mathbf{I}$ to the $[\mathbf{A}^t\mathbf{M}\mathbf{A}]$-term of equation (9) before the inversion, where σ is about 10^2 times the rms cosmological signal. In summary, it allows *any* observed data set to be put in the standard form of equation (1), described fully by the pair $(\tilde{\mathbf{x}}, \mathbf{\Sigma})$ regardless of any missing modes.

C. Measuring the power spectra with quadratic estimators

In this section, we discuss how to measure the four power spectra C_ℓ^T, C_ℓ^X, C_ℓ^E and C_ℓ^B from the map triplet $\tilde{\mathbf{x}}$ of equation (1). There are two basic approaches to this problem that ultimately give the same answer.

The first approach is to start by deconvolving the Q and U maps into E and B maps, and then use these as inputs to the power spectrum estimation. Since Q and U depend linearly (but non-locally) on E and E, this can always be done with the deconvolution method of [31]. Incomplete sky coverage will simply be reflected by near-infinite noise in certain modes in the resulting noise covariance matrix. An advantage of taking this route is that Wiener-filtered E and B maps can be plotted, whose spatial information may provide useful diagnostics for foreground contamination and systematic errors.

The second approach, which we will adopt here, is to skip the intermediate step of E and B maps and measure the power spectra directly from $\mathbf{x}$.

1. The definition of a quadratic estimator

Our basic problem is to estimate the parameters p_i in equation (2) from the observed data set $\mathbf{x}$ (we drop the tilde from Section II B for simplicity). Fortunately, this problem is mathematically identical to that for the unpolarized case, which has already been solved using so-called quadratic estimators [20,21]. This class of methods is closely related to the maximum-likelihood method — we return to this issue in Section IV. A quadratic estimator q_i is simply a quadratic function of the data vector $\mathbf{x}$, so the most general case can be written as

$$q_i \equiv \mathbf{x}^t\mathbf{Q}_i\mathbf{x} = \mathrm{tr}\,[\mathbf{Q}_i\mathbf{x}\mathbf{x}^t] \tag{10}$$

where $\mathbf{Q}_i$ an arbitrary symmetric $3N \times 3N$-dimensional matrix. We will often find it convenient to group the parameters p_i and the estimators q_i into N_b-dimensional vectors, denoted $\mathbf{p}$ and $\mathbf{q}$, where $N_b = 4\ell_{\max} - 3$ is the number of bands. The matrices $\mathbf{Q}_i$ should not be confused with the vector of stokes parameters $\mathbf{Q}$ from equation (1)!

2. The window function of a quadratic estimator

Since $\mathbf{Q}_i$ can be any symmetric matrix, one can write down infinitely many different quadratic estimators. Whether a given choice is useful or not depends on the mean and covariance of the vector $\mathbf{q}$. Equations (10) and (2) show that the mean of $\mathbf{q}$ is

$$\langle q_i\rangle = \mathrm{tr}\,[\mathbf{Q}_i\mathbf{C}] = \sum_{i'}\mathrm{tr}\,[\mathbf{Q}_i\mathbf{P}_{i'}]p_{i'} \tag{11}$$

$$= b + \sum_{P'=1}^{4}\sum_{\ell'=2}^{\ell_{\max}} W^{\ell P}_{\ell' P'} C^{P'}_{\ell'}, \tag{12}$$

where $i=(\ell_{\max}-1)(P-1)+\ell-1$ is the parameter number corresponding to polarization type P and multipole ℓ, $b\equiv\mathrm{tr}\,[\mathbf{Q}_i\mathbf{\Sigma}]$ is the contribution from experimental noise, and

$$W^{\ell P}_{\ell' P'} \equiv \mathrm{tr}\,[\mathbf{Q}_i\mathbf{P}_{i'}]. \tag{13}$$

These quantities can be viewed as a generalized form of window functions, since for a fixed (P,ℓ), they show the expected contributions to q_i not only from different ℓ-values, but also from different polarization types.

Ideally, we would be able to estimate C^P_ℓ by applying a quadratic estimator with the perfect window function $W^{\ell P}_{\ell' P'}=\delta_{PP'}\delta_{\ell\ell'}$, but this is often impossible or undesirable with incomplete sky coverage, shifting the aim to making the window functions narrow in both the ℓ-direction and the P-direction. Minimizing such unwanted mixing of different polarization types is one of the key topics of this paper, and numerous examples of such window functions will be plotted in Section III.

The covariance matrix of $\mathbf{q}$ is

$$\mathbf{M}_{ij} \equiv \langle\mathbf{q}\mathbf{q}^t\rangle - \langle\mathbf{q}\rangle\langle\mathbf{q}\rangle^t = 2\,\mathrm{tr}\,[\mathbf{Q}_i\mathbf{C}\mathbf{Q}_j\mathbf{C}] \tag{14}$$

for the case where $\mathbf{x}$ is Gaussian, and it is clearly desirable to make it small in some sense.

3. Quadratic estimators: specific examples

It can be shown [20] that the quadratic estimator defined by

$$\mathbf{Q}_i = \frac{1}{2}\mathcal{N}_i\sum_j(\mathbf{B})_{ij}\mathbf{C}^{-1}\mathbf{P}_j\mathbf{C}^{-1}, \tag{15}$$

distills all the cosmological information from $\mathbf{x}$ into the (normally much shorter) vector $\mathbf{q}$ if $\mathbf{C}$ is the true covariance matrix. Moreover, if $\mathbf{C}$ is a reasonable estimate of the true covariance matrix, say by computing it as in Appendix A using a prior power spectrum consistent with the actual measurements, then the data compression step of going from $\mathbf{x}$ to $\mathbf{q}$ destroys information only to second order. In equation (15), $\mathbf{B}$ is an arbitrary invertible matrix, and the normalization constants $\mathcal{N}_i$ are chosen so that all window functions sum to unity:

$$\sum_{i'=1}^{N_b-1}\mathrm{tr}\,[\mathbf{Q}_i\mathbf{P}_{i'}] = 1. \tag{16}$$

This means that we can interpret q_i as measuring a weighted average of our unknown parameters, the window giving the weights. We will discuss a number of choices of $\mathbf{B}$ in Section III that have various desirable properties. $\mathbf{B}=\mathbf{I}$ gives minimal but correlated error bars. $\mathbf{B}=\mathbf{F}^{-1}$ gives beautiful Kronecker-delta window functions, corresponding to $\langle\mathbf{q}\rangle=\mathbf{p}$ at the price of anti-correlated and typically very large error bars, where

$$\mathbf{F}_{ij} = \frac{1}{2}\mathrm{tr}\,\left[\mathbf{C}^{-1}\frac{\partial\mathbf{C}}{\partial p_i}\mathbf{C}^{-1}\frac{\partial\mathbf{C}}{\partial p_j}\right] \tag{17}$$

is the so-called Fisher information matrix [14,15] for the case where the CMB fluctuations are Gaussian. The intermediate choice $\mathbf{B}=\mathbf{F}^{-1/2}$ is normally a useful compromise [32], giving uncorrelated error bars and narrow window function with width $\Delta\ell$ of order the inverse map size. We will describe and test additional choices of $\mathbf{B}$ in Section III C and Section III D.

4. Broadening the bands

For CMB maps of small size where the window function width $\Delta\ell\gg 1$, it is unnecessary to oversample the measured power with a separate parameter p_i at each multipole ℓ. In such cases, it is useful to parametrize the power spectrum as a staircase-shaped (piecewise constant) function, with the parameters p_i giving the heights of the various steps [21]. We will occasionally do this in Section III.

TOY MODEL SPECIFICATIONS

Experiment	Coverage	FWHM	Noise/pixel
COBE	$b>20$	7°	$30\mu K$
MAP	$b>20$	$13'$	$71\mu K$
Planck	$b>20$	$8'$	$8\mu K$
B2001	$b>80$	$12'$	$16\mu K$
Circle	$b=80$	$12'$	$16\mu K$

III. RESULTS

In this section, we will apply the quadratic estimator method to a variety of fictitious data sets to quantify how experimental attributes such as sky coverage and sensitivity affect the ability to measure and separate the different power spectra. We will also describe two ways in which the basic quadratic estimator technique can in some circumstances be improved for polarization applications.

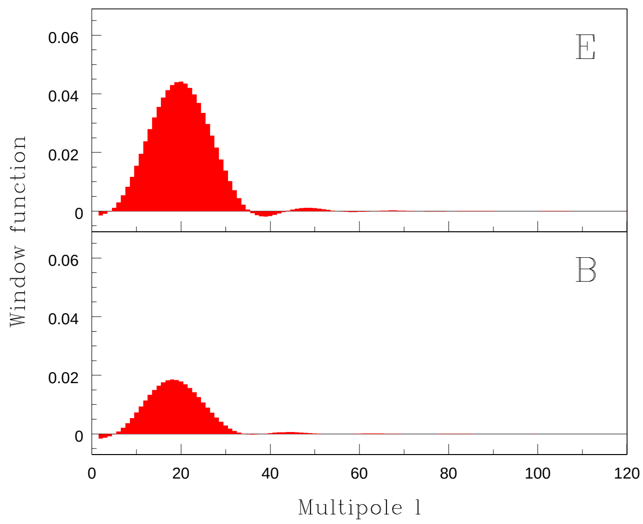

FIG. 1. The window function corresponding to the B2001 measurement of $\delta T^2{}_\ell^E$ for $\ell = 20$. Upper panel shows sensitivity to E-power (wanted) and lower panel shows sensitivity to B-power (unwanted — what we call "leakage").

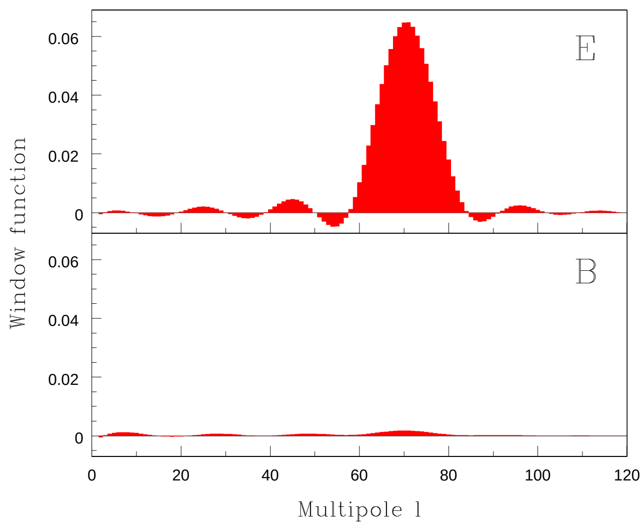

FIG. 2. Same as Figure 1, but for the B2001 E-measurement aimed at $\ell = 70$.

A. Case studies

Our five case studies are listed in Table 1. The first three cover the northern Galactic cap $b > 20°$ with successively higher sensitivity. The fourth covers a much smaller cap $b > 80°$, and the fifth covers merely a one-dimensional region: the circle defined by $b = 80°$. These case studies are not intended to be accurate forecasts for the actual performance of the experiments listed, but rather to span an interesting range in sensitivity, sky coverage and map shape. There are therefore numerous departures from realism. For instance, the actual maps from COBE, MAP and Planck will of course include the southern Galactic caps as well — apart from reducing the error bars, adding this reflection symmetry to the sky maps eliminates all leakage between even and odd ℓ-values [20], preserving the overall width of the window functions that we will present but giving them a jagged behavior where every other entry vanishes. The actual map from the Boomerang 2001 ("B2001") experiment will not be round and will have non-uniform sensitivity. We assume uncorrelated pixel noise for simplicity. Most of the experimental sensitivities we have used are likely to be slight underestimates, being based on a single frequency channel.

In all cases, we explicitly perform the various matrix computations described in Section II. The reason that this is numerically feasible within the scope of this paper is that the large angular scales of interest here allow us to use larger and fewer pixels than the experimental teams will employ in their actual data reduction.

We will first study window functions to quantify how accurately E and B can be separated in various cases. We will then discuss measurement of the cross power spectrum and finally investigate how accurately approximate error bars from the Fisher matrix formalism match the results from our full numerical calculation.

B. E and B window functions

We begin by quantifying the ability of $B2001$ and MAP to separate E and B using Q- and U-maps. We pixelize our sky patches using the equal-area icosahedron method [33] at resolution levels 35 and 7, respectively, corresponding to 361 B2001 pixels and 561 MAP pixels*. Since Q and U are measured for each pixel, the data vectors $\mathbf{x}$ have twice these lengths. We use the method

* We use the icosahedron pixelization since it has the roundest (mainly hexagonal) pixels and is highly uniform. Although we did not use it here, the HEALPIX package [34] offers a useful alternative, allowing azimuthal symmetry to be exploited for saving computer time.

given by equation (15) with $\mathbf{B} = \mathbf{F}^{-1/2}$ unless otherwise specified. We compute fiducial power spectra C_ℓ^T, C_ℓ^E and C_ℓ^X, with the CMBFAST software [35] using cosmological parameters from the "concordance" model from [36], which provides a good fit to existing CMB and large scale structure data. Although the true B-power spectrum may be close to zero, we set $C_\ell^B = C_\ell^E$ in our fiducial model since we wish to highlight geometrical effects. Since this prior is E/B-symmetric, any asymmetries between E and B in our resulting window functions and error bars will be due to geometry alone. We eliminate sensitivity to offsets by projecting out the mean (monopole) for the T, Q and U maps separately, as described in the Appendix of [20].

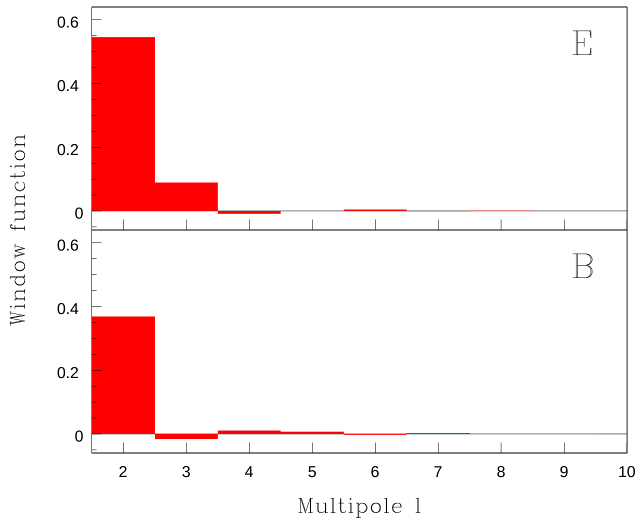

FIG. 3. Same as Figure 1, but for the MAP E-measurement aimed at $\ell = 2$.

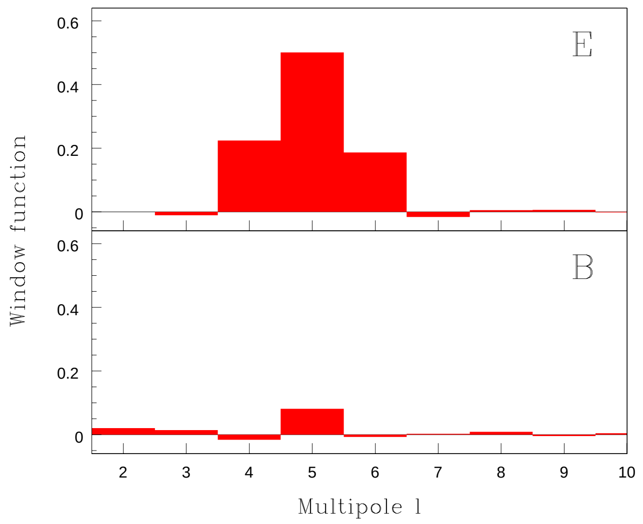

FIG. 4. Same as Figure 1, but for the MAP E-measurement aimed at $\ell = 5$.

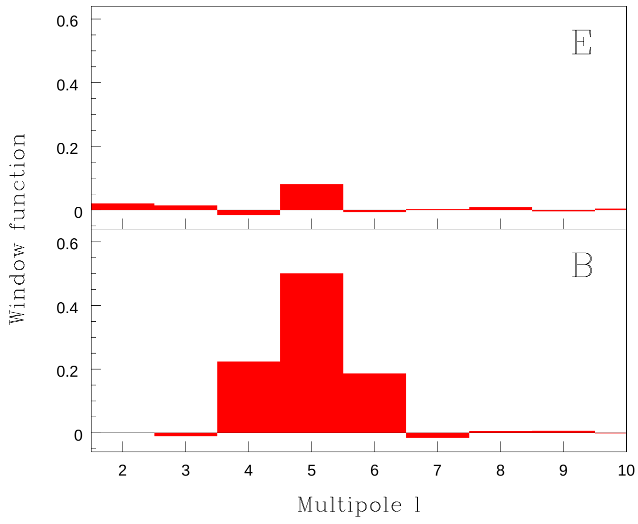

FIG. 5. Same as Figure 1, but for the MAP measurement of $\delta T^2{}_\ell^B$ for $\ell = 5$.

1. Dependence on sky coverage and angular scale

Figures 1-5 show a sequence of sample window functions for B2001 and MAP. Note that it is possible for window functions to go slightly negative for the decorrelation method $\mathbf{B} = \mathbf{F}^{-1/2}$ used here, whereas the method given by $\mathbf{B} = \mathbf{I}$ guarantees non-negative windows. Just as for the unpolarized case, the window function width $\Delta\ell$ is seen to be fairly independent of the target multipole ℓ, essentially scaling as the inverse size of the sky patch covered [37].

The amount of leakage of B-power into our E estimate is quantified by the lower panels in Figures 1-4, and is seen to decrease as smaller scales are probed. Figure 5 targets B-polarization and looks like Figure 4 with the two panels swapped, thereby showing that the leakage problems between E and B are quite symmetric. The smaller the area under the unwanted half of the window function, the better our method separates E and B. As a simple quantitative measure of this power leakage, let us therefore define a 2×2 *leakage matrix* $\mathbf{L}^\ell$ for each ℓ, given by

$$\mathbf{L}_{PP'}^\ell \equiv \sum_{\ell'=2}^{\ell_{\max}} W_{\ell' P'}^{\ell P}, \qquad (18)$$

where P and P' take the values E, B. In other words, the four components of this leakage matrix are the areas under the four histograms in Figure 4 and 5. If $\mathbf{L}_{EE}^\ell = \mathbf{L}_{BB}^\ell = 1$ and $\mathbf{L}_{EB}^\ell = \mathbf{L}_{BE}^\ell = 0$, *i.e.*, if $\mathbf{L}^\ell = \mathbf{I}$, then there is on average no leakage at all between E and B. For the simple case of complete sky coverage and uniform noise, all window functions become Kronecker delta functions, $W^{\ell P}_{\ell' P'} = \delta_{PP'}\delta_{\ell\ell'}$, and we verified that this happens numerically as a test of our software (in practice, it works only when ℓ, ℓ' are smaller than the scale corresponding to the pixel separation, *i.e.*, when the map is adequately oversampled). This simple case thus gives the ideal case $\mathbf{L}^\ell = \mathbf{I}$, but we will return in Section III C to a method producing this desirable result even for partial sky coverage.

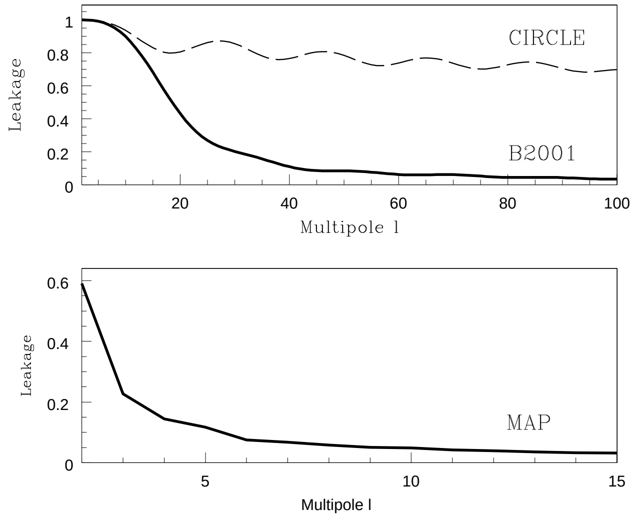

FIG. 6. The amount of leakage of B-power into the E-power spectrum estimates is shown for the cases of B2001 (top panel, solid), CIRCLE (top panel, dashed) and MAP (bottom panel, solid). These curves show the B/E ratio ($\mathbf{L}^\ell_{EB}/\mathbf{L}^\ell_{EE}$) for the E-estimates. The corresponding curves for leakage in the reverse direction (the E/B-ratio $\mathbf{L}^\ell_{BE}/\mathbf{L}^\ell_{BB}$ for the B-estimates) are visually identical.

To assess how the leakage depends on sky coverage and angular scale, we plot the leakage for B2001, MAP and CIRCLE as a function of ℓ in Figure 6. Specifically, we plot the ratios of unwanted to wanted contributions, *i.e.*, $\mathbf{L}^\ell_{EB}/\mathbf{L}^\ell_{EE}$ and $\mathbf{L}^\ell_{BE}/\mathbf{L}^\ell_{BB}$. These plots show three noteworthy results:

1. The situation for E and B is rather symmetric, with essentially equal leakage from B to E as vice versa.
2. The leakage drops with ℓ.
3. The B2001 and MAP curves have roughly similar shape apart from a scaling of the horizontal axis by a factor ~ 7, corresponding to the map size ratio.

Result 2 is expected since map boundary effects (incomplete sky coverage) are the reason that we cannot separate E and B perfectly — these boundary effects become less important as angular scales much smaller than the map are considered. In the small-scale limit where sky curvature and discreteness of ℓ become irrelevant, one would expect result 3 as well, since there is no other ℓ-scale in the problem than the window function width $\Delta\ell$, of order the inverse size of the map. If θ denotes the diameter of our circular sky patches in radians, then the FWHM window widths for B2001 and MAP are roughly fit by $\Delta\ell \approx 5/\theta$, and the figures show that the leakage ratio drops below 15% for $\ell \gtrsim 2\Delta\ell$. (Things are different for the CIRCLE case, which we defer to Section III B 2.)

In conclusion, we have found that E/B separation works well for $\ell \gg \Delta\ell$.

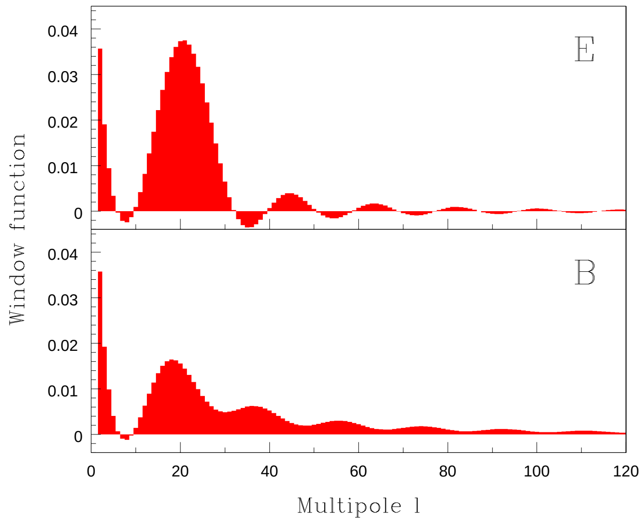

FIG. 7. Same as Figure 1, but for the CIRCLE E-measurement aimed at $\ell = 20$.

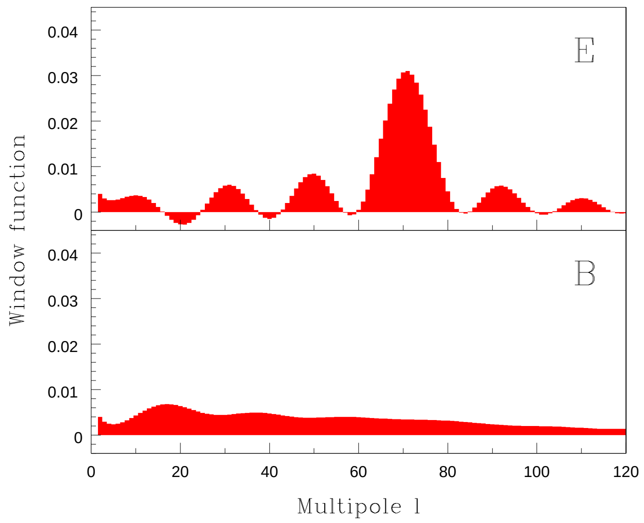

FIG. 8. Same as Figure 1, but for the CIRCLE E-measurement aimed at $\ell = 70$.

2. Dependence on map shape

Above we studied how leakage problems depend on map size. To assess how they depend on map shape, we will now compare two rather extreme examples: a disc and a circle. The B2001 and CIRCLE maps have the same diameter and thus probe comparable angular scales. However, whereas the B2001 map is truly two-dimensional, the CIRCLE map is essentially one-dimensional, containing merely a single strip of pixels along the circumference. The two current polarization experiments POLAR [10] and PIQUE [7] both use ring-shaped maps, and this important case has also been extensively studied theoretically [11].

We pixelize our CIRCLE map with 360 equispaced pixels around the circle. Sample CIRCLE window functions are shown in Figures 7 and 8. It is seen that the CIRCLE windows are generally broader (with larger $\Delta\ell$) than their B2001 counterpart, which agrees with the well-known rule of thumb [23] that the narrowest dimension of a map is the limiting factor. We also see that whereas the B2001 leakage reduced on smaller scales, things do not get correspondingly better for the CIRCLE case. This explains the dashed curve in Figure 6, which shows that substantial leakage persists even on small scales. In Section III C, we will describe how leakage can be further reduced.

In conclusion, we find that although ring maps do allow an interesting degree of E/B-separation, a two-dimensional map works better in this regard.

3. Dependence on sensitivity

Above we investigated how E/B-leakage was affected by map size and shape. To assess the effect of map sensitivity, we compare our COBE, MAP and Planck examples. These have identical sky coverage and pixelization, so the only difference is the sensitivity per unit area which increases dramatically from COBE to MAP to Planck. We refrain from plotting the three leakage curves, since they look visually identical to the MAP curve in Figure 6. This means that the effect of sensitivity on E/B separation is negligible compared to the effect of sky coverage. In other words, it depends mainly on geometry and only weakly on the (sensitivity-dependent) details of the pixel weighting.

Generally, the quadratic estimator method strives to minimize error bars by reducing leakage from multipoles and polarization types with substantial power. In a situation where sample variance is dominant, this tends to make windows slightly lopsided, with a wider wing towards the direction where power decreases — in most cases towards the right. Conversely, in a situation where detector noise is dominant, windows tend to be slightly lopsided in the opposite sense, since noise power normally increases on smaller scales.

C. Disentangling E and B better

Above we have seen that the E and B power spectra can be fairly accurately separated on angular scales $\ell \gg \Delta\ell$ with the decorrelated quadratic estimator method. Here we will argue that it is in some cases possible to do even better. Our motivation for pursuing this is that whereas broad windows in the ℓ-direction are easy to interpret (corresponding simply to a smoothing of the power spectrum), the mixing of different polarization types is rather annoying and complicates the interpretation of results.

Each choice of $\mathbf{B}$ in equation (15) corresponds to a different way of plotting the results. Which is the best choice?

If the goal is to use the measured band power estimates in $\mathbf{p}$ to constrain cosmological parameters, the choice is irrelevant. Any two methods using invertible $\mathbf{B}$-matrices of course retain exactly the same cosmological information, since it is possible to go back and forth between the corresponding two $\mathbf{p}$-vectors by multiplying by $\mathbf{B}_2\mathbf{B}_1^{-1}$ or $\mathbf{B}_1\mathbf{B}_2^{-1}$. This means that the likelihood function for cosmological parameters will be identical for the two methods.

The choice of $\mathbf{B}$ is therefore mainly a matter of presenting the power spectrum measurements in a clear and intuitive way. Ideally, a method would have the following properties

1. No leakage between E and B
2. Narrow window functions
3. Uncorrelated error bars

Unfortunately, these three properties are only achievable simultaneously (without information loss) for the case of complete sky coverage. As discussed in [32], the choice $\mathbf{B} = \mathbf{F}^{-1/2}$ in equation (15) achieves 3 and does a fairly good job on 2, giving window functions with the fundamental ℓ-resolution corresponding to the sky coverage. Above we saw that this gives a leakage between E and B that is substantial on scales corresponding to the size of the survey and approaches zero on substantially smaller scales.

In contrast, the choice $\mathbf{B} = \mathbf{F}^{-1}$ in equation (15) achieves 1 and 2 perfectly, but at the cost of producing measurements that are difficult to interpret because the error bars are typically anticorrelated and huge [32]. However, these problems are to a large extent caused by eliminating the rather benign leakage between different ℓ-values. We will now describe a less aggressive method that merely targets the E/B-leakage. Specifically, let us insist that the leakage matrices $\mathbf{L}^\ell = \mathbf{I}$ for all ℓ.

Let us define the *disentangling method* as the one given by equation (15) with the $\mathbf{B}$-matrix chosen as

$$\mathbf{B}_{ii'} = \mathbf{B}_{\mathbf{P}\ell\mathbf{P}'\ell'} \equiv \sum_{P''} (\mathbf{L}_\ell^{-1})_{PP''} \mathbf{F}^{-1/2}_{\mathbf{P}''\ell\mathbf{P}'\ell'}. \quad (19)$$

Here we have once again combined P and ℓ into a single index $i = (\ell_{\max} - 1)(P - 1) + \ell - 1$. It is easy to show that this method gives ideal leakage matrices $\mathbf{L}^\ell = \mathbf{I}$ if the leakage matrices in equation (19) were computed using $\mathbf{B} = \mathbf{F}^{-1/2}$. This means that the unwanted half of the window function averages to zero. A similar scheme was found to work well for disentangling three types of power in a galaxy redshift survey [38].

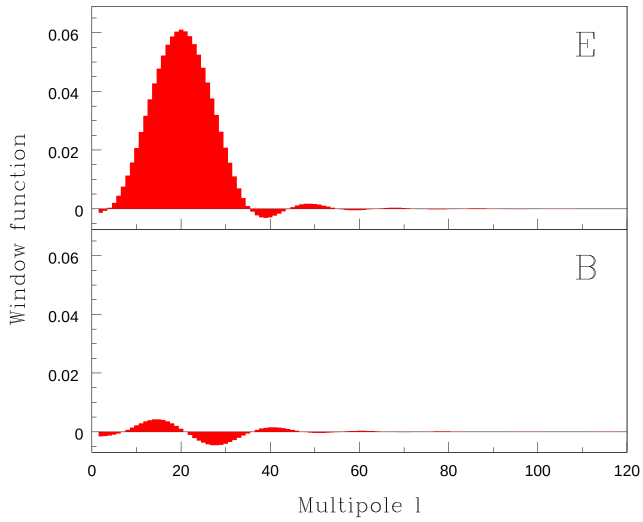

FIG. 9. Same as Figure 1, but after applying our disentanglement method.

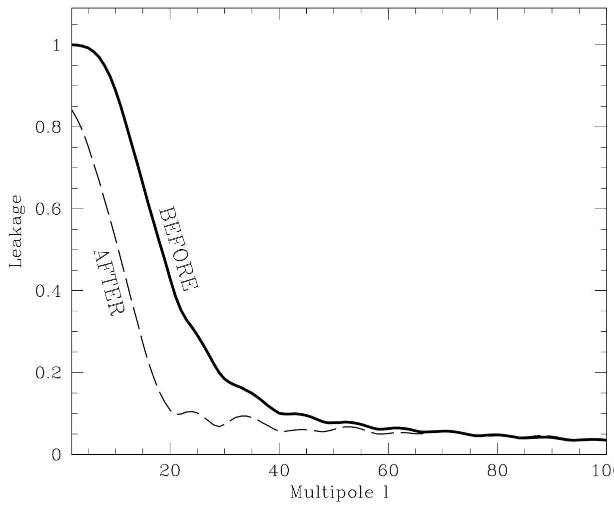

FIG. 10. Curves showing leakage of B-power into estimates of E are plotted for B2001 before (solid) and after (dashed) applying out disentanglement method. In this plot, leakage has been computed by taking absolute values of the window functions — otherwise the dashed curve would be identically zero. For example, the value of the dashed curve at $\ell = 20$ can be interpreted simply as the ratio of shaded area in the two panels of Figure 9.

We test this method for our B2001 example, and a sample disentangled window function is shown in Figure 9. Comparing this with Figure 1, which targeted the same multiple, we see that the leakage (lower panel) has been substantially reduced and oscillates around zero. On very large angular scales $\ell \sim \Delta\ell$, this undesirable half of the window function remains substantial even though it by construction averages to zero. On small scales, however, the unwanted part of the window function is found to be consistently near zero, not merely on average. This is because both the desirable and the undesirable halves of the initial window function before disentanglement have essentially the same shape, so that our disentanglement process will cancel them out almost completely. To quantify how well the this process works, Figure 10 shows leakage curves with the leakage matrix redefined with absolute values:

$$\tilde{\mathbf{L}}^\ell_{PP'} \equiv \sum_{\ell'=2}^{\ell_{\max}} \left| W^{\ell P}_{\ell' P'} \right|, \tag{20}$$

(Without taking absolute values, the disentanglement method would by construction be diagnosed with zero leakage.)

Although this method is likely to be adequate for practical applications, it is possible to disentangle E and B still better if necessary, at the price of larger error bars. The Fisher matrix $\mathbf{F}$ generally becomes singular if the bands used are much narrower than $\Delta\ell$. If the sky coverage is large enough that $\Delta\ell$ is narrower than features of interest in the power spectra, is it convenient to choose bands of width around $\Delta\ell$ instead of width one (solving for each multipole separately). Since this produces an invertible Fisher matrix, perfect disentanglement is achievable by setting $\mathbf{B} = \mathbf{F}^{-1}$ in equation (15), giving only a modest increase in error bars and rather slight anticorrelations between neighboring bands. The resulting measurements can then be made uncorrelated separately for E and B, by multiplying by the inverse square roots of the corresponding two covariance matrices, thereby broadening the ℓ-windows back to their natural widths.

D. Measuring the cross power spectrum

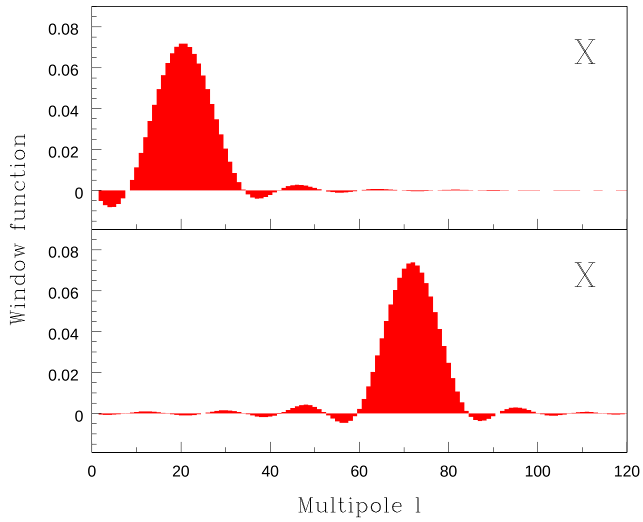

FIG. 11. Window functions for measuring the cross power spectrum are shown for the B2001 case, targeting $\ell = 20$ (top) and $\ell = 70$ (bottom).

The cross power spectrum C_ℓ^X can be measured using the basic decorrelated quadratic estimators given by equation (15) without any modification. However, as we will now discuss, this is not necessarily the most desirable approach.

To measure the cross power spectrum, our data vector $\mathbf{x}$ must contain both unpolarized and polarized measurements as in equation (1). These may be either from a single experiment or from separate unpolarized and polarized ones. As reviewed in Appendix A, this gives a covariance matrix of the form

$$\mathbf{C} \equiv \langle \mathbf{x}\mathbf{x}^t \rangle = \begin{pmatrix} \langle \mathbf{T}\mathbf{T}^t \rangle & \langle \mathbf{T}\mathbf{Q}^t \rangle & \langle \mathbf{T}\mathbf{U}^t \rangle \\ \langle \mathbf{Q}\mathbf{T}^t \rangle & \langle \mathbf{Q}\mathbf{Q}^t \rangle & \langle \mathbf{Q}\mathbf{U}^t \rangle \\ \langle \mathbf{U}\mathbf{T}^t \rangle & \langle \mathbf{U}\mathbf{Q}^t \rangle & \langle \mathbf{U}\mathbf{U}^t \rangle \end{pmatrix} \tag{21}$$

which generically has non-vanishing elements in all the off-diagonal blocks. When we try to measure one of the parameters p_i corresponding to the cross power spectrum, the matrix $\mathbf{P}_i = \partial\mathbf{C}/\partial p_i$ will vanish except in the $T-Q$ and $T-U$ cross-correlation blocks, since these are the only ones that depend on C_ℓ^X. However, since $\mathbf{C}$ is a full matrix, the quadratic estimator $\mathbf{Q}_i \propto \mathbf{C}^{-1}\mathbf{P}_i\mathbf{C}^{-1}$ (as well as decorrelated or disentangled variants thereof) will also be a full matrix, without any vanishing blocks. This means that the estimates q_i of the cross power spectrum will involve not only data combinations like T_jQ_k and T_jU_k, but also terms like T_jT_k and Q_jQ_k. In other words, the measured cross-correlation involves, among other things, the correlation of the temperature map with itself! The same peculiarity applies to the maximum-likelihood method, which is simply an iterated version of the quadratic method.

We will now give a simple example illustrating why this happens, as well as argue that it is avoidable and sometimes undesirable.

1. A complicated way to measure correlation

As an illustrative toy model, let us temporarily assume that our data vector $\mathbf{x}$ consists of merely two numbers, the measurements of T and E for a given multipole (ℓ, m) extracted from an all-sky map. Consider the simple case where $C_\ell^T = C_\ell^E = 1$ — the general case will follow from our result by a trivial scaling. This means that the data covariance matrix takes the form

$$\mathbf{C} \equiv \langle \mathbf{x}\mathbf{x}^t \rangle = \begin{pmatrix} \langle C_\ell^T \rangle & \langle C_\ell^X \rangle \\ \langle C_\ell^X \rangle & \langle C_\ell^E \rangle \end{pmatrix} = \begin{pmatrix} 1 & r \\ r & 1 \end{pmatrix}, \tag{22}$$

where $r \equiv C_\ell^X/[C_\ell^T C_\ell^E]^{1/2}$ is the dimensionless correlation coefficient between T and E. We have only three parameters to measure in our toy example, $\mathbf{p} = (C_\ell^T, C_\ell^E, C_\ell^X)$, so the matrices $\mathbf{P}_i = \partial\mathbf{C}/\partial p_i$ are

$$\mathbf{P}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \mathbf{P}_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathbf{P}_3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{23}$$

Substituting this and equation (22) into equation (17) gives the Fisher matrix

$$\mathbf{F} = \frac{1}{(1-r^2)^2} \begin{pmatrix} \frac{1}{2} & \frac{r^2}{2} & -r \\ \frac{r^2}{2} & \frac{1}{2} & -r \\ -r & -r & 1+r^2 \end{pmatrix}, \tag{24}$$

with inverse

$$\mathbf{F}^{-1} = \frac{1}{2} \begin{pmatrix} 1 & r^2 & r \\ r^2 & 1 & r \\ r & r & \frac{1+r^2}{2} \end{pmatrix}. \tag{25}$$

Substituting the above equations into equation (15) and using the method given by $\mathbf{B} = \mathbf{F}^{-1}$, the resulting unbiased estimators take the simple form

$$\mathbf{Q}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \mathbf{Q}_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathbf{Q}_3 = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{26}$$

This is not surprising, since no other estimators could possibly be unbiased ($\langle \mathbf{q} \rangle = \mathbf{p}$) when only one spherical harmonic mode is used. As soon as more modes are available, however, things can (and generally do) get more complicated. Consider, the case where $\mathbf{x}$ consists of four rather than two measurements for our multipole, — temperature measurements T_1 and T_2 and E-polarization measurements E_1 and E_2. The obvious guess would be that and estimator of C_ℓ^X should involve only cross terms (T_1E_1, T_1E_2 and T_2E_2). However, terms such as $T_1^2 - T_2^2$

and $E_1^2 - E_2^2$ will have a vanishing average, and can therefore be added to any estimator without biasing it. As we saw above, precisely this generically happens in our minimum-variance estimator, since it helps reduce the estimator variance if our four measurements have different noise variances.

This can also be understood as follows. Repeating the derivation of equation (26) when there are n measurement of T and E for our mode, perhaps with different variances, shows that the zeroes in equation (26) will be replaced by $n \times n$ blocks that generally not vanish — merely their traces will.

2. A simpler way

Apart from being surprisingly complicated to compute and interpret, the basic quadratic estimators of equation (10) also have drawback related to systematic errors. If the estimated cross-power spectrum involves components of the unpolarized and polarized power spectra that are supposed to cancel out to reduce the variance, there is a risk that systematic errors from these two autocorrelations propagate into and contaminate the cross-correlation measurements. For this reason, an interesting alternative is to use an estimator for the cross-power spectrum that does not have this funny property, *i.e.*, that contains only cross-terms. A simple way to construct such an estimator is to use equation (15) with the fiducial power spectrum C_ℓ^X set equal to zero. This will make $\mathbf{C}$ and $\mathbf{C}^{-1}$ block diagonal, so the $\mathbf{Q}$-matrices of equation (15) acquire the same block structure as the $\mathbf{P}$-matrices. The estimators of C_ℓ^X therefore involve only cross terms.

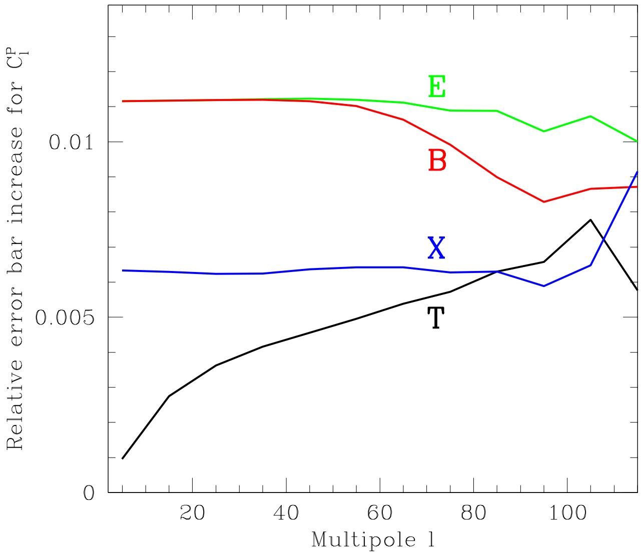

FIG. 12. The low cost of simplicity. The curves show the relative increase of the error bars on the four power spectra when the fiducial model is replaced by one with $\mathbf{C}_\ell^X = 0$, thereby eliminating potential systematic errors as described in the text.

It is noteworthy that this issue applies not only to estimation of C_ℓ^X, but to measuring C_ℓ^T, C_ℓ^E and C_ℓ^B as well. If the fiducial power spectrum has $C_\ell^X \neq 0$, then estimators of all four power spectra will involve using combinations of all three maps $(\mathbf{T}, \mathbf{Q}, \mathbf{U})$, so setting $C_\ell^X \neq 0$ when using equation (15) is an interesting simplifying option for all power spectrum estimation. We implicitly did so in Section III B by ignoring the T-map.

3. Which method is better?

The price we must pay for this simplification is a slight increase in error bars. This is quantified in Figure 12 for the B2001 case. We computed the error bars on the four power power spectra for the unbiased method with both weighting schemes (assuming the true C_ℓ^X for our concordance model and assuming $C_\ell^X = 0$)†. The figure shows the ratios (minus one), and illustrates that the simplification typically comes at a very low cost — an error bar increase of order a percent. In light of this and the potential peril of systematic errors, the simpler method appears preferable in practice.

E. Error bars

Up until now, we have focused on the issue of window functions. Let us now turn to the complementary issue of error bars. A large number of papers, for instance [2,5,6,16,39–43], have made forecasts for how accurately upcoming polarization experiments can constrain cosmological parameters. These estimates all assumed that the accuracy of the recovered polarized power spectra $(C_\ell^T, C_\ell^X, C_\ell^E, C_\ell^B)$ would be given by the 4×4 covariance matrix

$$\mathbf{M}_\ell \approx \frac{2 f_{\rm sky}^{-1}}{2\ell+1} \begin{pmatrix} T_\ell^2 & X_\ell^2 & 0 & T_\ell X_\ell \\ X_\ell^2 & E_\ell^2 & 0 & E_\ell X_\ell \\ 0 & 0 & B_\ell^2 & 0 \\ T_\ell X_\ell & E_\ell X_\ell & 0 & \frac{1}{2}[T_\ell E_\ell + X_\ell^2], \end{pmatrix}. \quad (27)$$

where $f_{\rm sky}$ is the fraction of the sky observed and

$$P_\ell \equiv C_\ell^P + (w^P)^{-1} e^{\theta^2 \ell(\ell+1)}, \quad P = T, X, E, B, \quad (28)$$

if the experimental beam is Gaussian with width θ in radians (the full-width-half-maximum is given by FWHM$=\sqrt{8 \ln 2}\,\theta$). Here the sensitivity measure $1/w^P$ is defined as [44] the noise variance per pixel times the pixel

† Specifically, we use broad ℓ-bins as described in Section II C 4 with $\Delta\ell = 10$ and use the method with $\mathbf{B} = \mathbf{F}^{-1}$ in equation (15) to ensure that we are comparing apples with apples, *i.e.*, to ensure that we are comparing error bars for measurements with identical window functions.

area in steradians for $P = T, E, B$. For the case of the cross power spectrum, $w^X = (w^T w^E)^{1/2}$. Equation (28) can be interpreted as simply giving the total power from CMB (first term) and detector noise (second term). For our examples, we take $w^T = w^E = w^B = 1/(\mathrm{FWHM}\,\sigma)^2$, where FWHM and σ are the resolution and noise values listed in Table 1.

The approximation of equation (27) has been shown to be exact for the special case of complete sky coverage ($f_{\rm sky} = 1$) [2], and the same result follows from the Fisher information matrix formalism [6]. The factor $(2\ell+1)f_{\rm sky}$ can be interpreted as the effective number of uncorrelated modes per multipole[‡], and the other factor as giving the covariance per mode.

The approximation that the number of uncorrelated modes scales as $f_{\rm sky}$ is both natural and well-motivated [45]. How accurate is it in practice? Our calculations enable us to address this issue quantitatively. Figure 13 shows the ratio of the approximate error bars from equation (27) to the exact error bars from equation (14) for the four power spectra (the ratios of the square roots of the corresponding elements on the covariance matrix diagonals). To ensure a fair comparison, we used the uncorrelated method given by equation (15) with $\mathbf{B} = \mathbf{F}^{-1/2}$ — the $\mathbf{B} = \mathbf{F}^{-1}$ and maximum-likelihood methods give larger anticorrelated error bars, whereas the $\mathbf{B} = \mathbf{I}$ method gives smaller correlated ones.[§] It is seen that the approximation of equation (27) is generally quite accurate when $\ell \gg \Delta\ell$ and $\ell \ll \ell_{\max} - \Delta\ell$, *i.e.*, for multipoles well away from the two endpoints of the range computed. This means that forecasts made using the approximation of equation (27) should be viewed as quite accurate except on scales comparable to the survey size.

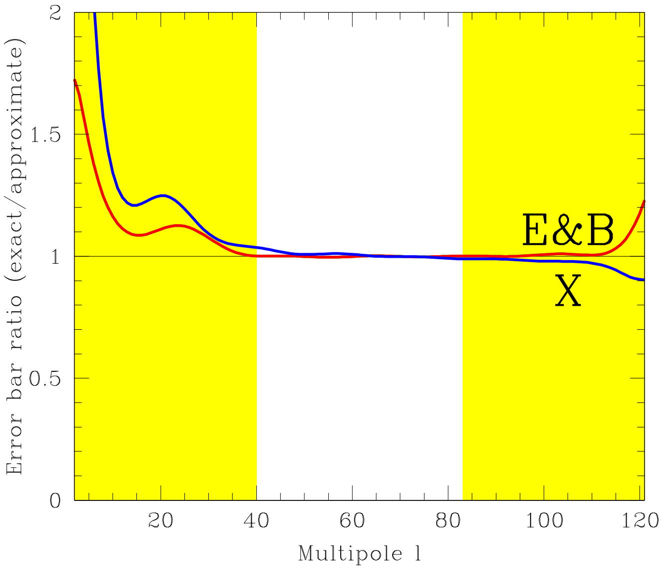

FIG. 13. The curves show the ratio of the actual error bars to those computed with the approximation of equation (27). The approximation is seen to be excellent for multipoles more than a couple of bandwidths $\Delta\ell$ away from the two endpoints of the ℓ-range computed (outside shaded regions).

IV. DISCUSSION

We have presented a method for measuring CMB polarization power spectra given incomplete sky coverage and tested it with a number of simulated examples.

A. What have we learned?

The issue of measuring the **T power spectrum** has been extensively treated in prior literature. An added challenge when measuring the **E and B power spectra** is leakage between the two caused by incomplete sky coverage. We quantified this leakage for the first time, and found that it is rather insensitive to experimental noise levels and depends mainly on geometry. Specifically, we found the leakage to depend mainly on the ratio $\ell/\Delta\ell$, where $\Delta\ell$ is the characteristic window function width and scales roughly as the inverse size of the sky patch in its narrowest direction. We introduced a disentanglement method which reduced the leakage to $5\% - 10\%$ for

‡ When the sky coverage $f_{\rm sky} < 1$, certain multipoles become correlated [37]. This reduces the effective number of uncorrelated modes by a factor $f_{\rm sky}^{-1}$, thereby increasing the sample variance on power measurements by the same factor [45,44]. It also smears out sharp features in the power spectrum by an amount $\Delta\ell$ comparable to the inverse size of the sky patch in its narrowest direction [23] and mixes E and B power as discussed in the previous sections.

§ The reader interested in implementing any of these methods in practice should note that care needs to be taken when the bands used are much narrower than $\Delta\ell$, since this makes $\mathbf{F}$ for all practical purposes singular, with many eigenvalues so close to zero that rounding errors make them slightly negative. Our B2001 example with bandwidth 1 (we are measuring for each ℓ separately and $\Delta\ell \sim 30$) is a case in point. For the $\mathbf{B} = \mathbf{F}^{-1/2}$ case, the window functions are simply proportional to the rows of $\mathbf{F}^{1/2}$, so they are readily computed by setting the offending eigenvalues to zero. This determines the normalization constants as $\mathcal{N}_i = 1/\sum_{j=1}^{N_b}(\mathbf{F}^{1/2})_{ij}$, since each window function must sum to unity. The error bar $\Delta\mathbf{q}_i$ is then equal to $\mathcal{N}_i$. In short, all plots in this paper remain well-defined even when $\mathbf{F}$ is singular. Evaluating $\mathbf{q}$ in practice, however, is not possible when $\mathbf{F}$ is singular, since it involves calculating $\mathbf{F}^{-1/2}$ numerically. If $\mathbf{B}$ is chosen to be a regularized version of $\mathbf{F}^{-1/2}$ in equation (15), the decorrelation method fails in the sense that the band power estimates q_i will exhibit a slight residual correlation. In conclusion, the power spectrum should not be too oversampled for actual data analysis. The same obviously applies to the $\mathbf{B} = \mathbf{F}^{-1}$ method.

$\ell \gtrsim \Delta\ell$, and described how it could be pushed to zero if necessary, at the cost of larger error bars.

We found that when measuring the **X power spectrum**, the basic quadratic estimator produces a surprisingly complicated answer, involving not only temperature-polarization cross terms, but also, *e.g.*, autocorrelations of the temperature map with itself. This is unfortunate, since it can make the measured power spectrum more sensitive to systematic errors. Especially if the temperature and polarization maps are made by two different experiments, systematics should be uncorrelated between the two and therefore not contribute to the cross term average. The maximum-likelihood method exhibits the same problem. This can affect E- and B-polarization estimation as well, by giving non-zero weight to the unpolarized map. The problem is eliminated by simply using vanishing cross-power in the fiducial model. We find that this is desirable in practice, since the resulting information loss causes error bars to increase only by negligle amounts, at the percent level.

Finally, we found that on scales substantially smaller than the sky patch, the error bars for the $F^{-1/2}$-method were accurately fit by the approximation of equation (27), where variance scales inversely with sky coverage.

B. Relation to other methods

The quadratic estimator (QE) method is closely related to the maximum-likelihood (ML) method: the latter is simply the quadratic estimator method with $\mathbf{B} = \mathbf{F}^{-1}$ in equation (15), iterated so that the fiducial ("prior") power spectrum equals the measured one [21]. The ML method has the advantage of not requiring any prior to be assumed. The QE method has the advantage of being faster (no iteration) and simpler to interpret — since it is quadratic rather than highly non-linear, the statistical properties the measured band power vector $\mathbf{q}$ can be computed analytically. This allows the likelihood function to be computed directly from $\mathbf{q}$ (as opposed to $\mathbf{x}$), in terms of generalized χ^2-distributions [46].

Both methods are unbiased, but they may differ as regards error bars. The QE method can produce inaccurate error bars if the prior is inconsistent with the actual measurement. The ML method Fisher matrix can produce inaccurate error bar estimates if the measured power spectra have substantial scatter due to noise or sample variance, in which case they are unlikely to describe the smoother true spectra. A good compromise is therefore to iterate the QE method once and choose the second prior to be a rather smooth model consistent with the original measurement. In addition, as mentioned above, we found that it is useful to set the X power spectrum to zero in the prior.

The difference between the QE and ML methods is often small in practice, which can be understood as follows. Since the QE method can be shown to be lossless if the prior equals the truth, thereby minimizing the error bars, small departures from the true prior merely destroy information to second order. This is also why adopting a prior without X power inflated the error bars so little.

The analysis of weak gravitational lensing data is rather analogous to that of CMB polarization, since the projected shear field can be decomposed into components corresponding to E and B. Recent lensing work cast in this language has included both theoretical predictions for various effects [47,48] and data analysis issues [49]. In particular, no E/B leakage was detected at the 10% level when the ML method was applied to simulated examples [49], which can be understood from our present results since the band power bins used were substantially broader than $\Delta\ell$ [49] and mainly on scales $\ell \gg \Delta\ell$.

C. Outlook

Our basic results are good news: although polarized power spectrum estimation adds several complications to the non-polarized case, they can all be dealt with using the techniques we have described. However, much theoretical work remains to be done. Here are a few examples of areas deserving further study:

- It will be interesting to measure six rather than four power spectra from future CMB data sets. Although the $T-B$ and $E-B$ cross-correlations are expected to vanish in the standard cosmological model, they could constrain or detect exotic parity-violating interactions [50].
- The effects of pixel shape and size in the mapmaking process needs to be quantified, and is likely to be more important for the polarized case.
- Although the methods we have discussed apply equally well to measuring the power spectra of non-Gaussian signals (the only change is that the error bars will no longer be given by equation (14)), non-Gaussian signals contain more information than is contained in their power spectra. Polarized non-Gaussian fluctuations are expected from microwave foregrounds [51–53], secondary anisotropies [54], topological defects [47] and CMB lensing [48], and only a small number of papers have so far addressed the issue of how to best measure such non-Gaussian signatures in practice [55–57].

First and foremost, however, we need a detection of CMB polarization!

The authors wish to thank Wayne Hu, Brian Keating and Matias Zaldarriaga for helpful comments, and the B2001, PIQUE and POLAR teams for encouraging us to write up these calculations. Support for this work was provided by NASA grant NAG5-9194, NSF grant AST00-71213 and the University of Pennsylvania Research Foundation.

APPENDIX A: COMPUTING THE POLARIZATION COVARIANCE MATRIX

This Appendix is intended for the reader who wishes to write software to explicitly compute the polarization covariance matrix. The complete formalism for describing CMB polarization was presented in [1,2], and extended with a number of useful explicit formulas in [11]. However, a number of practical details are not covered in the literature, *e.g.*, the rotation angles and degenerate cases, so we describe all steps explicitly for completeness.

Let the 3-dimensional vector $\mathbf{x}_i$ denote the three measurable quantities for the i^{th} pixel:

$$\mathbf{x}_i \equiv \begin{pmatrix} T_i \\ Q_i \\ U_i \end{pmatrix}. \quad \text{(A1)}$$

The 3×3 covariance matrix between two such vectors at different points can we written

$$\langle \mathbf{x}_i \mathbf{x}_j \rangle = \mathbf{R}(\alpha_{ij})\mathbf{M}(\widehat{\mathbf{r}}_i\cdot\widehat{\mathbf{r}}_j)\mathbf{R}(\alpha_{ji})^t, \quad \text{(A2)}$$

Here $\mathbf{M}$ is the covariance using a (Q,U)-convention where the reference direction is the great circle connecting the two points, and the rotation matrices given by

$$\mathbf{R}(\alpha) \equiv \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos 2\alpha & \sin 2\alpha \\ 0 & -\sin 2\alpha & \cos 2\alpha \end{pmatrix} \quad \text{(A3)}$$

accomplish a rotation into a global reference frame where the reference directions are meridians. The full $(3n)\times(3n)$ map covariance matrix is readily assembled out of the 3×3 blocks of equation (A2) by looping over all pixel pairs. A powerful probe for bugs is making sure that this large matrix is positive definite.

1. Computing the rotation angles

As we will see, computing the magnitudes of the rotation angles α_{ij} is straightforward, whereas getting the correct sign is somewhat subtle.

The great circle connecting the two pixels has the unit normal vector

$$\widehat{\mathbf{r}}_{ij} \equiv \frac{\widehat{\mathbf{r}}_i\times\widehat{\mathbf{r}}_j}{|\widehat{\mathbf{r}}_i\times\widehat{\mathbf{r}}_j|}. \quad \text{(A4)}$$

Similarly, the meridian passing though pixel i (its reference circle for our global (Q,U)-convention) has the unit normal vector

$$\widehat{\mathbf{r}}_i^* \equiv \frac{\widehat{\mathbf{z}}\times\widehat{\mathbf{r}}_i}{|\widehat{\mathbf{z}}\times\widehat{\mathbf{r}}_i|}, \quad \text{(A5)}$$

where $\widehat{\mathbf{z}}=(0,0,1)$ is the unit vector in the z-direction. The magnitude of α_{ij}, the rotation angle for pixel i, is simply the angle between these two great circles, so $\cos\alpha_{ij}=\widehat{\mathbf{r}}_{ij}\cdot\widehat{\mathbf{r}}_i^*$. The sign of α_{ij} is defined so that a positive angle corresponds to clockwise rotation at the pixel (at $\widehat{\mathbf{r}}_i$). We therefore compute the cross product of the two circle normals, which has the property that

$$\widehat{\mathbf{r}}_{ij}\times\widehat{\mathbf{r}}_i^* = \widehat{\mathbf{r}}_i \sin\alpha_{ij}. \quad \text{(A6)}$$

(Since both $\widehat{\mathbf{r}}_{ij}$ and $\widehat{\mathbf{r}}_i^*$ are by construction perpendicular to $\widehat{\mathbf{r}}_i$, their cross product will be either aligned or anti-aligned with $\widehat{\mathbf{r}}_i$.) Dotting equation (A6) with $\widehat{\mathbf{r}}_i$ and performing some vector algebra gives

$$\begin{aligned}\sin\alpha_{ij} &= (\widehat{\mathbf{r}}_{ij}\times\widehat{\mathbf{r}}_i^*)\cdot\widehat{\mathbf{r}}_i \propto [(\widehat{\mathbf{r}}_i\times\widehat{\mathbf{r}}_j)\times(\widehat{\mathbf{r}}_i\times\widehat{\mathbf{z}})]\cdot\widehat{\mathbf{r}}_i \\ &= [(\widehat{\mathbf{r}}_i\times\widehat{\mathbf{r}}_j)\cdot\widehat{\mathbf{z}}]\,\widehat{\mathbf{r}}_i\cdot\widehat{\mathbf{r}}_i \propto \widehat{\mathbf{r}}_{ij}\cdot\widehat{\mathbf{z}}, \end{aligned} \quad \text{(A7)}$$

where the omitted proportionality constants are positive. α_{ij} therefore has the same sign as the z-coordinate of $\widehat{\mathbf{r}}_{ij}$, and is given by

$$\alpha_{ij} = \begin{cases} +\cos^{-1}(\widehat{\mathbf{r}}_{ij}\cdot\widehat{\mathbf{r}}_i^*) & \text{if } \widehat{\mathbf{r}}_{ij}\cdot\widehat{\mathbf{z}}>0, \\ -\cos^{-1}(\widehat{\mathbf{r}}_{ij}\cdot\widehat{\mathbf{r}}_i^*) & \text{if } \widehat{\mathbf{r}}_{ij}\cdot\widehat{\mathbf{z}}<0. \end{cases} \quad \text{(A8)}$$

For generic pairs of directions, equation (A8) gives the two rotation angles α_{ij} and α_{ji} needed for equation (A2). However, it breaks down for the three special cases $\widehat{\mathbf{r}}_i\times\widehat{\mathbf{r}}_j=\mathbf{0}$, $\widehat{\mathbf{r}}_i\times\widehat{\mathbf{z}}=\mathbf{0}$ and $\widehat{\mathbf{r}}_j\times\widehat{\mathbf{z}}=\mathbf{0}$. If $\widehat{\mathbf{r}}_i\times\widehat{\mathbf{r}}_j=\mathbf{0}$, the two pixels are either identical or on diametrically opposite sides of the sky. Hence any great circle through $\widehat{\mathbf{r}}_i$ will go through $\widehat{\mathbf{r}}_2$ as well. We can choose this circle to be the meridian, so no rotation is needed, *i.e.*, $\alpha_{ij}=\alpha_{ji}=0$ for this case. Indeed, $\mathbf{M}$ comes out diagonal for this case by symmetry, with $\langle Q_iQ_j\rangle=\langle U_iU_j\rangle$, so rotations have no effect.

If $\widehat{\mathbf{r}}_i\times\widehat{\mathbf{z}}=\mathbf{0}$, then the pixel is at the North or South pole, making our the global (Q,U)-convention undefined. The simplest remedy to this problem is to move the pixel away from the pole by a tiny amount much smaller than the beam width of the experiment.

2. Computing the matrix M

The $\mathbf{M}$-matrix depends only on the angular separation between the two pixels. It is given by [11]

$$\mathbf{M}(\widehat{\mathbf{r}}_i\cdot\widehat{\mathbf{r}}_j) \equiv \begin{pmatrix} \langle T_iT_j\rangle & \langle T_iQ_j\rangle & 0 \\ \langle T_iQ_j\rangle & \langle Q_iQ_j\rangle & 0 \\ 0 & 0 & \langle U_iU_j\rangle \end{pmatrix}, \quad \text{(A9)}$$

$$\langle T_iT_j\rangle \equiv \sum_\ell \left(\frac{2\ell+1}{4\pi}\right) P_\ell(z) C_\ell^T, \quad \text{(A10)}$$

$$\langle T_iQ_j\rangle \equiv \sum_\ell \left(\frac{2\ell+1}{4\pi}\right) F_\ell^{10}(z) C_\ell^X, \quad \text{(A11)}$$

$$\langle Q_iQ_j\rangle \equiv \sum_\ell \left(\frac{2\ell+1}{4\pi}\right) \left[F_\ell^{12}(z)C_\ell^E - F_\ell^{22}(z)C_\ell^B\right], \quad \text{(A12)}$$

$$\langle U_iU_j\rangle \equiv \sum_\ell \left(\frac{2\ell+1}{4\pi}\right) \left[F_\ell^{12}(z)C_\ell^B - F_\ell^{22}(z)C_\ell^E\right], \quad \text{(A13)}$$

where $\mathbf{z} = \widehat{\mathbf{r}}_i \cdot \widehat{\mathbf{r}}_j$ is the cosine of the angle between the two pixels under consideration. P_ℓ denotes a Legendre polynomial, and the functions F^{10}, F^{12} and F^{22} are [11]

$$F^{10}(z) = 2\frac{\frac{\ell z}{(1-z^2)}P_{\ell-1} - \left(\frac{\ell}{1-z^2} + \frac{\ell(\ell-1)}{2}\right)P_\ell(\ell)}{[(\ell-1)\ell(\ell+1)(\ell+2)]^{1/2}} \quad \text{(A14)}$$

$$F^{12}(z) = 2\frac{\frac{(\ell+2)z}{(1-z^2)}P^2_{\ell-1} - \left(\frac{\ell-4}{1-z^2} + \frac{\ell(\ell-1)}{2}\right)P^2_\ell(\ell)}{(\ell-1)\ell(\ell+1)(\ell+2)} \quad \text{(A15)}$$

$$F^{22}(z) = 4\frac{(\ell+2)P^2_{\ell-1} - (\ell-1)zP^2_\ell(\ell)}{(\ell-1)\ell(\ell+1)(\ell+2)(1-z^2)} \quad \text{(A16)}$$

Here P_ℓ and P^2_ℓ denote Legendre polynomials P^m_ℓ for the cases $m = 0$ and $m = 2$, respectively, which are conveniently computed using the recursion relations [58]

$$P_\ell(z) = \begin{cases} 1 & \text{for } \ell = 0, \\ z & \text{for } \ell = 1, \\ \frac{(2\ell-1)zP_{\ell-1}(z)-(\ell-1)P_{\ell-2}(z)}{\ell} & \text{for } \ell \geq 2, \end{cases} \quad \text{(A17)}$$

$$P^2_\ell(z) = \begin{cases} 3(1-z^2) & \text{for } \ell = 2, \\ 5zP^2_2 & \text{for } \ell = 3, \\ \frac{(2\ell-1)zP^2_{\ell-1}(z)-(\ell+1)P^2_{\ell-2}(z)}{\ell-2} & \text{for } \ell \geq 4. \end{cases} \quad \text{(A18)}$$

The division by $(1-z^2)$ unfortunately causes expressions (A14)–(A16) to blow up numerically when $z = \pm 1$, *i.e.*, for correlations at zero separation or between diametrically opposite pixels in the sky. Taking the appropriate limits for these cases gives

$$F^{10}_\ell(z) = 0 \qquad \text{if } |z| = 1, \quad \text{(A19)}$$

$$F^{12}_\ell(z) = \begin{cases} \frac{1}{2} & \text{if } z = 1, \\ \frac{1}{2}(-1)^\ell & \text{if } z = -1, \end{cases} \quad \text{(A20)}$$

$$F^{22}_\ell(z) = \begin{cases} -\frac{1}{2} & \text{if } z = 1, \\ \frac{1}{2}(-1)^\ell & \text{if } z = -1. \end{cases} \quad \text{(A21)}$$

3. Including variable angular resolution

If the experimental beam is rotationally symmetric, its effect is straightforward to include even if the beam size and radial profile varies between pixels. This complication is particularly important for the case of cross-polarization, where it may be desirable to correlate polarization maps from one experiment with a temperature map from another experiment that happens to have different angular resolution.

Let $B_{i\ell}$ denote the coefficients obtained from expanding the beam profile corresponding to the measurement x_i in Legendre polynomials. For a Gaussian beam, these coefficients are accurately described by the well-known approximation

$$B_{i\ell} \approx e^{-\frac{1}{2}\theta_i^2\ell(\ell+1)}, \quad \text{(A22)}$$

where the rms beam width θ_i is related to the full-width-half-maximum by θ_i =FWHM$_i/\sqrt{8\ln 2}$. To compute the correlation $\langle x_i x_j \rangle$ in equation (A2), the four power spectra C^P_ℓ are simply replaced by $C^P_\ell B_{i\ell}B_{j\ell}$ in equations (A10) through (A13).

[1] M. Kamionkowski, A. Kosowsky, and A. Stebbins, Phys. Rev. D **55**, 7368 (1997).

[2] M. Zaldarriaga and U. Seljak, Phys. Rev. D **55**, 1830 (1997).

[3] W. Hu and M. White, New Astron **2**, 323 (1997).

[4] M. Zaldarriaga, D. N. Spergel, and U. Seljak, ApJ **488**, 1 (1997).

[5] D. J. Eisenstein, W. Hu, and Tegmark M, ApJ **518**, 2 (1999).

[6] M. Tegmark, D. J. Eisenstein, W. Hu, and A. de Oliveira-Costa, ApJ **530**, 133 (2000).

[7] M. M. Hedman, D. Barkats, J. O. Gundersen, S. T. Staggs, and B. Winstein, astro-ph/0010592 (2000).

[8] S. T. Staggs, J. O. Gundersen, and S. E. Church 1999, astro-ph/9904062, in *Microwave Foregrounds*, ed. A. de Oliveira-Costa & M. Tegmark (ASP: San Francisco)

[9] G. F. Smoot *et al.*, ApJ **396**, L1 (1992).

[10] B. Keating, P. Timbie, A. Polnarev, and J. Steinberger, ApJ **495**, 580 (1998).

[11] M. Zaldarriaga, ApJ **503**, 1 (1998).

[12] J. B. Peterson, J. E. Carlstrom, E. S. Cheng, M. Kamionkowski, A. E. Lange, M. Seiffert, D. N. Spergel, and Stebbins A, astro-ph/9907276 (1999).

[13] M. Kamionkowski and A. H. Jaffe, astro-ph/0011329 (2000).

[14] R. A. Fisher, J. Roy. Stat. Soc. **98**, 39 (1935).

[15] M. Tegmark, A. N. Taylor, and A. F. Heavens, ApJ **480**, 22 (1997).

[16] A. Jaffe, M. Kamionkowski, and L. Wang, astro-ph/9909281 (1999).

[17] A. Kosowsky *et al.* 2000, in preparation.

[18] M. Tegmark, ApJ **480**, L87 (1997).

[19] E. L. Wright, astro-ph/9612006 (1996).

[20] M. Tegmark, Phys. Rev. D **55**, 5895 (1997).

[21] J. R. Bond, A. H. Jaffe, and L. E. Knox, ApJ **533**, 19 (2000).

[22] E. L. Wright, G. Hinshaw, and C. L. Bennett, ApJ **458**, L53 (1996).

[23] M. Tegmark, Phys. Rev. D **56**, 4514 (1997).

[24] A. de Oliveira-Costa, M. Devlin, T. Herbig, A. D. Miller, C. B. Netterfield, L. A. Page, and M. Tegmark, ApJL **509**, L77 (1998).

[25] J. Borrill, astro-ph/9911389 (1999).

[26] B. Revenu, A. Kim, R. Ansari, F. Couchot, J. Delabrouille, and J. Kaplan, A& AS **142**, 499 (2000).

[27] P. de Bernardis *et al.*, Nature **404**, 955 (2000).
[28] A. E. Lange *et al.*, astro-ph/0005004 (2000).
[29] S. Hanany, astro-ph/0005123 (2000).
[30] I. Szapudi, S. Prunet, D. Pogosyan, A. S. Szalay, and J. R. Bond, astro-ph/0010256 (2000).
[31] Y. Xu, M. Tegmark, A. de Oliveira-Costa, M. Devlin, T. Herbig, A. Miller, C. B. Netterfield, and L. Page, astro-ph/0010552 (2000).
[32] M. Tegmark and A. J. S Hamilton, astro-ph/9702019 (1997).
[33] M. Tegmark, ApJL **470**, L81 (1996).
[34] K. M. Gorski, B. D. Wandelt, F. K. Hansen, E. Hivon, and A. J. Banday, astro-ph/9905275 (1999).
[35] U. Seljak and M. Zaldarriaga, ApJ **469**, 437 (1996).
[36] M. Tegmark, M. Zaldarriaga, and A. J. S Hamilton, astro-ph/0008167 (2000).
[37] M. Tegmark, MNRAS **280**, 299 (1996).
[38] A. J. S Hamilton, M. Tegmark, and N. Padmanabhan, astro-ph/0004334 (2000).
[39] W. H. Kinney, Phys. Rev. D **58**, 123506 (1998).
[40] L. Knox, astro-ph/9811358 (1999).
[41] S. Prunet, S. K. Sethi, and F. R. Bouchet, MNRAS **314**, 348 (2000).
[42] L. Popa, C. Burigana, F. Finelli, and N. Mandolesi, astro-ph/0009417 (2000).
[43] M. A. Bucher, K. Moodley, and N. Turok, astro-ph/0011025 (2000).
[44] L. Knox, Phys. Rev. D **48**, 3502 (1995).
[45] D. Scott, M. Srednicki, and M. White, ApJL **421**, L5 (1994).
[46] B. Wandeldt, E. Hivon, and K. M. Gorski;1998;astro-ph/9808292
[47] K. Benabed and F. Bernardeau, astro-ph/9906161 (1999).
[48] K. Benabed, F. Bernardeau , and L. van Waerbeke, astro-ph/astro-ph/0003038 (2000).
[49] W. Hu and M. White, astro-ph/0010352 (2000).
[50] A. Lue, L. Wang, and M. Kamionkowski, Phys. Rev. Lett. **83**, 1506 (1999).
[51] A. Kogut and G. Hinshaw, astro-ph/0006159 (2000).
[52] M. Tucci, E. Carretti, S. Cecchini, R. Fabbri, M. Orsini, and E. Pierpaoli, New Astron. **5**, 181 (2000).
[53] C. Baccigalupi, C. Burigana, F. Perrotta, G. De Zotti, L. La Porta, D. Maino, M. Maris, and R. Paladini, astro-ph/0009135 (2000).
[54] W. Hu, ApJ **529**, 12 (2000).
[55] L. Popa, P. Stefanescu, and R. Fabbri, astro-ph/9812462 (1998).
[56] A. Dolgov, A. Doroshkevich, D. Novikov, and I. Novikov, Int. J. Mod. Phys. D **8**, 189 (1999).
[57] E. Kotok, P. Naselsky, D. Novikov, and I. Novikov, astro-ph/0011521 (2000).
[58] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series and Products, 6th ed.* (New York, Academic Press, 2000).