PHYSICAL REVIEW D **94,** 103512 (2016)

Testing cosmological models with large-scale power modulation using microwave background polarization observations

Emory F. Bunn* and Qingyang Xue
Physics Department, University of Richmond, Richmond, Virginia 23173, USA

Haoxuan Zheng
Physics Department, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA
(Received 18 August 2016; published 14 November 2016)

We examine the degree to which observations of large-scale cosmic microwave background (CMB) polarization can shed light on the puzzling large-scale power modulation in maps of CMB anisotropy. We consider a phenomenological model in which the observed anomaly is caused by modulation of large-scale primordial curvature perturbations and calculate Fisher information and error forecasts for future polarization data, constrained by the existing CMB anisotropy data. Because a significant fraction of the available information is contained in correlations with the anomalous temperature data, it is essential to account for these constraints. We also present a systematic approach to finding a set of normal modes that maximize the available information, generalizing the well-known Karhunen-Loève transformation to take account of the constraints from the temperature data. A polarization map covering at least ∼60% of the sky should be able to provide a 3σ detection of modulation at the level favored by the temperature data. A significant fraction of the information in such a data set is contained in the single mode that optimally encapsulates the signal due to temperature-polarization correlation.

DOI: 10.1103/PhysRevD.94.103512

I. INTRODUCTION

The cosmic microwave background (CMB) radiation provides much of the most important evidence in support of the standard cosmological model [1–3]. However, there have been claims of various "anomalies" on large angular scales in the all-sky maps made by the WMAP and Planck satellites, which appear to be in tension with certain aspects of the model [4–9]. Some, such as the alignment of low-order multipoles [4–7,10–13] and the dipolar modulation of fluctuations [4–6,14–21], even appear to violate the assumptions of homogeneity and isotropy. If there is indeed strong evidence that these assumptions are violated, the effect on cosmology would be revolutionary.

The statistical significance of these anomalies is controversial (e.g., Refs. [4,22]), in large part because they are quantified with *a posteriori* statistics—that is, the anomalies were noticed in the data, and subsequently statistics were devised to quantify their improbability. Such statistics are problematic; in any large data set, some patterns will arise merely by chance, and statistics designed after the fact to characterize these patterns will have artificially low p-values. (This problem is sometimes described as the "look-elsewhere effect.") One might therefore choose to disregard the subject entirely. On the other hand, if the anomalies are not mere flukes, the consequences would be of the highest importance. It therefore seems reasonable to examine them closely while maintaining skepticism.

When faced with the problem of *a posteriori* statistics, the natural solution is to seek a new data set for which the questions can be addressed *a priori*. The large-angle CMB intensity has already been measured to the limit of cosmic variance, but CMB polarization on the largest angular scales has not yet been thoroughly characterized. In this paper, we examine the degree to which all-sky or partial-sky polarization data could help us to determine the significance of the dipolar modulation in fluctuation power.

Although measurements of CMB polarization have been made, there are none that have reliable information on the large angular scales of chief interest to us. We therefore do not use existing polarization data to constrain theories. (In contrast, see Refs. [23,24].) Rather, our focus is on the question of how much light *future* polarization data, with reliable large-angle information, would shed on the modulation.

We focus specifically on the dipolar power modulation—that is, the observation that the fluctuations in one half of the sky appear to be slightly larger in amplitude than in the other half. We choose to examine this anomaly because it appears in some ways more robust than the others. In particular, when maps that have been filtered to contain nonoverlapping multipole ranges are used to calculate the modulation direction, the results are remarkably consistent [17]. Under the hypothesis of statistical isotropy, these directions would be independent random variables. Even if we regard the first of these determinations as besmirched

*ebunn@richmond.edu

with the stain of *a posteriori* statistics, the remainder are untainted.

Possible explanations for any of the anomalies come in three categories: an anomaly can be a fluke, the result of a systematic error, or a sign of new physics. In this paper, we disregard the possibility of systematic error, as the modulation appears robustly in different data sets (WMAP and Planck), analyzed in different ways. We are therefore interested in the question of how well future polarization data could distinguish between the fluke hypothesis and the new-physics hypothesis.

Under the fluke hypothesis, the Universe is described by the standard statistically isotropic Gaussian model, and the probability distribution for polarization observations is straightforward to calculate. Note that the relevant probability distribution is the conditional probability for the future polarization data, constrained by the existing anisotropy data, as emphasized in Refs. [25–27].

Previous work [25–27] has assessed the ability of polarization data to test the fluke hypothesis. In this paper, we go further by comparing this hypothesis with an alternative in which statistical isotropy is broken. The most straightforward candidates for the new-physics hypothesis involve modulating the primordial perturbations with a long-wavelength mode [28–34], although there are other possibilities as well [35]. Such a modulation can be produced in inflationary scenarios with a curvaton field, among other ways. Rather than committing to a specific physical model, we represent the new-physics hypothesis with a phenomenological model originally explored by Dvorkin *et al.* [28]. We suppose that the primordial Newtonian curvature fluctuation Φ is modulated by a multiplicative perturbation that breaks statistical homogeneity and isotropy. To be specific, we suppose that

$$\Phi(\mathbf{r}) = g_1(\mathbf{r})[1 + h(\mathbf{r})] + g_2(\mathbf{r}). \tag{1}$$

Here, g_1 and g_2 are homogeneous and isotropic Gaussian random processes, of the sort found in the standard model. The modulating field h contains only very long-wavelength terms. In fact, to explain the dipolar modulation of the CMB, it can be taken to be a simple gradient,

$$h(\mathbf{r}) = \frac{\mathbf{w}\cdot\mathbf{r}}{R_{LSS}}, \tag{2}$$

for some constant vector $\mathbf{w}$. Here, R_{LSS} is the distance to the last-scattering surface and is introduced to make $\mathbf{w}$ dimensionless. Note that the magnitude of $\mathbf{w}$ differs from the parameter w_1 of Ref. [28] by a fixed normalization factor: $w \equiv |\mathbf{w}| = \sqrt{3/(4\pi)}w_1$.

The reason for the two fields g_1, g_2 is that the modulation does not appear to persist to arbitrarily small length scales [36,37]. As a result, we place the large-scale, modulated Fourier modes in g_1 and the small-scale, unmodulated modes in g_2. We adopt the simplest possible prescription: we let g_1 contain all Fourier modes below a fixed wave number cutoff $k_{\max}$ and place all modes with $k > k_{\max}$ in the unmodulated field g_2. The power spectra of g_1 and g_2 are taken to be the standard power spectrum $P(k) = Ak^{n-4}$ with spectral index $n = 1$, for $k < k_{\max}$ and $k > k_{\max}$ respectively.

We wish to quantify the new information that could be gained about this theory by a future polarization data set. Because CMB polarization is correlated with the temperature anisotropy, which has already been well measured, we should consider the conditional probability distribution of the future polarization data, given the temperature anisotropy data we already have. In the theories under consideration, all of the probability distributions are Gaussian. To be specific, the joint probability distribution for temperature and polarization (whether expressed in position space or in the spherical harmonic basis) is a multivariate normal distribution with zero mean. The conditional probability for polarization given temperature is then a normal distribution with nonzero mean. Both the mean and the covariance matrix of this distribution depend on the theory under consideration. To assess the ability of this data set to distinguish among competing theories, we will examine the theory dependence of the distribution.

In particular, we will calculate the Fisher information for the parameter w and show that for sufficiently large $k_{\max}$ a polarization data set could measure a w-value at the level suggested by the temperature data with $\sim 3\sigma$ significance, even with incomplete sky coverage. We will also show that a significant fraction of the information in such a data set is contained in a single mode, resulting from the correlation of the polarization with the known temperature data.

II. FORMALISM

A. Constrained Gaussian random processes

We begin by summarizing some results regarding constrained Gaussian random processes [38–43]. To be specific, we consider a set of data that can be modeled as a sample of a Gaussian random process with zero mean. Suppose that a subset of the data, represented by the vector $\vec{d}_1$, has been measured and that measurement of an additional data set $\vec{d}_2$ is planned. In the following sections of this paper, $\vec{d}_1$ will be the CMB intensity data already measured by Planck, and $\vec{d}_2$ will be a future set of polarization data. We combine the two data sets into a single vector:

$$\vec{d}_{\text{all}} = \begin{pmatrix} \vec{d}_1 \\ \vec{d}_2 \end{pmatrix}. \tag{3}$$

The combined data is a sample from a Gaussian random process with mean zero and covariance matrix

$$\mathbf{M}_{\rm all} = \langle \vec{d}_{\rm all} \vec{d}_{\rm all}^T \rangle = \begin{pmatrix} \mathbf{M}_{11} & \mathbf{M}_{21}^T \\ \mathbf{M}_{21} & \mathbf{M}_{22} \end{pmatrix}, \tag{4}$$

where T denotes the transpose and $\mathbf{M}_{jk} = \langle \vec{d}_j \vec{d}_k^T \rangle$. The matrix $\mathbf{M}_{\rm all}$ depends on a set of theory parameters $\vec{\theta}$. The likelihood function is $p(\vec{d}_{\rm all}|\vec{\theta}) \propto e^{-\chi^2/2}$, with

$$\chi^2 = \vec{d}_{\rm all}^T \mathbf{M}_{\rm all}^{-1} \vec{d}_{\rm all}. \tag{5}$$

Because we have already measured $\vec{d}_1$, we wish to know the conditional probability $p(\vec{d}_2|\vec{d}_1, \vec{\theta})$. This is still proportional to $e^{-\chi^2/2}$, but now we regard $\vec{d}_1$ as fixed.

Let

$$\mathbf{M}_{\rm all}^{-1} = \mathbf{N} = \begin{pmatrix} \mathbf{N}_{11} & \mathbf{N}_{21}^T \\ \mathbf{N}_{21} & \mathbf{N}_{22} \end{pmatrix}. \tag{6}$$

Then,

$$\chi^2 = \vec{d}_1^T \mathbf{N}_{11} \vec{d}_1 + \vec{d}_1^T \mathbf{N}_{21}^T \vec{d}_2 + \vec{d}_2^T \mathbf{N}_{21} \vec{d}_1 + \vec{d}_2^T \mathbf{N}_{22} \vec{d}_2 \tag{7}$$

$$= (\vec{d}_2 - \vec{\mu})^T \mathbf{N}_{22} (\vec{d}_2 - \vec{\mu}) + \text{constant}, \tag{8}$$

where

$$\vec{\mu} = -\mathbf{N}_{22}^{-1} \mathbf{N}_{21} \vec{d}_1. \tag{9}$$

In summary, when we take $\vec{d}_1$ as fixed, our theory gives a Gaussian likelihood function for $\vec{d}_2$ with mean $\vec{\mu}$ and covariance matrix $\mathbf{M}_c \equiv \mathbf{N}_{22}^{-1}$, both of which depend on the theory parameters $\vec{\theta}$.

It is often convenient to use the block inversion formulas,

$$\mathbf{N}_{22}^{-1} = \mathbf{M}_{22} - \mathbf{M}_{21} \mathbf{M}_{11}^{-1} \mathbf{M}_{21}^T, \tag{10}$$

$$\mathbf{N}_{22}^{-1} \mathbf{N}_{21} = -\mathbf{M}_{21} \mathbf{M}_{11}^{-1}, \tag{11}$$

to write

$$\vec{\mu} = \mathbf{M}_{21} \mathbf{M}_{11}^{-1} \vec{d}_1, \tag{12}$$

$$\mathbf{M}_c = \mathbf{M}_{22} - \mathbf{M}_{21} \mathbf{M}_{11}^{-1} \mathbf{M}_{21}^T. \tag{13}$$

The full expression for the constrained likelihood is

$$p(\vec{d}_2|\vec{d}_1, \vec{\theta}) = \frac{\exp\left(-\frac{1}{2}(\vec{d}_2 - \vec{\mu}(\vec{\theta}))^T \mathbf{M}_c(\vec{\theta})^{-1} (\vec{d}_2 - \vec{\mu}(\vec{\theta}))\right)}{(2\pi)^{N/2} \det^{1/2} \mathbf{M}_c(\vec{\theta})}, \tag{14}$$

where N is the dimension of $\vec{d}_2$.

Because we will in general be interested only in this constrained likelihood, we will simplify the notation by writing $\vec{d}$ instead of $\vec{d}_2$ and $\mathbf{M}$ instead of $\mathbf{M}_c$ wherever there is no risk of ambiguity.

B. Fisher information

Suppose that we are interested in measuring a single parameter θ, such as the modulation level w. The expected error on θ is $F^{-1/2}$, where the Fisher information is

$$F \equiv -\langle (\ln f)'' \rangle = \frac{1}{2} \mathrm{Tr}(\mathbf{M}^{-1} \mathbf{M}' \mathbf{M}^{-1} \mathbf{M}') + \vec{\mu}'^T \mathbf{M}^{-1} \vec{\mu}', \tag{15}$$

and the primes denote derivatives with respect to θ. In this expression, all quantities are to be evaluated at the "true" value of θ.

C. Information-maximizing modes

It may be of interest to know what aspects of the new data are most useful in measuring θ. Is it most useful to know large-scale or small-scale information, for instance? Are some parts of the sky more helpful than others? One way to address this sort of question is to suppose that, instead of measuring the entire N-dimensional data vector $\vec{d}$, we measure only its projection onto a small set of normal modes. To be specific, imagine that we measure $\delta_j \equiv \vec{v}_j \cdot \vec{d}$ for some small set of mode vectors $\vec{v}_1, \vec{v}_2, \ldots$. We can then ask which modes maximize the information in the resulting data set.

For a Gaussian random process whose mean is zero (or more generally, whose mean is independent of θ), the answer to this question is the Karhunen-Loève transform [44], which has a long history in cosmology [45–49]. The best modes are the "signal-to-noise eigenmodes" with the largest eigenvalues. For the constrained data we are considering, the situation is more complicated, as the mean of the distribution depends on the parameter θ.

If there are K mode vectors $\vec{v}_j$, arranged in the columns of an $N \times K$ matrix $\mathbf{V}$, then the Fisher information in the data set $\vec{\delta}$ is

$$F_V = \frac{1}{2} \mathrm{Tr}(\mathbf{M}_V^{-1} \mathbf{M}_V' \mathbf{M}_V^{-1} \mathbf{M}_V') + \vec{\mu}'^T \mathbf{V} \mathbf{M}_V^{-1} \mathbf{V}^T \vec{\mu}', \tag{16}$$

where

$$\mathbf{M}_V = \mathbf{V}^T \mathbf{M} \mathbf{V}, \qquad \mathbf{M}_V' = \mathbf{V}^T \mathbf{M}' \mathbf{V}. \tag{17}$$

The information depends only on the subspace spanned by the vectors; i.e., it is invariant under any invertible transformation $\mathbf{V} \to \mathbf{V}\mathbf{A}$. We can therefore without loss of

generality choose the vectors to be orthonormal with respect to the inner product

$$\langle \vec{x}, \vec{y} \rangle = \vec{x}^T \mathbf{M} \vec{y}. \tag{18}$$

With this choice, $\mathbf{M}_V$ is the identity matrix, and

$$F_V = \frac{1}{2}\mathrm{Tr}((\mathbf{V}^T \mathbf{M}' \mathbf{V})^2) + |\mathbf{V}^T \vec{\mu}'|^2. \tag{19}$$

Consider first the case of a single mode ($K = 1$), for which

$$F_V = \frac{1}{2}(\vec{v}^T \mathbf{M}' \vec{v})^2 + (\vec{v} \cdot \vec{\mu}')^2, \tag{20}$$

subject to the constraint $\vec{v}^T \mathbf{M} \vec{v} = 1$. (When considering the case $K = 1$, we omit the subscript on $\vec{v}_1$.) We can solve the problem of maximizing F_V by a variety of standard numerical methods, but if one of the two terms in this expression is much larger than the other, then an approximate solution is easily found. The first term satisfies

$$\frac{1}{2}(\vec{v}^T \mathbf{M}' \vec{v})^2 \le \frac{1}{2}\lambda_{\mathrm{max}}^2, \tag{21}$$

where λ_{max} is the largest eigenvalue in the generalized eigenvalue problem $\mathbf{M}' \vec{u} = \lambda \mathbf{M} \vec{u}$, with equality when $\vec{v}$ is the corresponding eigenvector. The second term satisfies

$$(\vec{v} \cdot \vec{\mu}')^2 \le (\vec{\mu}'^T \mathbf{M}^{-1} \vec{\mu}')^2 \tag{22}$$

with equality when

$$\vec{v} = \frac{\mathbf{M}^{-1} \vec{\mu}'}{\sqrt{\vec{\mu}'^T \mathbf{M}^{-1} \vec{\mu}'}}. \tag{23}$$

If one of the two expressions on the right side of these inequalities is much larger than the other, then a good approximation to the information-maximizing mode is the mode that saturates that inequality.

As we will describe in detail in the next section, for all of the cases we consider, the second term is much larger than the first one, and the information-maximizing mode is therefore well approximated by Eq. (23). Moreover, this mode often contains a significant fraction of the total information. This mode fully captures all of the information that is contained in the way the mean of the probability distribution varies as the parameter θ is changed. All of the remaining modes will contain information associated only with variations in the covariance matrix.

Having chosen the first mode, we can then seek a second mode that supplies the most additional information. To be specific, let $\vec{v}_1$ be given by Eq. (23), and let $\vec{v}_2$ be orthonormal to $\vec{v}_1$ according to the inner product (18). Then, Eq. (19) can be rewritten:

$$F_V = F_{\vec{v}_1} + \frac{1}{2}(\vec{v}_2^T \mathbf{M}'_\perp \vec{v}_2)^2 + (\vec{v}_2 \cdot \vec{w})^2. \tag{24}$$

Here, $F_{\vec{v}_1}$ is the information contained in mode $\vec{v}_1$ alone. The matrix $\mathbf{M}'_\perp$ is the projection of $\mathbf{M}'$ onto the subspace orthogonal to $\vec{v}_1$, and $\vec{w} = \mathbf{M}' \vec{v}_1$. Choosing the optimal $\vec{v}_2$ therefore involves a maximization precisely analogous to that required to find $\vec{v}_1$ [Eq. (20)]. This time, however, as we will see in the next section, the two contributions are comparable for the cases we consider, so neither simple approximate vector is a good solution.

By an argument analogous to that which led to inequalities (21) and (22), the new information is bounded by

$$F_{\vec{v}_2} \le \frac{1}{2}\lambda_{\perp\,\mathrm{max}}^2 + (\vec{w}^T \mathbf{M}_\perp^{-1} \vec{w})^2, \tag{25}$$

where $\lambda_{\perp\,\mathrm{max}}$ is the maximum eigenvalue for the generalized eigenvalue problem $\mathbf{M}'_\perp \vec{u} = \lambda \mathbf{M}_\perp \vec{u}$. As we will see, this quantity is small in the cases we will consider, so the second-best mode is of little interest in comparison to the first.

D. Application to CMB polarization

We will take the previously measured data $\vec{d}_1 \equiv \vec{t}$ to be CMB temperature data, measured over a masked sky. The jth measurement can be written

$$t_j = \sum_{l,m} a_{lm}^T Y_{lm}(\hat{\mathbf{r}}_j) + n_j^T, \tag{26}$$

where a_{lm}^T are the spherical harmonic coefficients, $\hat{\mathbf{r}}_j$ is the location of the jth pixel, and n_j^T is the noise. (For the low-resolution maps we will consider, noise is quite small. Its primary effect is to regularize the inversion of the covariance matrix.) We write this compactly as

$$\vec{t} = \mathbf{Y} \vec{a} + \vec{n}^T. \tag{27}$$

Here, the vector $\vec{a}$ contains the real and imaginary parts of the spherical harmonic coefficients a_{lm}^T, and the matrix $\mathbf{Y}$ contains the real and imaginary parts of the corresponding spherical harmonics evaluated at the pixel locations.

The covariance matrix is

$$\mathbf{M}_{11} \equiv \langle \vec{t}\vec{t}^T \rangle = \mathbf{Y} \mathbf{C}^{TT} \mathbf{Y}^T + \mathbf{N}_T, \tag{28}$$

where $\mathbf{N}_T$ is the noise covariance matrix and $\mathbf{C}^{TT}$ is the covariance matrix of the a_{lm}^T coefficients.

The polarization data $\vec{d}_2$ will consist of polarization measurements, which can be written

$$\vec{d}_2 = \mathbf{Z} \vec{e} + \vec{n}^P. \tag{29}$$

Here, $\vec{e}$ contains the real and imaginary parts of the E-mode polarization coefficients a^E_{lm}, and $\vec{n}^P$ is the noise. (We neglect the contribution of B modes.) The vector $\vec{d}_2$ contains the Stokes parameters Q, U for each pixel. The matrix $\mathbf{Z}$ contains the real and imaginary parts of the contributions of the spin-2 spherical harmonics to each Q- and U-value.

The remaining blocks of the covariance matrix are

$$\mathbf{M}_{22} = \mathbf{Z}\mathbf{C}^{EE}\mathbf{Z}^T + \mathbf{N}_P, \tag{30}$$

$$\mathbf{M}_{21} = \mathbf{Z}\mathbf{C}^{ET}\mathbf{Y}^T. \tag{31}$$

The matrices $\mathbf{C}^{EE}$ and $\mathbf{C}^{ET}$ characterize the covariances of the E coefficients and the ET cross-covariance, and $\mathbf{N}_P$ is the noise covariance matrix.

In the standard, statistically isotropic model, $\mathbf{C}^{TT}$, $\mathbf{C}^{EE}$, and $\mathbf{C}^{ET}$ are diagonal matrices containing the three power spectra. When isotropy is broken, these matrices acquire off-diagonal elements, which are computed according to the detailed recipe in Ref. [28]. In a coordinate system aligned with the direction of isotropy breaking, the off-diagonal elements are nonzero only when the two m-values are equal and the l's differ by 1.

III. RESULTS

We have performed computations for future polarization data sets, constrained by the existing Planck temperature maps [1]. To be specific, we used the Planck COMMANDER CMB map with HEALPix [50] $N_{\rm side} = 256$, downgraded to $N_{\rm side} = 32$ and smoothed with a Gaussian beam of width $\sigma_{\rm beam} = 2°$. We keep all pixels within the High Frequency Instrument Galactic emission mask, retaining 80% of the pixels.

We imagine future polarization data with the same smoothing and $N_{\rm side}$, with signal-to-noise ratio per pixel of 3. We consider five different sky coverage scenarios:

(i) An all-sky map.
(ii) A map with sky coverage $f_{\rm sky} = 0.8$, with a straight Galactic latitude cut, including all pixels whose Galactic latitude satisfies $|b| > \sin^{-1}(1 - f_{\rm sky})$.
(iii) A map with $f_{\rm sky} = 0.6$, with a similar Galactic latitude cut.
(iv) Two maps with $f_{\rm sky} = 0.3$, consisting of spherical caps centered on the north and south Galactic poles. (These two maps combined cover the same area as the $f_{\rm sky} = 0.6$ map.)

For the broken-isotropy hypothesis, we consider five values for the cutoff wave number, namely $k_{\rm max}c/H_0 = 10$, 20, 30, 40, 50. We hold the modulation direction fixed at Galactic coordinates $(l, b) = (226°, -17°)$. All computations are performed after rotating the maps to a coordinate system with this direction at the pole, so that the covariances among the $a^{(T,E)}_{lm}$ coefficients are as simple as possible.

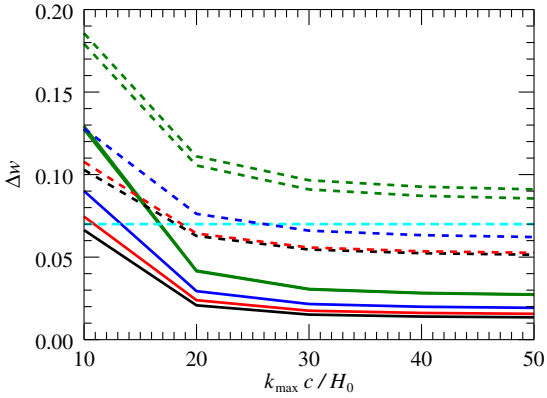

FIG. 1. The error forecast Δw. From bottom to top, the solid curves correspond to polarization data sets with $f_{\rm sky} = 1$ (black), 0.8 (red), 0.6 (blue), 0.3 (green). For $f_{\rm sky} = 0.3$, two virtually identical curves are shown, corresponding to the northern and southern caps. The dashed curves show the error forecasts for a hypothetical experiment in which only the single "best" mode of the polarization data is measured. The horizontal long-dashed line shows the value preferred by the existing temperature data.

The solid curves in Fig. 1 show the projected error $\Delta w \equiv F^{-1/2}$, where F is the Fisher information. The long-dashed line is the value $w = 0.07$, which is roughly the best-fit value from the temperature data. For $k_{\rm max} \gtrsim 30H_0/c$, a strong detection is possible even with relatively low sky coverage.

The dashed curves in the figure show the projected error in the hypothetical scenario where only the information-maximizing mode $\vec{v}_1$, defined in Eq. (23), is measured. Although this single mode is never enough to provide a definitive measurement, it contains a significant fraction of the total Fisher information, ranging from approximately 48% when $k_{\rm max} = 10H_0/c$ to 9% when $k_{\rm max} = 50H_0/c$.

As noted in Sec. II C, $\vec{v}_1$ is in fact an approximation to the information-maximizing mode. The quality of the approximation is determined by the ratio of the two bounds in inequalities (21) and (22). For the models plotted in the figure, this ratio always exceeds 30, which implies that the information contained in mode $\vec{v}_1$ is within 3% of the maximum possible.

The information contained in the second-best mode is bounded by the inequality (25). In almost all of the cases plotted, the ratio of this bound to the information contained in the first mode is less than 8%, indicating that far more information is contained in the first mode than in any other individual mode. The only exceptions occur when $k_{\rm max} = 10H_0/c$ and $f_{\rm sky} = 0.3$, in which case the total information is quite low. In all cases considered, the two terms in (25) are comparable, differing by no more than a factor of 3, so neither simple approximation would work well for finding the

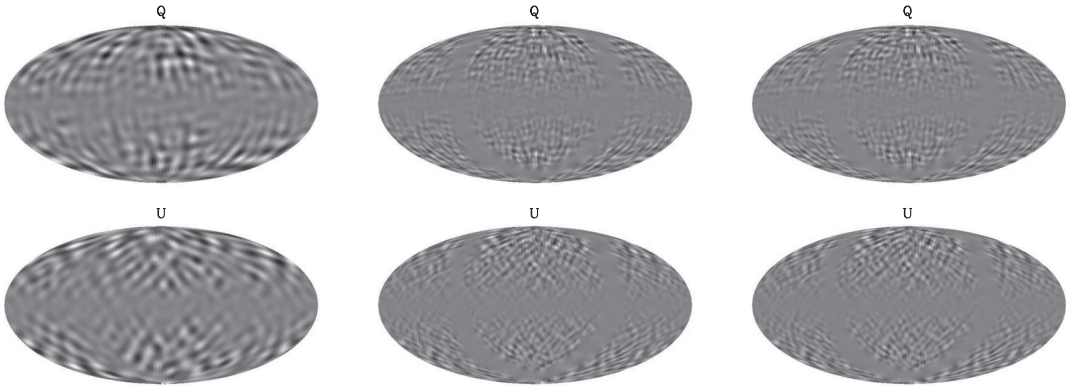

FIG. 2. The single mode that best constrains w, for an all-sky polarization data set. The top panel shows Stokes Q, and the bottom is Stokes U. These modes are for an all-sky polarization data set, constrained by the existing temperature data with 80% sky coverage as described in the text. The maps are oriented so that the modulation direction is at the top. The two bands where the mode is nearly zero are the region of zero modulation (horizontal) and the vicinity of the Galactic plane. The mode has little power near the Galactic plane because this region is unconstrained by temperature data. From left to right, $k_{\rm max} = 10H_0/c, 30H_0/c, 50H_0/c$.

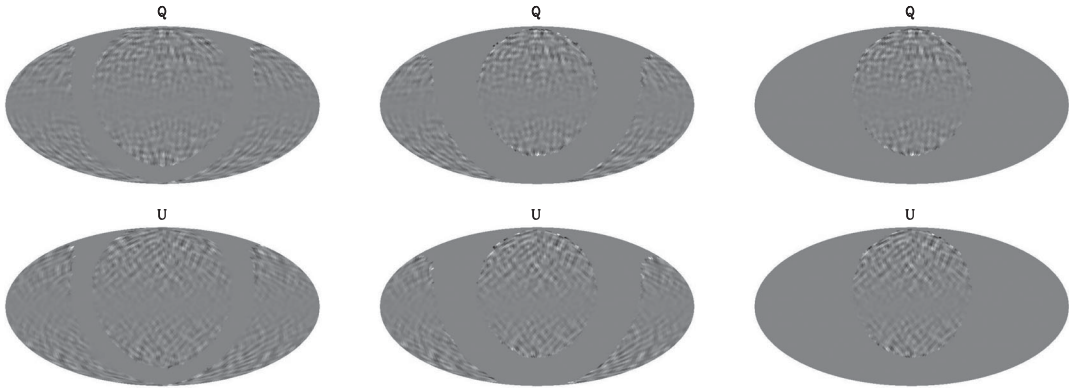

FIG. 3. The best w-constraining mode for polarization data sets with sky coverage $f_{\rm sky} = 0.8, 0.6, 0.3$. For $f_{\rm sky} = 0.3$, the data are assumed to cover a spherical cap centered on the Galactic north pole; in the other cases, a cut in $|b|$ is used. The maps are oriented as in the previous figure.

second-best mode. Since this mode is known to contain little information, we do not pursue its calculation further.

The information-maximizing modes themselves are shown in Figs. 2 and 3. Note that these maps are oriented with the modulation direction, rather than the Galactic north pole, at the top. Because these modes are measuring primarily the correlation with the existing temperature data, they have little power in the Galactic plane. Unsurprisingly, they also have little power in the plane perpendicular to the modulation direction, where the modulation is zero.

IV. DISCUSSION

We have presented Fisher matrix calculations and resulting error forecasts for a future large-scale CMB polarization data set, to assess the degree to which such a data set can shed light on the puzzling large-scale power modulation in the CMB temperature anisotropy. The calculations are based on the probability distribution of the polarization data, conditioned on the already-measured temperature data.

If the observed hemispherical modulation in CMB anisotropy power is not a fluke, a detection of it in CMB polarization should be possible. In the model we have considered, a roughly 3σ detection is possible with a data set that covers at least ~60% of the sky, as long as the modulation affects modes with wave numbers up to $\sim 30H_0/c$. Because these calculations are for the polarization data conditioned on the existing temperature data, this error is associated with the *additional* information available in polarization, on top of what we have already seen in temperature.

The future of CMB polarization measurement on large scales is uncertain, but plans are underway for a ground-based initiative known as CMB-S4 [51]. This experiment would be based in the Atacama desert and would survey a large fraction of the southern sky. The experimenters are also considering including a second telescope in the

northern hemisphere to increase the sky coverage. With only southern sky coverage, this experiment would be roughly comparable to our $f_{\rm sky} = 0.3$ calculations. With the addition of a northern instrument, the sky coverage would be closer to our $f_{\rm sky} = 0.6$. The calculations for larger sky fractions would correspond roughly to a hypothetical satellite mission, which would be in the more distant future. These comparisons are of course extremely rough, as we have not tailored our calculations to match any particular experiment design in detail.

Although the calculations presented herein are based on a simple phenomenological model, we may expect qualitatively similar results from any model in which the observed temperature power modulation is caused by modulation of the large-scale density perturbation modes.

We have presented a formalism for identifying the spatial modes in the data that best constrain the theory. A significant fraction (9% or more) of the information in such a data set is contained in the single "information-maximizing" mode that optimally captures the correlation between the known temperature data and the polarization. The rest of the information is contained in the covariances of the polarization measurements (e.g., the predicted mean-square amplitudes of the various modes), although no single mode in this category contains an amount of information comparable to the information-maximizing mode. Because the correlation with the anomalous temperature data is the source of much of the available information, it is necessary to perform constrained calculations as we have done in order to get reliable forecasts.

ACKNOWLEDGMENTS

This work was supported by NSF Grants No. 0922748 and No. 1410133.

[1] R. Adam *et al.* (Planck Collaboration), Planck 2015 results. I. Overview of products and scientific results, Astron. Astrophys. **594**, A1 (2016).
[2] C. L. Bennett *et al.*, Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) observations: Final maps and results, Astrophys. J. Suppl. Ser. **208**, 20 (2013).
[3] M. Bucher, Physics of the cosmic microwave background anisotropy, Int. J. Mod. Phys. D **24**, 1530004 (2015).
[4] C. L. Bennett *et al.*, Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) observations: Are there cosmic microwave background anomalies?, Astrophys. J. Suppl. Ser. **192**, 17 (2011).
[5] P. A. R. Ade *et al.* (Planck Collaboration), Planck 2013 results. XXIII. Isotropy and statistics of the CMB, Astron. Astrophys. **571**, A23 (2014).
[6] P. A. R. Ade *et al.* (Planck Collaboration), Planck 2015 results. XVI. Isotropy and statistics of the CMB, Astron. Astrophys. **594**, A16 (2016).
[7] D. J. Schwarz, G. D. Starkman, D. Huterer, and C. J. Copi, Is the Low-ℓ Microwave Background Cosmic? Phys. Rev. Lett. **93**, 221301 (2004).
[8] C. J. Copi, D. Huterer, D. J. Schwarz, and G. D. Starkman, On the large-angle anomalies of the microwave sky, Mon. Not. R. Astron. Soc. **367**, 79 (2006).
[9] D. J. Schwarz, C. J. Copi, D. Huterer, and G. D. Starkman, CMB Anomalies after Planck, Classical Quantum Gravity **33**, 184001 (2016).
[10] A. de Oliveira-Costa, M. Tegmark, M. Zaldarriaga, and A. Hamilton, Significance of the largest scale CMB fluctuations in WMAP, Phys. Rev. D **69**, 063516 (2004).
[11] C. J. Copi, D. Huterer, and G. D. Starkman, Multipole vectors: A new representation of the CMB sky and evidence for statistical anisotropy or non-Gaussianity at $2 \leq l \leq 8$, Phys. Rev. D **70**, 043515 (2004).
[12] K. Land and J. Magueijo, The multipole vectors of the Wilkinson Microwave Anisotropy Probe, and their frames and invariants, Mon. Not. R. Astron. Soc. **362**, 838 (2005).
[13] C. J. Copi, D. Huterer, D. J. Schwarz, and G. D. Starkman, Large-scale alignments from WMAP and Planck, Mon. Not. R. Astron. Soc. **449**, 3458 (2015).
[14] H. K. Eriksen, F. K. Hansen, A. J. Banday, K. M. Górski, and P. B. Lilje, Asymmetries in the cosmic microwave background anisotropy field, Astrophys. J. **605**, 14 (2004).
[15] H. K. Eriksen, A. J. Banday, K. M. Górski, F. K. Hansen, and P. B. Lilje, Hemispherical power asymmetry in the Third-Year Wilkinson Microwave Anisotropy Probe sky maps, Astrophys. J. Lett. **660**, L81 (2007).
[16] C. Gordon, Broken isotropy from a linear modulation of the primordial perturbations, Astrophys. J. **656**, 636 (2007).
[17] F. K. Hansen, A. J. Banday, K. M. Górski, H. K. Eriksen, and P. B. Lilje, Power asymmetry in cosmic microwave background fluctuations from full sky to sub-degree scales: Is the Universe isotropic?, Astrophys. J. **704**, 1448 (2009).
[18] J. Hoftuft, H. K. Eriksen, A. J. Banday, K. M. Górski, F. K. Hansen, and P. B. Lilje, Increasing evidence for hemispherical power asymmetry in the Five-Year WMAP data, Astrophys. J. **699**, 985 (2009).
[19] Y. Akrami, Y. Fantaye, A. Shafieloo, H. K. Eriksen, F. K. Hansen, A. J. Banday, and K. M. Górski, Power asymmetry in WMAP and Planck temperature sky maps as measured by a local variance estimator, Astrophys. J. Lett. **784**, L42 (2014).
[20] S. Adhikari, Local variance asymmetries in Planck temperature anisotropy maps, Mon. Not. R. Astron. Soc. **446**, 4232 (2015).
[21] S. Aiola, B. Wang, A. Kosowsky, T. Kahniashvili, and H. Firouzjahi, Microwave background correlations from dipole anisotropy modulation, Phys. Rev. D **92**, 063008 (2015).

[22] G. Efstathiou, The statistical significance of the low cosmic microwave background mulitipoles, Mon. Not. R. Astron. Soc. **346**, L26 (2003).
[23] F. Paci, A. Gruppuso, F. Finelli, P. Cabella, A. de Rosa, N. Mandolesi, and P. Natoli, Power asymmetries in the cosmic microwave background temperature and polarization patterns, Mon. Not. R. Astron. Soc. **407**, 399 (2010).
[24] F. Paci, A. Gruppuso, F. Finelli, A. De Rosa, N. Mandolesi, and P. Natoli, Hemispherical power asymmetries in the WMAP 7-year low-resolution temperature and polarization maps, Mon. Not. R. Astron. Soc. **434**, 3071 (2013).
[25] C. J. Copi, D. Huterer, D. J. Schwarz, and G. D. Starkman, Large-angle cosmic microwave background suppression and polarization predictions, Mon. Not. R. Astron. Soc. **434**, 3590 (2013).
[26] A. Yoho, S. Aiola, C. J. Copi, A. Kosowsky, and G. D. Starkman, Microwave background polarization as a probe of large-angle correlations, Phys. Rev. D **91**, 123504 (2015).
[27] M. O'Dwyer, C. J. Copi, L. Knox, and G. D. Starkman, CMB-S4 and the Hemispherical Variance Anomaly, arXiv: 1608.02234.
[28] C. Dvorkin, H. V. Peiris, and W. Hu, Testable polarization predictions for models of CMB isotropy anomalies, Phys. Rev. D **77**, 063008 (2008).
[29] A. L. Erickcek, S. M. Carroll, and M. Kamionkowski, Superhorizon perturbations and the cosmic microwave background, Phys. Rev. D **78**, 083012 (2008).
[30] A. L. Erickcek, M. Kamionkowski, and S. M. Carroll, A hemispherical power asymmetry from inflation, Phys. Rev. D **78**, 123520 (2008).
[31] A. L. Erickcek, C. M. Hirata, and M. Kamionkowski, A scale-dependent power asymmetry from isocurvature perturbations, Phys. Rev. D **80**, 083507 (2009).
[32] A. Moss, D. Scott, J. P. Zibin, and R. Battye, Tilted physics: A cosmologically dipole-modulated sky, Phys. Rev. D **84**, 023014 (2011).
[33] J. P. Zibin and D. Contreras, Testing physical models for dipolar asymmetry: from temperature to k space to lensing, arXiv:1512.02618.
[34] C. T. Byrnes, D. Regan, D. Seery, and E. R. M. Tarrant, Implications of the cosmic microwave background power asymmetry for the early universe, Phys. Rev. D **93**, 123003 (2016).
[35] L. Dai, D. Jeong, M. Kamionkowski, and J. Chluba, The pesky power asymmetry, Phys. Rev. D **87**, 123005 (2013).
[36] C. M. Hirata, Constraints on cosmic hemispherical power anomalies from quasars, J. Cosmol. Astropart. Phys. 09 (2009) 011.
[37] D. Hanson and A. Lewis, Estimators for CMB statistical anisotropy, Phys. Rev. D **80**, 063004 (2009).
[38] G. B. Rybicki and W. H. Press, Interpolation, realization, and reconstruction of noisy, irregularly sampled data, Astrophys. J. **398**, 169 (1992).
[39] Y. Hoffman and E. Ribak, Constrained realizations of Gaussian fields—A simple algorithm, Astrophys. J. Lett. **380**, L5 (1991).
[40] Y. Hoffman and E. Ribak, Primordial Gaussian perturbation fields—Constrained realizations, Astrophys. J. **384**, 448 (1992).
[41] E. F. Bunn, K. B. Fisher, Y. Hoffman, O. Lahav, J. Silk, and S. Zaroubi, Wiener filtering of the COBE Differential Microwave Radiometer data, Astrophys. J. Lett. **432**, L75 (1994).
[42] O. Lahav, K. B. Fisher, Y. Hoffman, C. A. Scharf, and S. Zaroubi, Wiener reconstruction of all-sky galaxy surveys in spherical harmonics, Astrophys. J. Lett. **423**, L93 (1994).
[43] E. F. Bunn, Y. Hoffman, and J. Silk, The Wiener-filtered COBE DMR data and predictions for the Tenerife experiment, Astrophys. J. **464**, 1 (1996).
[44] K. Karhunen, Über lineare methoden in der wahrscheinlichkeitsrechnung, Ann. Acad. Sci. Fennicae Ser. A I. Math.-Phys. **37**, 1 (1947).
[45] J. R. Bond, Signal-to-Noise Eigenmode Analysis of the Two-Year COBE Maps, Phys. Rev. Lett. **74**, 4369 (1995).
[46] E. F. Bunn, D. Scott, and M. White, The COBE normalization for standard cold dark matter, Astrophys. J. Lett. **441**, L9 (1995).
[47] E. F. Bunn and N. Sugiyama, Cosmological constant cold dark matter models and the COBE two-year sky maps, Astrophys. J. **446**, 49 (1995).
[48] M. Vogeley, Constraints on cosmological models from once and future surveys, in *Wide Field Spectroscopy and the Distant Universe*, edited by S. J. Maddox and A. Aragon-Salamanca (World Scientific, Singapore, 1995), p. 142.
[49] M. S. Vogeley and A. S. Szalay, Eigenmode Analysis of galaxy redshift surveys. I. Theory and methods, Astrophys. J. **465**, 34 (1996).
[50] K. M. Górski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelmann, HEALPix: A framework for high-resolution discretization and fast analysis of data distributed on the sphere, Astrophys. J. **622**, 759 (2005).
[51] K. N. Abazajian *et al.*, CMB-S4 Science Book, First Edition, arXiv:1610.02743.