

CONSTRAINTS ON THE TOPOLOGY OF THE UNIVERSE FROM THE *WILKINSON MICROWAVE ANISOTROPY PROBE* FIRST-YEAR SKY MAPS

N. G. PHILLIPS¹ AND A. KOGUT²
 Received 2004 April 20; accepted 2005 June 17

ABSTRACT

We compute the covariance expected between the spherical harmonic coefficients $a_{\ell m}$ of the cosmic microwave temperature anisotropy if the universe had a compact topology. For a fundamental cell size smaller than the distance to the decoupling surface, off-diagonal components carry more information than the diagonal components (the power spectrum). We use a maximum likelihood analysis to compare the *Wilkinson Microwave Anisotropy Probe* (*WMAP*) first-year data to models with a cubic topology. The data are compatible with finite flat topologies with fundamental domain $L > 1.2$ times the distance to the decoupling surface at a 95% confidence level. The *WMAP* data show reduced power at the quadrupole and octopole, but they do not show the correlations expected for a compact topology and are indistinguishable from infinite models.

Subject headings: cosmic microwave background — cosmology: observations

1. INTRODUCTION

The simplest model for the universe is a spatially homogeneous, isotropic spacetime with a Euclidian (flat) geometry. This simple model is consistent with observations but leaves unaddressed the question of topology or the connectedness of spacetime. Schwarzschild (1900) first noted the possibility of nontrivial topology for the universe even before Einstein’s discovery of his field equations. Almost immediately after Einstein’s discovery, de Sitter (1917) pointed out that the field equations did not constrain the topology. Since general relativity provides no theoretical guidance, we turn to observations to constrain topology.

Observational tests of topology all rely on multiple imaging of distant objects. If the universe is multiply connected with a cell size smaller than the distance to some object, photons from that object can reach the observer via multiple paths. Simply searching the sky for multiply imaged point sources, e.g., quasars, is problematic: since the travel time to each image is different, each image shows the same object at a different time. If source evolution is important, the multiple images may no longer be recognizable as such. The ideal source for topological tests would fill the whole sky with a pattern centered on the observer and emitted at a single time. The cosmic microwave background (CMB) is an excellent approximation to this ideal source.

A number of authors have used the CMB to constrain the topology of the universe. These tests fall into two general categories. A compact topology cannot support spatial structure with wavelength longer than the cell size. The CMB temperature anisotropy will thus be suppressed on angular scales larger than the (projected) cell size. The first category tests use the CMB power spectrum (or its Legendre transform, the two-point correlation function) to test for nontrivial topology.

The CMB in fact shows significantly less power in the quadrupole and octopole than would be expected for a model based on higher order moments. The discrepancy was first detected by the *Cosmic Background Explorer* (*COBE*; Bennett et al. 1996) and verified at much high signal-to-noise ratio by the *Wilkinson*

Microwave Anisotropy Probe (*WMAP*; Bennett et al. 2003a). Figure 1 shows the angular power spectrum of the *WMAP* first-year data compared to the best-fit Λ CDM model (Spergel et al. 2003). Models with compact topology $L \sim 1$ provide a good match to the observed power spectrum, motivating tests of finite-universe models (de Oliveira-Costa & Smoot 1995; Inoue & Sugiyama 2003; Luminet et al. 2003).

The suppression of power on large angular scales is a necessary but not sufficient condition for the existence of a compact topology. The power spectrum is rotationally invariant, averaging over any phase information in the pattern of CMB anisotropy. Such phase information must exist for compact topologies and forms the basis for a second class of tests. A “circles-on-the-sky” search (Cornish et al. 1998) provides a more stringent test for compact topologies. The CMB decoupling surface is a sphere centered on the observer. If the cell size is smaller than the distance to the decoupling surface, the multiple images of this sphere induced by a compact topology will intersect to produce patterns that match along circles. Such circles are not observed, limiting the cell size $L > 1.7$ for a wide class of models (Cornish et al. 2004).

Additional tests are possible. Compact topologies will not produce circles on the sky if the cell size is larger than the distance to the source, since the resulting images will not intersect. Compact topologies with $L > 2$ may still be distinguished using phase information. In this paper, we describe the correlations imposed on the microwave background by the topology. We use this formalism to compare the *WMAP* first-year data to a model with cubic topology and derive constraints on the cell size L .

2. COVARIANCE OF SPHERICAL HARMONIC COEFFICIENTS

On large scales, the CMB temperature anisotropy is given by

$$\Delta T(\hat{x}) = -\frac{1}{3} \Phi(\Delta\tau \hat{x}), \quad (1)$$

where $\Phi(\mathbf{x})$ is the gravitational potential and $\Delta\tau$ is the radius of the decoupling surface. The potential has the harmonic expansion

$$\Phi(\mathbf{x}) = \int d\mu(\mathbf{k}) \Phi_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{x}}, \quad (2)$$

¹ Science Systems and Applications, Inc. (SSAI), NASA Goddard Space Flight Center, Greenbelt, MD 20771; nicholas.g.phillips.1@gssc.nasa.gov.

² Infrared Astrophysics, NASA Goddard Space Flight Center, Code 685, Goddard Space Flight Center, Greenbelt, MD 20771.

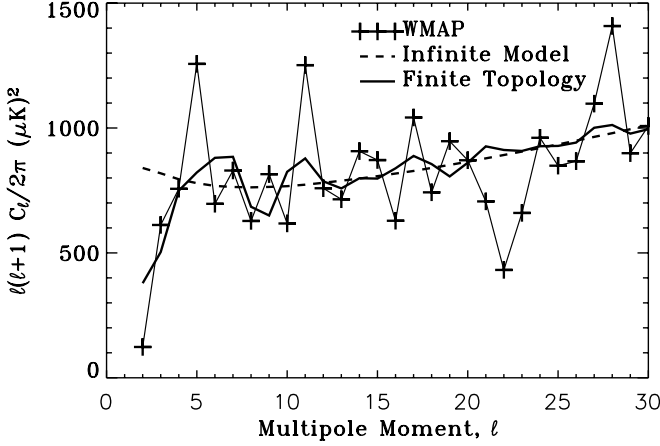


FIG. 1.—Power spectrum of the *WMAP* first-year data, compared to a model for a flat universe with a finite topology. The topology has a cubic fundamental domain with a side length $L = 1.1$ times the distance to the decoupling surface, the best-fit value if only the power spectrum is considered. The dashed line shows the best-fit power spectrum for an open flat universe. The lowest two multipoles are suppressed for the finite topology model, since the universe in such a model is too small to support power at such large scales.

where $d\mu(\mathbf{k})$ reflects the density of states, determined by the topology. For an infinite universe, \mathbf{k} is continuous, while for a compact topology only discrete values,

$$\mathbf{k} = \frac{2\pi}{L} \mathbf{n}, \quad \mathbf{n} = (n_x, n_y, n_z), \quad (3)$$

are allowed, where n_x , n_y , and n_z are integers and L is the cell size in units of the conformal time to the decoupling surface (Zel'dovich 1973; Fang & Mo 1987; Starobinsky 1993; Sokolov 1993). The cutoff in the discrete spectrum, $|\mathbf{k}| > (2\pi/L)$, suppresses power on large angular scales.

In inflationary cosmologies, the gravitational potential's $\Phi_{\mathbf{k}}$ are random Gaussian variables with zero mean, and the covariance

$$\langle \Phi_{\mathbf{k}} \Phi_{\mathbf{k}'}^* \rangle = \frac{2\pi^2}{k^3} \delta^3(\mathbf{k} - \mathbf{k}') \mathcal{P}(k). \quad (4)$$

We expand the corresponding temperature fluctuations,

$$\Delta T(\hat{\mathbf{x}}) = \sum_{\ell m} a_{\ell m} Y_{\ell m}(\hat{\mathbf{x}}), \quad (5)$$

where

$$a_{\ell m} = -\frac{(-i)^\ell 4\pi}{3} \int d\mu(\mathbf{k}) \Phi_{\mathbf{k}} j_\ell(k\Delta\tau) Y_{\ell m}(\hat{\mathbf{k}}) \quad (6)$$

and $j_\ell(k\Delta\tau)$ is a Bessel function of order ℓ . The $a_{\ell m}$ values have zero mean. Using equation (4), we find their covariance,

$$\mathbf{M}_{\ell m, \ell' m'}^L = \langle a_{\ell m} a_{\ell' m'}^* \rangle \quad (7)$$

$$\begin{aligned} &= \frac{(-i)^\ell i^{\ell'} 32\pi^4}{9} \int d\mu(\mathbf{k}) j_\ell(k\Delta\tau) j_{\ell'}(k\Delta\tau) \\ &\quad \times \frac{\mathcal{P}(k)}{k^3} Y_{\ell m}(\hat{\mathbf{k}}) Y_{\ell' m'}^*(\hat{\mathbf{k}}), \end{aligned} \quad (8)$$

which is not necessarily diagonal.

In the limit of a flat open universe, $d\mu(\mathbf{k}) \rightarrow k^2 dk d\hat{\mathbf{k}}$, and the orthonormality of the spherical harmonics yields

$$\lim_{L \rightarrow \infty} \mathbf{M}_{\ell m, \ell' m'}^L = \frac{32\pi^4}{9} \delta_{\ell\ell'} \delta_{mm'} \int_0^\infty dk j_\ell(k\Delta\tau)^2 \frac{\mathcal{P}(k)}{k}. \quad (9)$$

For a compact topology, the integral over the continuous variables \mathbf{k} becomes a sum over discrete $\mathbf{k}_n = (2\pi/L)\mathbf{n}$, with $\mathbf{n} = (n_x, n_y, n_z)$ a triplet of integers. Just as we break the continuous integration $d\mu(\mathbf{k})$ into magnitude and angular parts, we do the same for the discrete case, $d\mu(\mathbf{k}) = \sum_{\mathbf{n} \in \mathcal{N}} \sum_{\{|\mathbf{n}|=n\}}$, where \mathcal{N} is the set of all possible magnitudes of the integer triplets \mathbf{n} and $\{|\mathbf{n}|=n\}$ are all \mathbf{n} with magnitude n . Thus $\sum_{\mathbf{n} \in \mathcal{N}}$ is the sum over magnitude, and $\sum_{\{|\mathbf{n}|=n\}}$ is, for each magnitude, the angular sum. For example, the first value in \mathcal{N} is $n = 1$, with the corresponding $\{|\mathbf{n}|=1\}$ containing the six vectors $(\pm 1, 0, 0)$, $(0, \pm 1, 0)$, and $(0, 0, \pm 1)$. The next value in \mathcal{N} is $n = \sqrt{2}$, and $\{|\mathbf{n}|=\sqrt{2}\}$ contains 12 vectors, each with ± 1 in two places and 0 in the third. Writing the measure $d\mu(\mathbf{k})$ in this form, for compact topologies we have

$$\mathbf{M}_{\ell m, \ell' m'}^L = \frac{32\pi^4}{9} \sum_{\mathbf{n} \in \mathcal{N}} j_\ell(k_n \Delta\tau) j_{\ell'}(k_n \Delta\tau) \frac{\mathcal{P}(k)}{k^3} A_{\ell m, \ell' m'}^L(\mathbf{n}), \quad (10)$$

where

$$A_{\ell m, \ell' m'}^L(\mathbf{n}) = (-i)^\ell i^{\ell'} \sum_{\{|\mathbf{n}|=n\}} Y_{\ell m}(\hat{\mathbf{n}}) Y_{\ell' m'}^*(\hat{\mathbf{n}}). \quad (11)$$

The momentum k_n is related to n via $k_n = (2\pi/L)|\mathbf{n}|$.

There is a great computational advantage to making this split in the sum. All the cosmology is contained in the magnitude sum, while the topology is reflected in the sum for $A_{\ell m, \ell' m'}^L(\mathbf{n})$, which requires the lion's share of CPU time. When numerically evaluating equation (13), we compute and store all the required matrices $A_{\ell m, \ell' m'}^L(\mathbf{n})$, which only depend on the relative ratios of the fundamental domain sizes L_x , L_y , and L_z . For this work, we have assumed they are all the same and equal to L . As we vary the rest of the model parameters, including the topology scale L , we need only evaluate the factors in the first sum and use the stored matrices $A_{\ell m, \ell' m'}^L(\mathbf{n})$ to complete the sum.

Thus far, we have ignored contributions from the time evolution of the gravitational potential (the integrated Sachs-Wolfe [ISW] effect) on large scales and from the Boltzmann physics of the coupled photon-baryon fluid on small scales. Both effects may readily be included. The ‘‘line-of-sight’’ approach (Seljak & Zaldarriaga 1996) to computing the anisotropies for a flat open universe first calculates the (scalar) response function $\Delta_{T\ell}^{(S)}(k, \Delta\tau)$ and then gives the (diagonal) covariance,

$$\mathbf{M}_{\ell m, \ell' m'}^{\text{CMBFAST}} = (4\pi)^2 \delta_{\ell\ell'} \delta_{mm'} \int_0^\infty dk |\Delta_{T\ell}^{(S)}(k, \Delta\tau)|^2 \frac{\mathcal{P}(k)}{k}. \quad (12)$$

This suggests that we can include the ISW and acoustic effects of the plasma motion by writing covariance (8) as

$$\begin{aligned} \mathbf{M}_{\ell m, \ell' m'}^L &= (4\pi)^2 \sum_{\mathbf{n} \in \mathcal{N}} \Delta_{T\ell}^{(S)}(k_n, \Delta\tau) \Delta_{T\ell'}^{(S)}(k_n, \Delta\tau) \\ &\quad \times \frac{\mathcal{P}(k)}{k^3} A_{\ell m, \ell' m'}^L(\mathbf{n}), \end{aligned} \quad (13)$$

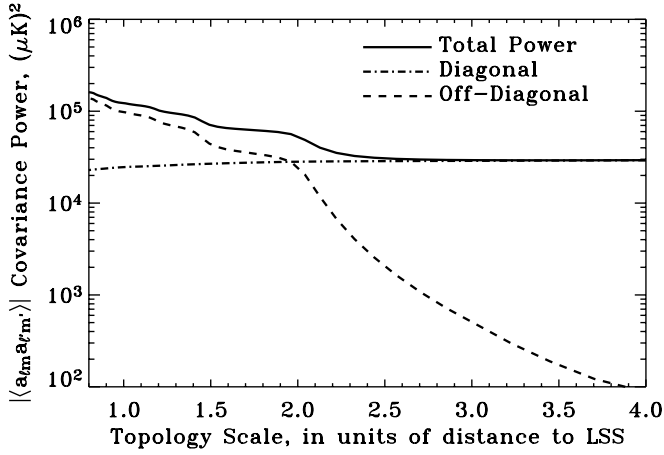


FIG. 2.—Comparison of diagonal vs. off-diagonal power in the spherical expansion coefficient correlation matrix, $|\mathbf{M}_{\ell m, \ell' m'}^L| = |\langle a_{\ell m} a_{\ell' m'}^* \rangle|$, as a function of topology scale L . The power measures the relative importance of terms. The total power in the diagonal components, $\sum_{\ell m} |\mathbf{M}_{\ell m, \ell m}^L|$, measures the information considered by the power spectrum and is flat with only a small decrease at small L due to suppression of power at the largest scales for small topologies. The off-diagonal power, $\sum_{(\ell m) \neq (\ell' m')} |\mathbf{M}_{\ell m, \ell' m'}^L|$, arises from the global structure present for finite topologies. This varies from almost negligible at the largest topology scales to being the dominant contribution for small topologies. An analysis looking for topology via a spherical expansion must consider the correlations between different expansion coefficients.

with $A_{\ell m, \ell' m'}^L(n)$ given by equation (11). We obtain $\Delta_{T\ell}^{(S)}(k, \Delta\tau)$ from CMBFAST (Seljak & Zaldarriaga 1996), by way of the program CMBEASY (Doran 2003), along with the associated value of $\Delta\tau$.

Equation (13) gives the covariance of the temperature coefficients $a_{\ell m}$ as a function of cosmological parameters and global topology. Figure 2 compares the diagonal elements to the off-diagonal elements for the case of a cubic topology. The total power on the diagonal, $\sum_{\ell m} |\mathbf{M}_{\ell m, \ell m}^L|$, reflects the content of power spectrum C_ℓ^L and is rotationally invariant. The off-diagonal power, $\sum_{(\ell m) \neq (\ell' m')} |\mathbf{M}_{\ell m, \ell' m'}^L|$, measures the correlations between the different angular scales and is generated by the global properties of the topology. For cell size L less than twice the distance to the decoupling surface, the off-diagonal elements dominate. Analyses based solely on the power spectrum thus ignore the main information content of the map. The off-diagonal correlations persist at larger cell size but become increasingly less important. In contrast to the circles-on-the-sky test, the off-diagonal correlations smoothly decrease past $L > 2$.

2.1. Sky Map Generation

The covariance matrix $\mathbf{M}_{\ell m, \ell' m'}^L$ for the $a_{\ell m}$ coefficients fully describes a cosmological model and topology. Since the $a_{\ell m}$ values are still Gaussian variables, albeit now correlated, all higher moments are given in terms of their covariance. Equation (11) thus allows the rapid generation of simulated sky maps for a given topology. For a given topology scale L and cosmology $\Delta_{T\ell}^{(S)}(k)$, we first compute the Cholesky decomposition $\mathbf{L}_{\ell m, \ell' m'}$ of the $a_{\ell m}$ covariance matrix, $\mathbf{M}_{\ell m, \ell' m'} = \mathbf{L}_{\ell m, \ell' m'} \mathbf{L}_{\ell' m', \ell m}$. Then if $x_{\ell m}$ are a set of uncorrelated unit variance Gaussian variables, we set $a_{\ell m} = \sum_{\ell' m'} \mathbf{L}_{\ell m, \ell' m'} x_{\ell' m'}$ to obtain the sky map $\Delta T(\hat{x}) = \sum_{\ell m} a_{\ell m} Y_{\ell m}(\hat{x})$. A circles-in-the-sky test (Cornish et al. 1998) run on the resulting sky maps verifies that the algorithm correctly reproduces all features for compact topologies.

The only remaining free parameter in our model is the overall amplitude of the fluctuations. We fix this by comparing the power spectrum for the compact model to that for an infinite

universe with the same response function $\Delta_{T\ell}^{(S)}(k)$. The CMB power spectrum is given by

$$C_\ell = \frac{1}{2\ell + 1} \sum_m \langle |a_{\ell m}|^2 \rangle = \frac{1}{2\ell + 1} \sum_m M_{\ell m, \ell m}^L. \quad (14)$$

The covariance of the power spectrum becomes

$$\mathbf{M}_{\ell\ell'}^{L, C_\ell} = \frac{2}{(2\ell + 1)(2\ell' + 1)} \sum_{m=-\ell}^{\ell} \sum_{m'=-\ell'}^{\ell'} |\mathbf{M}_{\ell m, \ell' m'}^L|^2. \quad (15)$$

We normalize the covariance matrices $\mathbf{M}_{\ell m, \ell' m'}^L$ using the amplitude that minimizes $\sum_{\ell_{\text{lower}}}^{\ell_{\text{max}}} (C_\ell^L - C_\ell^{\text{CMBFAST}})^2$, where we use $\ell_{\text{lower}} = 20$, as this is a small enough scale to be beyond where we expect to find interesting topological effects. The simulated maps correctly show the rise to the first Doppler peak. Since topological effects are most apparent on large scales, we take $\ell_{\text{max}} = 30$ for all analyses in this paper.

3. LIKELIHOOD ANALYSIS

We compare data from the *WMAP* first-year sky maps to a set of models with a cubic fundamental domain described by the cell size L in units of the conformal time to the decoupling surface. We specialize to the case of a cubic topology, identifying opposing faces of the unit cell without twists or rotation, so that the topology is fully specified by the cell size L . For each value L we compute the likelihood $\log \mathcal{L} = -\frac{1}{2}(\chi^2 + \log \det \mathbf{M})$, where

$$\chi^2 = \sum_{\ell\ell', mm'} a_{\ell m} (\mathbf{M}^{-1})_{\ell m, \ell' m'} a_{\ell' m'} \quad (16)$$

and $\mathbf{M}_{\ell m, \ell' m'}$ is given by equation (13).

We use temperature coefficients $a_{\ell m}$ derived from the internal linear combination (ILC) map from the *WMAP* first-year data release (Bennett et al. 2003b). This map reduces foreground emission at the cost of a complicated window function and instrument noise on angular scales $\theta < 2^\circ$. Topology is important only at much larger angular scales. We limit the likelihood calculation to $2 < \ell < 30$. On these scales, the dominant uncertainty is cosmic variance; effects from the instrument noise and beam profiles are negligible.

The ILC map reduces foreground emission but does not eliminate it completely. We impose a cut in Galactic latitude $|b| > 5^\circ$ and compute the $a_{\ell m}$ using unit weight for all pixels outside the cut. Our results are stable, as the cut is varied from 2.5 to 15° . This stability arises because we apply the same cut to the models as we apply to the data. This is done before the data are compared to the model via the likelihood \mathcal{L} . Thus the aliasing of power that occurs due to the cut happens equally for both the data and the model. The variations in likelihood as the topology scale changes will reflect the relative probabilities of various topology scales, above and beyond the effect of the Galactic cut.

The likelihood analysis includes off-diagonal correlations between different $a_{\ell m}$ and is not rotationally invariant. Equation (13) was derived for the case when the faces of the fundamental domain align with the data coordinate system. We must thus consider different possible orientations between the data and the unit cell of the topology. It does not matter whether we rotate the model or the data, so we rotate the data. In terms of the Euler angles $\boldsymbol{\xi} = (\phi, \theta, \psi)$, we use the rotation matrix

$$\mathbf{R} = \mathbf{R}_{\ell m, \ell' m'}(\boldsymbol{\xi}) = \delta_{\ell\ell'} e^{-i\phi m'} d_{mm'}^\ell(\theta) e^{-i\psi m}, \quad (17)$$

where $d_{mm'}^\ell(\theta)$ are the Wigner rotation functions. We apply the rotation on a grid of 1665 possible orientations ξ over the range $0 \leq \phi, \theta, \psi \leq \pi/2$; ϕ and θ are uniformly distributed on the sphere, while the azimuthal rotation ψ is uniformly distributed over its range. Since the fundamental domain for the topology is cubic, the noted range covers all possible orientations.

For a fixed cosmological model, the likelihood becomes a function of four parameters,

$$\log \mathcal{L}(\mathbf{a}|L, \xi) = -\frac{1}{2} [(\mathbf{R}\mathbf{a})^T (\mathbf{M})^{-1} (\mathbf{R}\mathbf{a}) + \log \det \mathbf{M}], \quad (18)$$

with one parameter specifying the size of the unit cell and three parameters specifying the orientation, where for clarity we have suppressed the ℓm subscripts. Even in the absence of topological effects, chance alignments between the random $a_{\ell m}$ can create nonzero correlations in a single realization. If the relative orientation of the data and model is held constant, the likelihood will approximate a normal distribution over an ensemble of CMB maps. Selecting the maximum likelihood over many possible orientations of a *single* CMB map, however, selects the lowest χ^2 at each cell size L . If the data do not represent a model with compact topology, maximizing the likelihood over the (now nuisance) rotational parameters ξ will select from the tail of the χ^2 distribution (eq. [16]), leading to a biased estimate for L . The bias is unimportant for large L where the scale is too large for the topology to break global isotropy, but it becomes significant for $L \sim 2$ where the off-diagonal terms begin to dominate (Fig. 2).

We quantify the effect of maximizing the likelihood over orientation using Monte Carlo simulations. We generate 1000 realizations of compact topologies at 51 different values of L on a grid from $L = 0.8\Delta\tau$ to $4.0\Delta\tau$ uniformly spaced in $1/L$. For simplicity, we work directly with the $a_{\ell' m'}$ coefficients to avoid the intermediate steps of generating sky maps, masking pixels near the Galactic plane, and computing the $a_{\ell' m'}$ from the unmasked pixels. We instead impose the Galactic cut using the matrix

$$\begin{aligned} \mathbf{P} &= \mathbf{P}_{\ell m, \ell' m'}(b) \\ &= \left[1 + (-1)^{\ell+\ell'+m+m'} \right] \int_{\cos(\pi/2-b)}^1 P_\ell^m(y) P_{\ell'}^{m'}(y) dy, \quad (19) \end{aligned}$$

which removes all power from the azimuthally symmetric region $|\theta - \pi/2| \leq b$, where $P_\ell^m(y)$ are the associated Legendre functions. Each realization thus generates a set of correlated $a_{\ell' m'}$,

$$\mathbf{a}^L = \mathbf{PRL}^L \mathbf{x}, \quad (20)$$

where \mathbf{L}^L is the Cholesky decomposition of the $a_{\ell m}$ covariance matrix (eq. [13]) and \mathbf{x} is a vector of zero mean, unit variance Gaussian random variables. We select a grid uniform in $1/L$ because the topology effects the covariance matrix \mathbf{M} as $k_n = (2\pi/L)|\mathbf{n}| \propto 1/L$. There is also the advantage of placing the infinite case at a finite distance from the region of interest, $L \sim 1$. A model with $L = 3\Delta\tau$ is nearly indistinguishable from $L = \infty$. We vary the topology scale L while keeping the ‘‘background’’ cosmology fixed. Since we only consider large angular scales, the results are insensitive to the cosmological parameters. We run the k_n integration in equation (13) for $n \leq 90$, which is sufficient for convergence.

Figure 3 shows the likelihood of the *WMAP* first-year data as a function of domain size L . When maximized over orientation, the likelihood peaks sharply at $L \sim 2$. We test for the significance of this peak by comparing the *WMAP* results to 1000 Monte

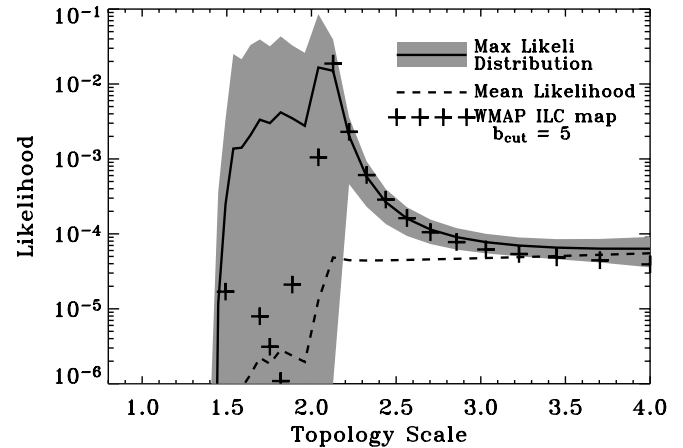


FIG. 3.—Likelihood of the *WMAP* first-year data as a function of topology scale. At each topology scale L , the likelihood is maximized over the possible orientations of the fundamental domain. For $L \sim 2$, chance alignments create a bias in the likelihood estimator. The solid line and gray band show the mean and standard deviation of 1000 simulations drawn from an infinite flat model. The *WMAP* data are consistent with the infinite model. The dashed line shows the mean likelihood, marginalized over orientation. For the largest topology scale, there is no difference between maximizing or marginalizing over orientation; the scale is too large for the topology to break the global isotropy. At smaller scales, orientation starts to matter, and although a parent population may lack global isotropy, any given realization may ‘‘appear’’ to break it. The resulting bias can be quantified using Monte Carlo simulations.

Carlo simulations of an infinite flat universe. The simulations also peak at $L = 2$, demonstrating the bias incurred when maximizing over orientation. The *WMAP* data fall near the mean of the simulations, suggesting that the data are consistent with an infinite universe. In this case, the orientation ξ becomes a nuisance parameter. We may then marginalize over the nuisance parameter by averaging the likelihood over all orientations (as opposed to selecting the best orientation). The marginalized likelihood shows a plateau for $L > 2.1$, with a sharp drop at smaller cell size. Note that for $L > 3$ the orientation ceases to be important as the maximum likelihood asymptotically approaches the marginalized likelihood.

Figure 3 shows that the *WMAP* likelihood falls near the mean of Monte Carlo simulations drawn from a parent population with an infinite topology. We use additional Monte Carlo simulations to set upper limits to the allowed size L of the fundamental domain. We repeat the maximum likelihood analysis for 19,000 Monte Carlo simulations, 1000 at each of 19 values for L ranging from 1.01 to 3.7 uniformly spaced in $1/L$. For each simulation, we get a ‘‘best-fit’’ topology scale L_{out} , the value that maximizes $\mathcal{L}(\mathbf{a}^{L_{\text{in}}}|L, \xi)$ (eq. [18]). We then tabulate the probability $\mathcal{P}(L_{\text{out}}|L_{\text{in}})$ for each input realization with the ‘‘true’’ topology scale L_{in} to produce the best-fit output scale L_{out} . We invert this relationship to derive the probability for a given best-fit output L_{out} to be drawn from a parent population with topology scale L_{in} ,

$$\mathcal{P}(L_{\text{in}}|L_{\text{out}}) = \frac{\mathcal{P}(L_{\text{out}}|L_{\text{in}})}{\int_0^\infty \mathcal{P}(L_{\text{out}}|L_{\text{in}}) d(L_{\text{in}})}, \quad (21)$$

where the factor $d(L_{\text{in}})$ explicitly accounts for the simulations’ uniform distribution in $1/L$.

Figure 4 shows both distributions. For $L_{\text{in}} < 2$, $L_{\text{out}} \simeq L_{\text{in}}$ and our likelihood analysis successfully identifies the topology scale. On these scales, the off-diagonal components of \mathbf{M}_L become dominant (Fig. 2); the correlations between the different $a_{\ell m}$ values are important, and the likelihood function strongly discriminates

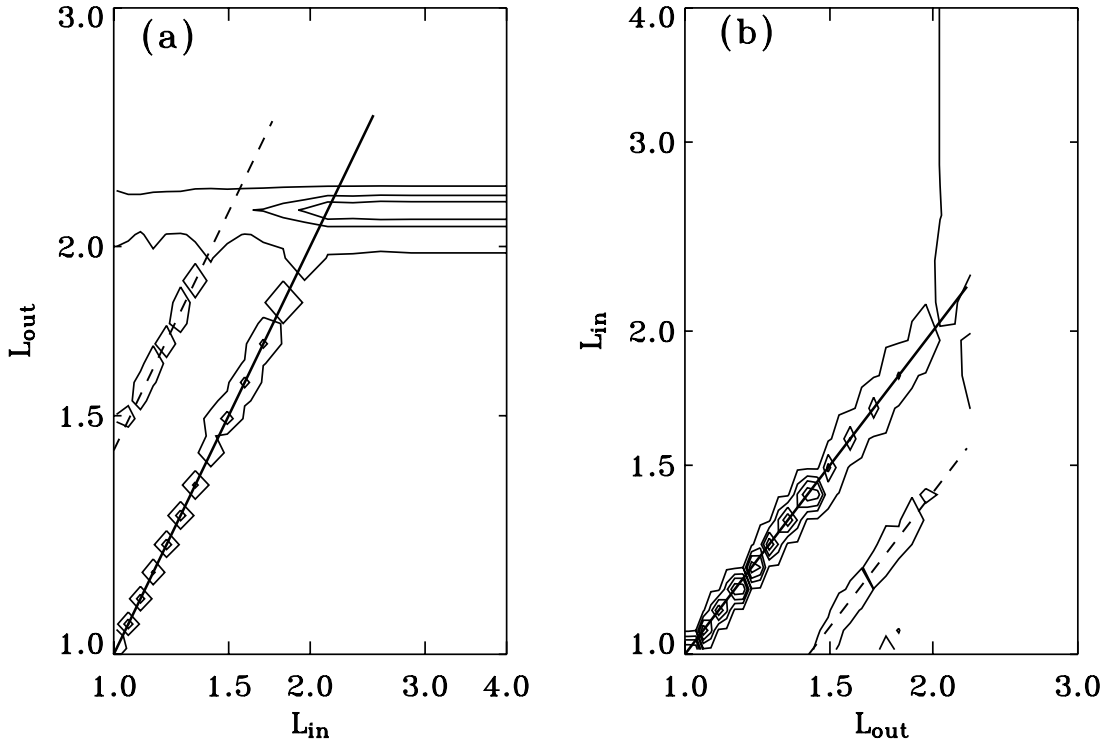


FIG. 4.—Probability distributions from Monte Carlo simulations. (a) Probability to obtain maximum likelihood output L_{out} as a function of the input scale size L_{in} after maximizing over orientation. The solid line shows $L_{\text{out}} = L_{\text{in}}$. For $L_{\text{in}} < 2$, the fundamental domain fits inside the decoupling surface, and the distributions are strongly peaked at $L_{\text{out}} \simeq L_{\text{in}}$. For $L_{\text{in}} > 2$, maximizing over orientation causes all the simulations to fall into a group centered around $L_{\text{out}} \simeq 2.2$. (b) By inverting the relationship in (a), we obtain the probability that an observed output L_{out} was drawn from a parent population with L_{in} . Values $L_{\text{out}} > 1.96$ are indistinguishable from infinite models.

models. A minor identification $L_{\text{out}} \simeq \sqrt{2}L_{\text{in}}$ is also apparent. This aliasing of scale is typical and in this case corresponds to the ratio in sizes between a circle that just circumscribes a square and the circle that is just contained in a square. Such aliasing of topology scales is expected and is also seen during a circles-in-the-sky analysis.

For $L_{\text{in}} > 2$, maximizing over orientation produces output $L_{\text{out}} \simeq 2.2$ independent of the actual value of $L_{\text{in}} > 2$. The correlations between the $a_{\ell m}$ induced by topology are weak on scales $L > 2$ compared to chance alignments. Individual realizations with $L > 2$ will thus appear to have a slight preference

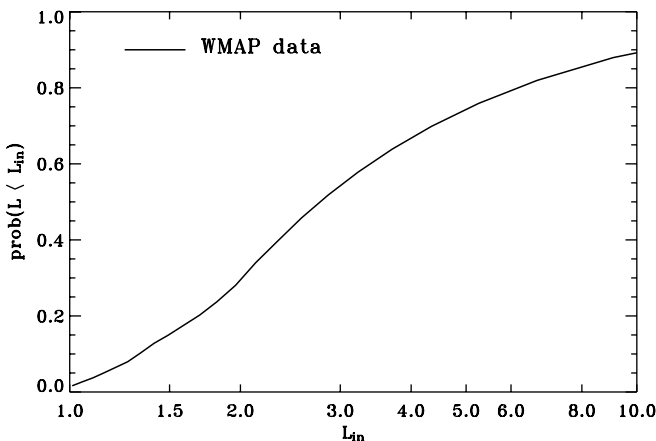


FIG. 5.—Cumulative probability that the true topology scale is below a given value for *WMAP* data with Galaxy cut $|b| > 5^\circ$. We obtain limits $L > 2.1$ for the distance to the decoupling surface at 68% confidence, and $L > 1.2$ at 95% confidence. The *WMAP* data are consistent, with a lack of any finite topology for the cubic flat topologies we consider.

for orientation, although the parent population is nearly indistinguishable from the infinite model. We account for this bias using equation (21). An observed best-fit value $L_{\text{out}} \sim 2$ does *not* mean the parent population must necessarily have $L_{\text{in}} = 2$, but rather that the parent population has nearly uniform probability to represent any scale $L_{\text{in}} \geq 2$.

The maximum likelihood for the *WMAP* data occurs for $L_{\text{out}}^{\text{WMAP}} = 2.1$. From equation (21), the cumulative probability is $P(L_{\text{in}} < L_{\text{WMAP, in}}) = \int_0^{L_{\text{in}}} \mathcal{P}(L'_{\text{in}} | L_{\text{WMAP, out}}) d(L'_{\text{in}})$. We obtain the 68% confidence limit that the topology scale is greater than 2.1 times the distance to the decoupling surface, and 95% confidence that it is greater than 1.2 times the distance. We place a 68% (95%) confidence that the topology scale is greater than 29 Gpc (17 Gpc).

Figure 4a shows that a certain number of realizations with $L < 2$ “scatter” into the region $L_{\text{out}} \sim 2$ where the *WMAP* data show maximum likelihood. We test the null hypothesis (that the universe is infinite) by taking all simulations with $L_{\text{out}} > 1.96$ (the horizontal band in Fig. 4a) and computing the cumulative probability for these simulations only. Figure 5 shows the resulting curve. The cumulative probability for the *WMAP* data, computed using all simulations, has identical confidence intervals, as the probability derived using only those simulations whose likelihood peak occurs at $L_{\text{out}} > 1.96$. We therefore accept the null hypothesis to conclude that the *WMAP* data are consistent with an infinite universe.

4. DISCUSSION

A compact topology imposes a specific pattern of correlations $\langle a_{\ell m} a_{\ell' m'} \rangle$ between the spherical harmonic expansion of the CMB temperature. We compute the expected correlations for the simplest nontrivial topology, the cubic torus, and compare a

range of cell size L to the *WMAP* first-year data using a maximum likelihood algorithm. The covariance matrix explicitly includes the contribution from the ISW effect on large angular scales and the acoustic peaks at small scales. We separate the covariance into a piece dependent on the topology and a piece dependent on the cosmology. Although the transfer functions $\Delta_{T\ell}^{(S)}(k)$ for the cosmology assume isotropy in k -space, which is no longer exact for compact models, the effect is predominantly in the cosmology with negligible effect on the topology.

The algorithm is sensitive both to the power spectrum of the data (diagonal elements of the covariance matrix for different $a_{\ell m}$) and the phase information contained in the off-diagonal elements. For cell size $L < 2$ the off-diagonal elements are larger than the diagonal elements. A comparison of the data to topological models that uses only the power spectrum can produce false positives by ignoring the additional information in the off-diagonal elements. We demonstrate this using Monte Carlo simulations. The power spectrum is rotationally invariant and does not specify orientation. We may thus modify equation (16) to use the power spectrum C_ℓ and its covariance (eqs. [14] and [15]) in place of the spherical harmonic coefficients $a_{\ell m}$. When only considering the power spectrum, the maximum likelihood for the *WMAP* data occurs at $L = 1.1\Delta\tau$; this is the “finite” model power spectrum plotted in Figure 1. Does this imply a positive detection of finite topology? To test this, we generate 1000 Monte Carlo realizations drawn from a parent population with $L = 1.1$ and generate the likelihood for each realization using the full covariance matrix (eqs. [13] and [19]). For such a small topology scale, almost all realizations have their likelihood peak at $L = 1.1$. This scale is small enough so that the bias from maximizing over orientation is not important. For each realization, we also generate a “companion” realization with exactly the same

power spectrum, but with completely uncorrelated $a_{\ell m}$ values. The two realizations by construction must give the same results for an analysis based solely on the power spectrum. When we analyze the companion realizations using the full $a_{\ell m}$ covariance matrix, we obtain results similar to the infinite models displayed in Figure 3. A likelihood analysis using the full $a_{\ell m}$ covariance matrix successfully distinguishes models with compact topology from models with identical power spectra but without the correlations between different $a_{\ell m}$ required by the topology. Suppression of power in the quadrupole and octopole moments is a necessary but not sufficient condition for a compact topology. The *WMAP* data show reduced power at $\ell = 2$ and 3 but do not show the correlations expected for a compact topology and are indistinguishable from infinite models.

For cell sizes comparable to the distance to the decoupling surface, the correlations become weaker. Maximizing the likelihood over orientation then allows chance alignments to introduce a bias in the likelihood estimator. We quantify this using Monte Carlo simulations. The *WMAP* first-year data are consistent with input models drawn from parent populations with infinite fundamental domain. We establish 95% confidence limit $L > 17$ Gpc for the cell size of a cubic topology, in agreement with the result of 24 Gpc obtained by Cornish et al. (2004).

We thank G. Hinshaw for helpful discussions. N. G. P. thanks Joseph Silk for first suggesting this topic of investigation, along with many fruitful discussions. N. G. P. also thanks M. Doran for his help in using CMBEASY. This work was supported by the National Aeronautics and Space Administration under the Astrophysics Data program of the Office of Space Science.

REFERENCES

- Bennett, C. L., et al. 1996, *ApJ*, 464, L1
 ———. 2003a, *ApJS*, 148, 1
 ———. 2003b, *ApJS*, 148, 97
 Cornish, N. J., Spergel, D. N., & Starkman, G. D. 1998, *Classical Quantum Gravity*, 15, 2657
 Cornish, N. J., Spergel, D. N., Starkman, G. D., & Komatsu, E. 2004, *Phys. Rev. Lett.*, 92, 201302
 de Oliveira-Costa, A., & Smoot, G. F. 1995, *ApJ*, 448, 477
 de Sitter, W. 1917, *Proc. R. Acad. Amsterdam*, 20, 29
 Doran, M. 2003, preprint (astro-ph/0302138)
 Fang, L. Z., & Mo, H. J. 1987, in *IAU Symp. 124, Observational Cosmology*, ed. A. Hewitt, G. Burbidge, & L. Z. Fang (Dordrecht: Reidel), 461
 Inoue, K. T., & Sugiyama, N. 2003, *Phys. Rev. D*, 67, 043003
 Luminet, J.-P., Weeks, J., Riazuelo, A., Lehoucq, R., & Uzan, J.-P. 2003, *Nature*, 425, 593
 Schwarzschild, K. 1900, *Vierteljahrsschrift d. Astronom. Gesellschaft*, 35, 1900 (English transl. in *Classical Quantum Gravity*, 15, 2539 [1998])
 Seljak, U., & Zaldarriaga, M. 1996, *ApJ*, 469, 437
 Sokolov, I. Y. 1993, *Soviet Phys.—JETP Lett.*, 57, 617
 Spergel, D. N., et al. 2003, *ApJS*, 148, 175
 Starobinsky, A. A. 1993, *Soviet Phys.—JETP Lett.*, 57, 622
 Zel’dovich, Y. B. 1973, *Comments Astrophys. Space Phys.*, 5, 169