The Astrophysical Journal Supplement Series, 148:195–211, 2003 September

FIRST-YEAR *WILKINSON MICROWAVE ANISOTROPY PROBE* (*WMAP*)[1] OBSERVATIONS: PARAMETER ESTIMATION METHODOLOGY

L. Verde,[2,3] H. V. Peiris,[2] D. N. Spergel,[2] M. R. Nolta,[4] C. L. Bennett,[5] M. Halpern,[6] G. Hinshaw,[5] N. Jarosik,[4] A. Kogut,[5] M. Limon,[5,7] S. S. Meyer,[8] L. Page,[4] G. S. Tucker,[5,7,9] E. Wollack,[5] and E. L. Wright[10]

Received 2003 February 11; accepted 2003 May 30

ABSTRACT

We describe our methodology for comparing the *Wilkinson Microwave Anisotropy Probe* (*WMAP*) measurements of the cosmic microwave background (CMB) and other complementary data sets to theoretical models. The unprecedented quality of the *WMAP* data and the tight constraints on cosmological parameters that are derived require a rigorous analysis so that the approximations made in the modeling do not lead to significant biases. We describe our use of the likelihood function to characterize the statistical properties of the microwave background sky. We outline the use of the Monte Carlo Markov Chains to explore the likelihood of the data given a model to determine the best-fit cosmological parameters and their uncertainties. We add to the *WMAP* data the $\ell \gtrsim 700$ Cosmic Background Imager (CBI) and Arcminute Cosmology Bolometer Array Receiver (ACBAR) measurements of the CMB, the galaxy power spectrum at $z \sim 0$ obtained from the Two-Degree Field Galaxy Redshift Survey (2dFGRS), and the matter power spectrum at $z \sim 3$ as measured with the Lyα forest. These last two data sets complement the CMB measurements by probing the matter power spectrum of the nearby universe. Combining CMB and 2dFGRS requires that we include in our analysis a model for galaxy bias, redshift distortions, and the nonlinear growth of structure. We show how the statistical and systematic uncertainties in the model and the data are propagated through the full analysis.

Subject headings: cosmic microwave background — cosmological parameters — cosmology: observations — methods: data analysis — methods: statistical

1. INTRODUCTION

Cosmic microwave background (CMB) experiments are powerful cosmological probes because the early universe is particularly simple and because the fluctuations over angular scales $\theta > 0.^\circ 2$ are described by linear theory (Peebles & Yu 1970; Bond & Efstathiou 1984; Zaldarriaga & Seljak 2000). Exploiting this simplicity to obtain precise constraints on cosmological parameters requires that we accurately characterize the performance of the instrument (Jarosik et al. 2003b; Page et al. 2003b; Barnes et al. 2003; Hinshaw et al. 2003a), the properties of the foregrounds (Bennett et al. 2003a), and the statistical properties of the microwave sky.

The primary goal of this paper is to present our approach to extracting the cosmological parameters from the temperature-temperature angular power spectrum (TT) and the temperature-polarization angular cross-power spectrum (TE). In companion papers, we present the TT (Hinshaw et al. 2003a) and TE (Kogut et al. 2003) angular power spectra and show that the CMB fluctuations may be treated as Gaussian (Komatsu et al. 2003).

Our basic approach is to constrain cosmological parameters with a likelihood analysis first of the *Wilkinson Microwave Anisotropy Probe* (*WMAP*) TT and TE spectra alone, then jointly with other CMB angular power spectrum determinations at higher angular resolution, and finally of all CMB power spectra data jointly with the power spectrum of the large-scale structure (LSS). In § 2 we describe the use of the likelihood function for the analysis of microwave background data. This builds on the Hinshaw et al. (2003b) methodology for determining the TT spectrum and its curvature matrix and Kogut et al. (2003), who describe our methodology for determining the TE spectrum. In § 3 we describe our use of Markov Chain Monte Carlo (MCMC) techniques to evaluate the likelihood function of model parameters. While *WMAP*'s measurements are a powerful probe of cosmology, we can significantly enhance their scientific value by combining the *WMAP* data with other astronomical data sets. This paper also presents our approach for including external CMB data sets (§ 4), LSS data (§ 5), and Lyα forest data (§ 6). When including external data sets, the reader should keep in mind that the physics and the instrumental effects involved in the interpretation of these external data sets (especially 2dFGRS and Lyα) are much more complicated and less well understood than for WMAP data. Nevertheless, we aim to match the rigorous treatment of uncertainties in the *WMAP* angular power spectrum with the inclusion of known statistical and systematic effects (of the data and of the theory), in the complementary data sets.

[1] *WMAP* is the result of a partnership between Princeton University and the NASA Goddard Space Flight Center. Scientific guidance is provided by the *WMAP* Science Team.

[2] Department of Astrophysical Sciences, Princeton University, Peyton Hall, Princeton, NJ 08544; lverde@astro.princeton.edu.

[3] *Chandra* Fellow.

[4] Department of Physics, Princeton University, Jadwin Hall, P.O. Box 708, Princeton, NJ 08544.

[5] NASA Goddard Space Flight Center, Code 685, Greenbelt, MD 20771.

[6] Department of Physics and Astronomy, University of British Columbia, 6224 Agricultural Road, Vancouver, BC V6T 1Z1, Canada.

[7] National Research Council (NRC) Fellow.

[8] Departments of Astrophysics and Physics, EFI, and CfCP, University of Chicago, 5640 South Ellis Avenue, Chicago, IL 60637.

[9] Department of Physics, Brown University, Providence, RI 02912.

[10] Department of Astronomy, UCLA, P.O. Box 951562, Los Angeles, CA 90095-1562.

2. LIKELIHOOD ANALYSIS OF *WMAP* ANGULAR POWER SPECTRA

The first goal of our analysis program is to determine the values and confidence levels of the cosmological parameters that best describe the *WMAP* data for a given cosmological model. We also wish to discriminate between different classes of cosmological models, in other words, to assess whether a cosmological model is an acceptable fit to *WMAP* data.

The ultimate goal of the likelihood analysis is to find a set of parameters that give an estimate of $\langle \mathscr{C}_\ell \rangle$, the ensemble average of which the realization on our sky[11] is $\mathscr{C}_\ell^{\rm sky}$. The likelihood function, $\mathscr{L}[\hat{\mathscr{C}}_\ell | \mathscr{C}_\ell^{\rm th}(\boldsymbol{a})]$, yields the probability of the data given a model and its parameters ($\boldsymbol{a}$). In our notation $\hat{\mathscr{C}}_\ell$ denotes our best estimator of $\mathscr{C}_\ell^{\rm sky}$ (Hinshaw et al. 2003a) and $\mathscr{C}_\ell^{\rm th}$ is the theoretical prediction for angular power spectrum. From Bayes's theorem, we can split the expression for the probability of a model given the data as

$$\mathscr{P}\left(\alpha | \hat{\mathscr{C}}_\ell\right) = \mathscr{L}\left[\hat{\mathscr{C}} | \mathscr{C}_\ell^{\rm th}(\alpha)_\ell\right] \mathscr{P}(\alpha) \, , \tag{1}$$

where $\mathscr{P}(\boldsymbol{a})$ describes our *priors* on cosmological parameters and we have neglected a normalization factor that does not depend on the parameters . Once the choice of the priors is specified, our estimator of $\langle \mathscr{C}_\ell \rangle$ is given by $\mathscr{C}_\ell^{\rm th}$ evaluated at the maximum of $\mathscr{P}(\boldsymbol{a} | \hat{\mathscr{C}}_\ell)$.

2.1. *Likelihood Function*

One of the generic predictions of inflationary models is that fluctuations in the gravitational potential have Gaussian random phases. Since the physics that governs the evolution of the temperature and metric fluctuations is linear, the temperature fluctuations are also Gaussian. If we ignore the effects of nonlinear physics at $z < 10$ and the effect of foregrounds, then all of the cosmological information in the microwave sky is encoded in the temperature and polarization power spectra. The leading-order low-redshift astrophysical effect is expected to be gravitational lensing of the CMB by foreground structures. We ignore this effect here as it generates a less than 1% covariance in the TT angular power spectrum on *WMAP* angular scales (Hu 2001; see also Spergel et al. 2003, § 3).

There are several expected sources of noncosmological signal and of non-Gaussianity in the microwave sky. The most significant sources on the full sky are Galactic foreground emission, radio sources, and galaxy clusters. Bennett et al. (2003b) show that these contributions are greatly reduced *if* we restrict our analysis to a cut sky that masks bright sources and regions of bright Galactic emission. The residual contribution of these foregrounds is further reduced by the use of external templates to subtract foreground emission from the Q-, V-, and W-band maps. Komatsu et al. (2003) find no evidence for deviations from Gaussianity on this template-cleaned cut sky. While the sky cut greatly reduces foreground emission, it has the unfortunate effect of coupling multipole modes on the sky so that the power spectrum covariance matrix is no longer diagonal. The goal of this section is to include this covariance in the likelihood function.

[11] Throughout this paper we use the convention that $\mathscr{C}_\ell = \ell(\ell+1)C_\ell/(2\pi)$.

The likelihood function for the temperature fluctuations observed by a noiseless experiment with full-sky coverage has the form

$$\mathscr{L}\left(\boldsymbol{T} | \mathscr{C}_l^{\rm th}\right) \propto \frac{\exp[-(\boldsymbol{T}\boldsymbol{S}^{-1}\boldsymbol{T})/2]}{\sqrt{\det \boldsymbol{S}}} \, , \tag{2}$$

where $\boldsymbol{T}$ denotes our temperature map and $S_{ij} = \sum_\ell (2\ell+1) C_\ell^{\rm th} P_\ell(\hat{\boldsymbol{n}}_i \hat{\boldsymbol{n}}_j)/(4\pi)$, where the P_ℓ are the Legendre polynomials and $\hat{\boldsymbol{n}}_i$ is the pixel position on the map. If we expand the temperature map in spherical harmonics, $T(\hat{\boldsymbol{n}}) = \sum_{\ell m} a_{\ell m} Y_{\ell m}$, then the likelihood function for each $a_{\ell m}$ has a simple form:

$$\mathscr{L}\left(\boldsymbol{T} | C_\ell^{\rm th}\right) \propto \prod_{\ell m} \frac{\exp[-|a_{\ell m}|^2/(2C_\ell^{\rm th})]}{\sqrt{C_\ell^{\rm th}}} . \tag{3}$$

Since we assume that the universe is isotropic, the likelihood function is independent of m. Thus, we can sum over m and rewrite the likelihood function as

$$-2 \ln \mathscr{L} = \sum_\ell (2\ell+1) \left[\ln\left(\frac{\mathscr{C}_\ell^{\rm th}}{\hat{\mathscr{C}}_\ell}\right) + \frac{\hat{\mathscr{C}}_\ell}{\mathscr{C}_\ell^{\rm th}} - 1 \right] \tag{4}$$

up to an irrelevant additive constant. Here, for a full-sky, noiseless experiment, we have identified $\sum_m |a_{\ell m}|^2/(2\ell+1)$ with $\hat{\mathscr{C}}_\ell$. Note that the likelihood function depends only on the angular power spectrum. In this limit, the *angular power spectrum* encodes all of the cosmological information in the CMB.

Characteristics of the instrument are also included in the likelihood analysis. Jarosik et al. (2003a) show that the detector noise is Gaussian (see their Fig. 6 and § 3.4); consequently, the pixel noise in the sky map is also Gaussian (Hinshaw et al. 2003b). The resolution of *WMAP* is quantified with a window function, w_ℓ (Page et al. 2003a). Thus, the likelihood function for our CMB map has the same form as equation (2), but with $\boldsymbol{S}$ replaced by $\boldsymbol{C} = \hat{\boldsymbol{S}} + \boldsymbol{N}$, where $\boldsymbol{N}$ is the nearly diagonal noise correlation matrix[12] and $\hat{\boldsymbol{S}}_{ij} = \sum_\ell (2\ell+1) C_\ell^{\rm th} w_\ell P_\ell(\hat{\boldsymbol{n}}_i \cdot \hat{\boldsymbol{n}}_j)/(4\pi)$.

If foreground removal did not require a sky cut and if the noise were uniform and purely diagonal, then the likelihood function for the *WMAP* experiment would have the form (Bond, Jaffe, & Knox 2000)

$$-2 \ln \mathscr{L} = \sum_\ell (2\ell+1) \left[\ln\left(\frac{\mathscr{C}_\ell^{\rm th} + \mathscr{N}_\ell}{\tilde{\mathscr{C}}_\ell}\right) + \frac{\tilde{\mathscr{C}}_\ell}{\mathscr{C}_\ell^{\rm th} + \mathscr{N}_\ell} - 1 \right] , \tag{5}$$

where the effective bias $\mathscr{N}_\ell$ is related to the noise bias N_ℓ as

$$\mathscr{N}_\ell = N_\ell / w_\ell \ell(\ell+1)/(2\pi) \quad \text{and}$$
$$\tilde{\mathscr{C}}_\ell = \ell(\ell+1)/(2\pi) \sum_m |a_{\ell m}|^2/(2\ell+1)/w_\ell \, .$$

Note that $\mathscr{N}_\ell$ and $\mathscr{C}_\ell^{\rm th}$ appear together in equation (5) because the noise and cosmological fluctuations have the

[12] $1/f$ noise makes a non–random-phase contribution to the detector noise and leads to off-diagonal terms in the noise matrix. By making the noise N_0 a function of ℓ (denoted by N_ℓ) we include this effect to leading order (Hinshaw et al. 2003b).

same statistical properties: they both are Gaussian random fields.

Because of the foreground sky cut, different multipoles are correlated and only a fraction of the sky, $f_{\rm sky}$, is used in the analysis. In this case, it becomes computationally prohibitive to compute the exact form of the likelihood function. There are several different approximations used in the CMB literature for the likelihood function. At large ℓ, equation (5) is often approximated as Gaussian:

$$\ln \mathscr{L}_{\rm Gauss} \propto -\tfrac{1}{2} \sum_{\ell\ell'} (\mathscr{C}_\ell^{\rm th} - \hat{\mathscr{C}}_\ell) Q_{\ell\ell'} (\mathscr{C}_{\ell'}^{\rm th} - \hat{\mathscr{C}}_{\ell'}) \ , \qquad (6)$$

where $Q_{\ell\ell'}$, the curvature matrix, is the inverse of the power spectrum covariance matrix.

The power spectrum covariance encodes the uncertainties in the power spectrum due to cosmic variance, detector noise, point sources, the sky cut, and systematic errors. Hinshaw et al. (2003a) and § 2.2 describe the various terms that enter into the power spectrum covariance matrix.

Since the likelihood function for the power spectrum is slightly non-Gaussian, equation (6) is a systematically biased estimator. Bond et al. (2000) suggest using a lognormal distribution, $\mathscr{L}_{\rm LN}$ (Bond et al. 2000; Sievers et al. 2002):

$$-2 \ln \mathscr{L}_{\rm LN} = \sum_{\ell\ell'} \left(z_\ell^{\rm th} - \hat{z}_\ell\right) \mathscr{Q}_{\ell\ell'} \left(z_{\ell'}^{\rm th} - \hat{z}_{\ell'}\right) \ , \qquad (7)$$

where $z_\ell^{\rm th} = \ln(\mathscr{C}_\ell^{\rm th} + \mathscr{N}_\ell)$, $\hat{z}_\ell = \ln(\hat{\mathscr{C}}_\ell + \mathscr{N}_\ell)$, and $\mathscr{Q}_{\ell\ell'}$ is the local transformation of the curvature matrix Q to the lognormal variables z_ℓ,

$$\mathscr{Q}_{\ell\ell'} = (\hat{\mathscr{C}}_\ell + \mathscr{N}_\ell) Q_{\ell\ell'} (\hat{\mathscr{C}}_{\ell'} + \mathscr{N}_{\ell'}) \ . \qquad (8)$$

We find that, for the *WMAP* data, both equations (6) and (7) are biased estimators. We use an alternative approximation of the likelihood function for the $\mathscr{C}_\ell$ values (eq. [11]) motivated by the following argument.

We can expand the exact expression for the likelihood (eq. [4]) around its maximum by writing $\hat{\mathscr{C}}_\ell = \mathscr{C}_\ell^{\rm th}(1+\epsilon)$. Then, for a single multipole ℓ,

$$\begin{aligned} -2 \ln \mathscr{L}_\ell &= (2\ell+1)[\epsilon - \ln(1+\epsilon)] \\ &\simeq (2\ell+1)\left[\frac{\epsilon^2}{2} - \frac{\epsilon^3}{3} + \mathscr{O}\left(\epsilon^4\right)\right] . \end{aligned} \qquad (9)$$

We note that the Gaussian likelihood approximation is equivalent to the above expression truncated at ϵ^2:

$$\begin{aligned} -2 \ln \mathscr{L}_{{\rm Gauss},\ell} &\propto (2\ell+1)/2[(\hat{\mathscr{C}}_\ell - \mathscr{C}_\ell^{\rm th})/\mathscr{C}_\ell^{\rm th}]^2 \\ &\simeq (2\ell+1)\epsilon^2/2 \ . \end{aligned}$$

The Bond, Jaffe, & Knox (1998) expression for the lognormal likelihood for the equal variance approximation is

$$-2 \ln \mathscr{L}'_{\rm LN} = \frac{(2\ell+1)}{2}\left[\ln\left(\frac{\hat{\mathscr{C}}_\ell}{\mathscr{C}_\ell^{\rm th}}\right)\right]^2 \simeq (2\ell+1)\left(\frac{\epsilon^2}{2} - \frac{\epsilon^3}{2}\right) . \qquad (10)$$

Thus, our approximation of the likelihood function is given by the form

$$\ln \mathscr{L} = \tfrac{1}{3} \ln \mathscr{L}_{\rm Gauss} + \tfrac{2}{3} \ln \mathscr{L}'_{\rm LN} \ , \qquad (11)$$

where $\mathscr{L}'_{\rm LN}$ has the form of equation (7) apart from $\mathscr{Q}_{\ell\ell'}$ that is not given by equation (8) but by

$$\mathscr{Q}_{\ell\ell'} = \left(\mathscr{C}_\ell^{\rm th} + \mathscr{N}_\ell\right) Q_{\ell\ell'} \left(\mathscr{C}_{\ell'}^{\rm th} + \mathscr{N}_{\ell'}\right) . \qquad (12)$$

We tested this form of the likelihood by making 100,000 full-sky realizations of the TT angular power spectrum $\mathscr{C}_\ell^{\rm th}$. For each realization, the maximum likelihood amplitude of fluctuations in the underlying model was found and the mean value was computed. Since we kept all other model parameters fixed, this one-dimensional maximization was computationally trivial. The Gaussian approximation (eq. [6]) was found to systematically overestimate the amplitude of the fluctuations by $\simeq$0.8%, while the lognormal approximation underestimates it by $\simeq$0.2%. Equation (11) was found to be accurate to better than 0.1%.

2.2. *Curvature Matrix*

We obtain the curvature matrix in a form that can be used in the likelihood analysis from the power spectrum covariance matrix for $\hat{\mathscr{C}}_\ell$ computed in Hinshaw et al. (2003a). The matrix is composed of several terms of the following form:

$$\Sigma_{\ell\ell'} = \sqrt{D_\ell D_{\ell'}}\left(\delta_{\ell\ell'}^{\rm K} - r_{\ell\ell'}\right) + \epsilon_{\ell\ell'} \ , \qquad (13)$$

where $\epsilon_{\ell\ell'}$ is the coupling introduced by the beam uncertainties and point-source subtraction ($\epsilon_{\ell\ell'} = 0$ if $\ell = \ell'$), $\delta^{\rm K}$ denotes the Kronecker delta function, and D_ℓ denotes the diagonal terms,

$$D_\ell = 2\frac{\left(\mathscr{C}_\ell^{\rm th} + \mathscr{N}_\ell\right)^2}{(2\ell+1) f_{\rm sky}^2} \ . \qquad (14)$$

The quantity $r_{\ell\ell'}$ encodes the mode coupling due to the sky cut and is the dominant off-diagonal term (it is set to be 0 if $\ell = \ell'$). The mode-coupling coefficient, $r_{\ell\ell'}$, is most easily defined in terms of the curvature matrix, $Q_{\ell\ell'} = D_\ell^{-1}\delta_{\ell\ell'}^{\rm K} + r_{\ell\ell'}/\sqrt{D_\ell D_{\ell'}}$ (see Hinshaw et al. 2003b).[13]

The sky cut has two significant effects on the power spectrum covariance matrix. Because less data are used, the covariance matrix is increased by a factor of $f_{\rm sky}$. An additional factor of $f_{\rm sky}$ arises from the coupling to nearby ℓ-modes. The additional term does not lead to a loss of information as nearby ℓ-modes are slightly anticorrelated.

Hinshaw et al. (2003b) describe the beam uncertainty and point-source terms included in $\mathscr{N}_\ell$ and $\epsilon_{\ell\ell'}$. The beam and calibration uncertainties depend on the realization of the angular power spectrum on the sky $\mathscr{C}_\ell^{\rm sky}$, not on the theoretical angular power spectrum $\mathscr{C}_\ell^{\rm th}$; thus, they should not change as, in exploring the likelihood surface, we change $\mathscr{C}_\ell^{\rm th}$ in the expression for D_ℓ. This differs from other approaches (e.g., Bridle et al. 2002). Rescaling all the contributions to the off-diagonal terms in the covariance matrix with $\mathscr{C}_\ell^{\rm th}$ is not correct and leads to a 2% bias in our estimator of $\langle \mathscr{C}_\ell \rangle$, which propagates, for example, into a $\sim$2% error on the matter density parameter Ω_m or $\sim$2% error on the spectral slope n_s.

[13] In this equation we have set to zero the beam and point-source uncertainties. This is because the coupling coefficient is computed for an ideal cut sky.

We find the curvature matrix by inverting equation (13):

$$Q_{\ell\ell'} = D_\ell^{-1}\delta^{\rm K}_{\ell\ell'} - \frac{\epsilon_{\ell\ell'}}{D_\ell D_{\ell'}} + \frac{r_{\ell\ell'}}{\sqrt{D_\ell D_{\ell'}}}\ , \qquad (15)$$

where we have assumed that the off-diagonal terms are small. For cosmological models that have $\mathscr{C}_\ell^{\rm th}$ very different from the best-fit $\hat{\mathscr{C}}_\ell$, equation (15) does not yield the inverse of equation (13): in these cases the inversion of $\Sigma_{\ell\ell'}$ needs to be computed explicitly.

We do not propagate the *WMAP* 0.5% calibration uncertainty in the covariance matrix as this uncertainty does not affect cosmological parameter determinations. This systematic only affects the power spectrum amplitude constraint at the 0.5% level, while the statistical error on this quantity is $\sim$10%.

2.2.1. *Calibration with Monte Carlo Simulations*

The angular power spectrum is computed using three different weightings: uniform weighting in the signal-dominated regime ($\ell < 200$), an intermediate weighting scheme for $200 < \ell < 450$, and $N_{\rm obs}$ weighting (for the noise-dominated regime $450 < \ell \leq 900$; Hinshaw et al. 2003b). Uniform weighting is a minimum variance weighting in the signal-dominated regime, and $N_{\rm obs}$ weighting is a minimum variance in the noise-dominated regime. However, in the intermediate regime the weighting schemes are not necessarily optimal and the analytic expression for the covariance matrix might thus *underestimate* the errors. To ensure that we have the appropriate errors, we calibrate the covariance matrix from 100,000 Monte Carlo realizations of the sky with the *WMAP* noise level, symmetrized beams, and the Kp2 sky cut. A good approximation of the curvature matrix can be obtained by using equations (13)–(15), but substituting $\mathscr{N}_\ell$ and $f_{\rm sky}$ with $\mathscr{N}_\ell^{\rm eff}$ and $f_{\rm sky}^{\rm eff}$ calibrated from the Monte Carlo simulations, as shown in Figures 1 and 2.

We find that for $\ell < 200$ the weighting scheme is nearly optimal. The power spectrum covariance matrix (eq. [13]) gives a correct estimate of the error bars; thus, we do not need to calibrate $\mathscr{N}_\ell$ or $f_{\rm sky}$. We have computed an effective reduced χ^2,[14] $\chi^2_{\rm eff}/\nu \equiv -2\ln\mathscr{L}/\nu$, where ν is the number of degrees of freedom. The effective reduced χ^2 from the Monte Carlo simulations in this ℓ range is consistent with unity.

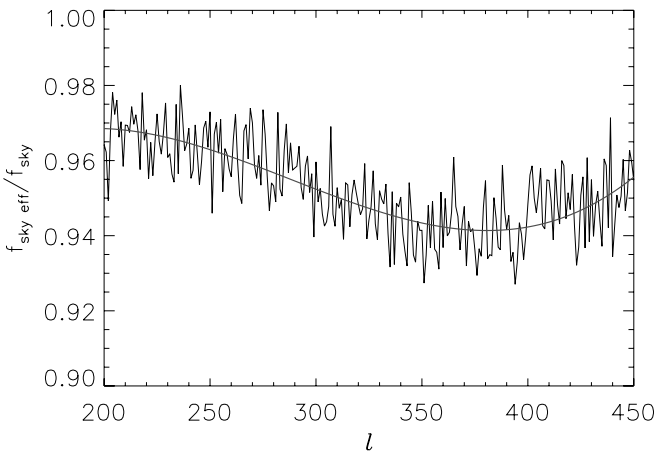

FIG. 1.—Ratio of the effective sky coverage to the actual sky coverage. This correction factor calibrates the expression for the Fisher matrix to the value obtained from the Monte Carlo approach. Here we show the ratio obtained from 100,000 simulations (*jagged line*); the smooth curve shows the fit we use, eq. (16). Note that, since we are switching between weighting schemes, the correction factors are not expected to smoothly interpolate between regimes.

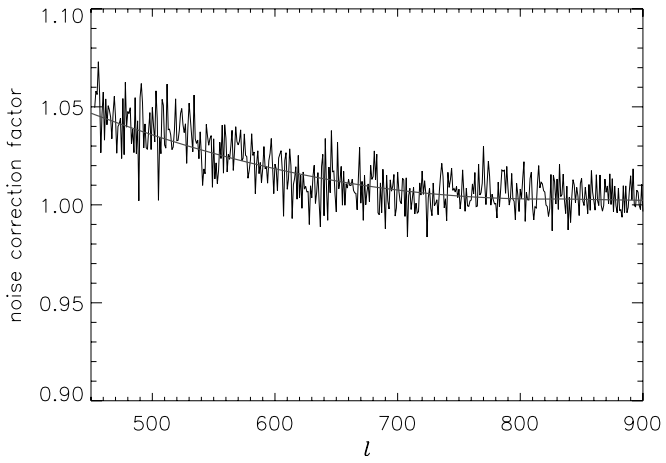

FIG. 2.—Correction factor for the noise. The lines are as in Fig. 1. Note that, since we are switching between weighting schemes, the correction factors are not expected to smoothly interpolate between regimes.

In the intermediate regime our *Ansatz* power spectrum covariance matrix (eq. [13]) slightly underestimates the errors. This can be corrected by computing the covariance matrix for an effective fraction of the sky $f_{\rm sky}^{\rm eff}$ as shown in Figure 1. The jagged line is the ratio obtained from the Monte Carlo simulations, while the smooth curve shows the fit to $f_{\rm sky}^{\rm eff}$ we adopt,

$$\frac{f_{\rm sky}^{\rm eff}}{f_{\rm sky}} = 0.813 + 0.001914\ell - 7.405\times10^{-6}\ell^2 + 8.65\times10^{-9}\ell^3 \qquad (16)$$

for $200 < \ell < 450$.

For $\ell > 450$, in the noise-dominated regime, the weighting is asymptotically optimal for $\ell \to \infty$. However, since we are using a smaller fraction of the sky, we need again to correct the $f_{\rm sky}$ factor. This numerical factor describes the reduction in effective sky coverage due to weighting the well-observed ecliptic poles more heavily than the ecliptic plane (see Fig. 3 of Bennett et al. 2003a). We fit this factor to the numerical simulations of the TT spectrum covariance matrix. Kogut et al. (2003) note that this same factor is also a good fit to the Monte Carlo simulations of the TE spectrum covariance matrix. For the noise-dominated regime, we define an effective sky fraction $f_{\rm sky}^{\rm eff} = f_{\rm sky}/1.14$ and an effective noise given by

$$\mathscr{N}_\ell^{\rm eff} = [\Sigma^{\rm sim}_{\ell\ell}(f_{\rm sky}^{\rm eff})^2(2\ell+1)/2]^{1/2} - \mathscr{C}_\ell^{\rm sim}\ ,$$

which can be obtained from the noise bias of the maps $\mathscr{N}_\ell$ by a noise correction factor $\mathscr{N}_\ell^{\rm eff}/\mathscr{N}_\ell$. This is shown in Figure 2, where the smooth curve is the fit we adopt to this

[14] This is not exactly the reduced χ^2 because the likelihood is non-Gaussian especially at low ℓ.

correction factor,

$$\frac{\mathcal{N}_\ell^{\rm eff}}{\mathcal{N}_\ell} = 1.046 - 0.0002346(\ell - 450) + 3.204 \times 10^{-7}(\ell - 450)^2 \quad (17)$$

for $\ell > 450$.

This calibration of the covariance matrix from the Monte Carlo simulations allows us to use the effective reduced χ^2 as a tool to assess goodness of fit. It can also be used to determine the relative likelihood of different models (e.g., Peiris et al. 2003).

2.3. *Likelihood for the TE Angular Power Spectrum*

Since the TE signal is noise dominated, we adopt a Gaussian likelihood, where the curvature matrix is given by

$$Q_{\ell\ell'}^{\rm TE} = \left(D_\ell^{\rm TE}\right)^{-1}\delta_{\ell\ell'}^{\rm K} + \frac{r_{\ell\ell'}^{\rm TE}}{\sqrt{D_\ell^{\rm TE} D_{\ell'}^{\rm TE}}} \, . \quad (18)$$

The expression for $D_\ell^{\rm TE}$ is given by equation (10) of Kogut et al. (2003), and the coupling coefficient due to the sky cut, $r_{\ell\ell'}^{\rm TE}$, is obtained from 100,000 Monte Carlo realizations of the sky with *WMAP* mask and noise level. The TE spectrum is computed with noise inverse weighting; in this regime $r_{\ell\ell'}$ depends only on the difference $\Delta_\ell = \ell - \ell'$ and is set to be 0 at separations $\Delta_\ell > 15$. We use all multipoles $2 \le \ell \le 450$, as comparison with the Monte Carlo realizations shows that in this regime equation (18) correctly estimates the TE uncertainties. We have also verified on the simulations that the Gaussian likelihood is an unbiased estimator and that the effective reduced χ^2 is centered around 1.

The amplitude of the covariance between TT and TE power spectra is $\sim\tilde{r}/(1 + C_\ell^{\rm EE}/n_\ell^{\rm EE})$, where $\tilde{r}$ is the correlation term $(C_\ell^{\rm TE})^2(C_\ell^{\rm EE}C_\ell^{\rm TT})^{-1} \simeq 0.2$. Since $C_\ell^{\rm EE}/n_\ell^{\rm EE} \ll 0.25$ for 1 yr data, we neglect this term, but we will include it in the 2+ yr analysis as it becomes increasingly important.

We provide a subroutine[15] that reads in a set of $\mathcal{C}_\ell^{\rm th}$ (TT, or TE, or both) and returns the likelihood for the *WMAP* data set, including all the effects described in this section.

3. MARKOV CHAIN MONTE CARLO LIKELIHOOD ANALYSIS

The analysis described in Spergel et al. (2003) and Peiris et al. (2003) is numerically demanding. At each point in the six-dimensional (or more) parameter space a new model from CMBFAST[16] (Seljak & Zaldarriaga 1996) is computed. Our version of the code incorporates a number of corrections and uses the RECFAST (Seager, Sasselov, & Scott 1999) recombination routine. Most of the likelihood calculations were done with four shared memory 32 CPU SGI Origin 300 with 600 MHz processors. With eight processors per calculation, each evaluation of CMBFAST for $\ell < 1500$ for a flat reionized Λ-dominated universe requires 3.6 s. (The scaling is not linear; with 32 processors each evaluation requires 1.62 s.)

A grid-based likelihood analysis would have required prohibitive amounts of CPU time. For example, a coarse grid ($\sim$20 grid points per dimension) with six parameters requires $\sim 6.4 \times 10^7$ evaluations of the power spectra. At 1.6 s per evaluation, the calculation would take $\sim$1200 days. Christensen & Meyer (2000) proposed using MCMC to investigate the likelihood space. This approach has become the standard tool for CMB analyses (e.g., Christensen et al. 2001; Knox, Christensen, & Skordis 2001; Lewis & Bridle 2002; Kosowsky, Milosavljevic, & Jimenez 2002) and is the backbone of our analysis effort. For a flat reionized Λ-dominated universe, we can evaluate the likelihood $\sim$120,000 times in less than 2 days using four sets of eight processors. As we explain below, this is adequate for finding the best-fit model and for reconstructing the 1 and 2 σ confidence levels for the cosmological parameters.

We refer the reader to Gilks, Richardson, & Spiegelhalter (1996) for more information about MCMC. Here we only provide a brief introduction to the subject and concentrate on the issue of convergence.

3.1. *Markov Chain Monte Carlo*

MCMC is a method to simulate posterior distributions. In particular, we simulate observations from the posterior distribution $\mathcal{P}(\alpha|x)$, of a set of parameters α given event x, obtained via Bayes's theorem,

$$\mathcal{P}(\alpha|x) = \frac{\mathcal{P}(x|\alpha)\mathcal{P}(\alpha)}{\int \mathcal{P}(x|\alpha)\mathcal{P}(\alpha)d\alpha} \, , \quad (19)$$

where $\mathcal{P}(x|\alpha)$ is the likelihood of event x given the model parameters α and $\mathcal{P}(\alpha)$ is the prior probability density. For our application the *WMAP* α denotes a set of cosmological parameters (e.g., for the standard, flat ΛCDM model these could be the cold dark matter density parameter Ω_c, the baryon density parameter Ω_b, the spectral slope n_s, the Hubble constant [in units of 100 km s^{-1} Mpc^{-1}] h, the optical depth τ, and the power spectrum amplitude A), and event x will be the set of observed $\acute{\mathcal{C}}_\ell$.

The MCMC generates random draws (i.e., simulations) from the posterior distribution that are a "fair" sample of the likelihood surface. From this sample, we can estimate all of the quantities of interest about the posterior distribution (mean, variance, confidence levels). The MCMC method scales approximately linearly with the number of parameters, thus allowing us to perform likelihood analysis in a reasonable amount of time.

A properly derived and implemented MCMC draws from the joint posterior density $\mathcal{P}(\alpha|x)$ once it has converged to the stationary distribution. The primary consideration in implementing MCMC is determining when the chain has *converged*. After an initial "burn-in"period, all further samples can be thought of as coming from the stationary distribution. In other words, the chain has no dependence on the starting location.

Another fundamental problem of inference from Markov chains is that there are always areas of the target distribution that have not been covered by a finite chain. If the MCMC is run for a very long time, the ergodicity of the Markov chain guarantees that eventually the chain will cover all the target distribution, but in the short term the simulations cannot tell us about areas where they have not been. It is thus crucial that the chain achieves good "mixing." If the Markov chain does not move rapidly

[15] The routine is available at http://lambda.gsfc.nasa.gov.

[16] We used the parallelized ver. 4.1 of CMBFAST developed in collaboration with Uros Seljak and Matias Zaldarriaga.

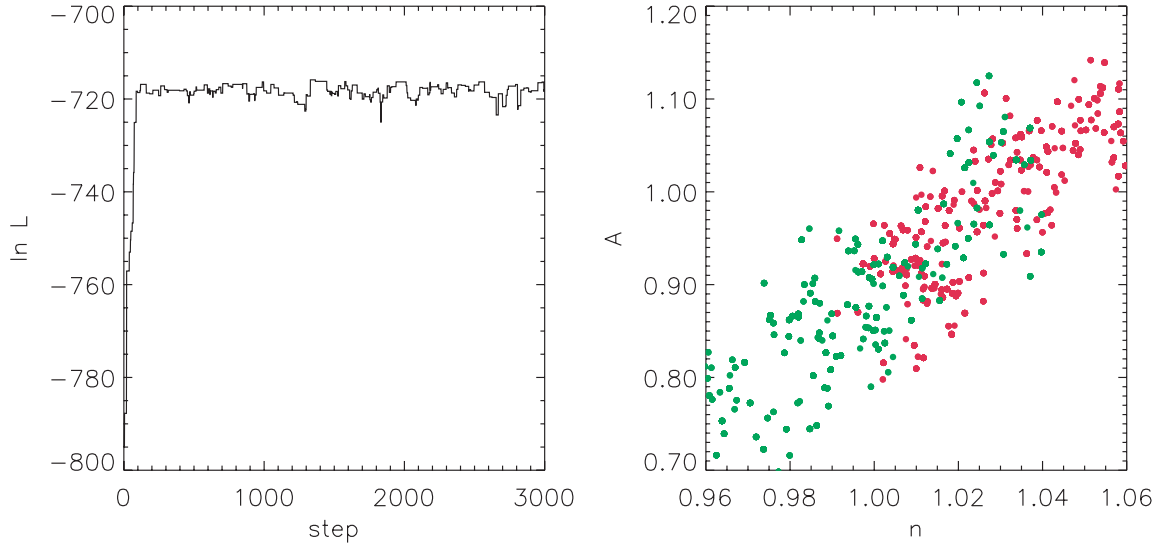

FIG. 3.—Unconverged Markov chains. The left-hand panel shows a *trace plot* of the likelihood values vs. iteration number for one MCMC (these are the first 3000 steps from one of our ΛCDM model runs). Note the burn-in for the first ∼100 steps. In the right-hand panel, red dots are points of the chain in the (n, A)-plane after discarding the burn-in. Green dots are from another MCMC for the same data set and the same model. It is clear that, although the trace plot may appear to indicate that the chain has converged, it has not fully explored the likelihood surface. Using either of these two chains at this stage will give *incorrect* results for the best-fit cosmological parameters and their errors.

throughout the support of the target distribution because of poor mixing, it might take a prohibitive amount of time for the chain to fully explore the likelihood surface. Thus, it is important to have a convergence criterion and a mixing diagnostic. Plots of the sampled MCMC parameters or likelihood values versus iteration number are commonly used to provide such criteria (Fig. 3, *left-hand panel*). However, samples from a chain are typically serially correlated; very high autocorrelation leads to little movement of the chain and thus makes the chain "appear" to have converged. For a more detailed discussion see Gilks et al. (1996). Using an MCMC that has not fully explored the likelihood surface for determining cosmological parameters will yield *wrong* results. We describe below the method we use to ensure convergence and good mixing.

3.2. *Convergence and Mixing*

We use the method proposed by Gelman & Rubin (1992) to test for convergence and mixing. They advocate comparing several sequences drawn from different starting points and checking to see that they are indistinguishable. This method not only tests convergence but can also diagnose poor mixing. *For any analysis of the WMAP data, we strongly encourage the use of a convergence criterion.*

Let us consider M chains (the analyses in Spergel et al. 2003 and Peiris et al. 2003 use four chains unless otherwise stated) starting at well-separated points in parameter space; each has $2N$ elements, of which we consider only the last N: $\{y_i^j\}$, where $i = 1, \ldots, N$ and $j = 1, \ldots, M$; i.e., y denotes a chain element (a point in parameter space), the index i runs over the elements in a chain, and the index j runs over the different chains. We define the mean of the chain

$$\bar{y}^j = \frac{1}{N}\sum_{i=1}^{N} y_i^j \tag{20}$$

and the mean of the distribution

$$\bar{y} = \frac{1}{NM}\sum_{ij=1}^{NM} y_i^j \,. \tag{21}$$

We then define the variance between chains as

$$B_n = \frac{1}{M-1}\sum_{j=1}^{M}(\bar{y}^j - \bar{y})^2 \tag{22}$$

and the variance within a chain as

$$W = \frac{1}{M(N-1)}\sum_{ij}(y_i^j - \bar{y}^j)^2 \,. \tag{23}$$

The quantity

$$\hat{R} = \frac{[(N-1)/N]W + B_n(1+1/M)}{W} \tag{24}$$

is the ratio of two estimates of the variance in the target distribution: the numerator is an estimate of the variance that is unbiased if the distribution is stationary, but it is otherwise an overestimate. The denominator is an underestimate of the variance of the target distribution if the individual sequences did not have time to converge.

The convergence of the Markov chain is then monitored by recording the quantity $\hat{R}$ for all the parameters and running the simulations until the values for $\hat{R}$ are always less than 1.1. A. Gelman (Kaas et al. 1997) suggests to use values for $\hat{R} < 1.2$. Here we conservatively adopt the criterion $\hat{R} < 1.1$ as our definition of convergence. We have found that the four chains will sometimes go in and out of convergence as they explore the likelihood surface, especially if the number of points already in the chain is small. To avoid this, one could run many chains simultaneously or run one chain for a very long time (e.g., Panter, Heavens, & Jimenez 2002). Because of CPU-time constraints, we run four chains

until they fulfill both of the following criteria: (1) they have reached convergence and (2) each chain contains at least 30,000 points. In addition to minimizing chance deviations from convergence, we find that this many points are needed to be able to robustly reconstruct the 1 and 2 σ levels of the marginalized likelihood for all the parameters. For most chains, the burn-in time is relatively rapid, so that we only discard the first 200 points in each chain; however, the results are not sensitive to this procedure.

3.3. *Markov Chains in Practice*

In this section we explain the necessary steps to run an MCMC for the CMB temperature power spectrum. It is straightforward to generalize these instructions to include the temperature-polarization power spectrum and other data sets. The MCMC is essentially a random walk in parameter space, where the probability of being at any position in the space is proportional to the posterior probability.

Here is our basic approach:

1. Start with a set of cosmological parameters $\{\alpha_1\}$ and compute the $\mathscr{C}_\ell^1$ and the likelihood $\mathscr{L}_1 = \mathscr{L}(\mathscr{C}_\ell^{1\rm th}|\hat{\mathscr{C}}_\ell)$.
2. Take a random step in parameter space to obtain a new set of cosmological parameters $\{\alpha_2\}$. The probability distribution of the step is taken to be Gaussian in each direction i with rms given by σ_i. We will refer below to σ_i as the "step size." The choice of the step size is important to optimize the chain efficiency (see § 3.4.2).
3. Compute the $\mathscr{C}_\ell^{2\rm th}$ for the new set of cosmological parameters and their likelihood $\mathscr{L}_2$.

4a. If $\mathscr{L}_2/\mathscr{L}_1 \geq 1$, "take the step," i.e., save the new set of cosmological parameters $\{\alpha_2\}$ as part of the chain, then go to step 2 after the substitution $\{\alpha_1\} \rightarrow \{\alpha_2\}$.

4b. If $\mathscr{L}_2/\mathscr{L}_1 < 1$, draw a random number x from a uniform distribution from 0 to 1. If $x \geq \mathscr{L}_2/\mathscr{L}_1$, "do not take the step," i.e., save the parameter set $\{\alpha_1\}$ as part of the chain and return to step 2. If $x < \mathscr{L}_2/\mathscr{L}_1$, "take the step," i.e., do as in step 4a.

5. For each cosmological model run four chains starting at randomly chosen, well-separated points in parameter space. When the convergence criterion is satisfied and the chains have enough points to provide reasonable samples from the a posteriori distributions (i.e., enough points to be able to reconstruct the 1 and 2 σ levels of the marginalized likelihood for all the parameters), stop the chains.

It is clear that the MCMC approach is easily generalized to compute the joint likelihood of *WMAP* data with other data sets.

3.4. *Improving MCMC Efficiency*

The Markov chain efficiency can be improved in different ways. We have tuned our algorithm by reparameterization and optimization of the step size.

3.4.1. *Reparameterization*

Degeneracies and poor parameter choices slow the rate of convergence and mixing of the Markov chain. There is one near-exact degeneracy (the geometric degeneracy) and several approximate degeneracies in the parameters describing the CMB power spectrum (Bond et al. 1994; Efstathiou & Bond 1999). The numerical effects of these degeneracies are reduced by finding a combination of cosmological parameters (e.g., Ω_c, Ω_b, h) that have essentially orthogonal effects on the angular power spectrum. The use of such parameter combinations removes or reduces degeneracies in the MCMC and hence speeds up convergence and improves mixing because the chain does not have to spend time exploring degeneracy directions. Kosowsky et al. (2002) introduced a set of reparameterizations to do just this. In addition, these new parameters reflect the underlying physical effects determining the form of the CMB power spectrum (we refer to these as physical parameters). This leads to particularly intuitive and transparent parameter dependencies of the CMB power spectrum.

Following Kosowsky et al. (2002), we use a core set of six physical parameters. There are two parameters for the physical energy densities of cold dark matter, $\omega_c \equiv \Omega_c h^2$, and baryons, $\omega_b \equiv \Omega_b h^2$. There is a parameter for the characteristic angular scale of the acoustic peaks,

$$\theta_A = \frac{r_s(a_{\rm dec})}{d_A(a_{\rm dec})} , \tag{25}$$

where $a_{\rm dec}$ is the scale factor at decoupling,

$$r_s(a_{\rm dec}) = \frac{c}{H_0\sqrt{3}} \int_0^{a_{\rm dec}} \left\{ \left(1 + \frac{3\Omega_b}{4\Omega_\gamma}\right) \left[(1-\Omega)x^2 + \Omega_\Lambda x^{1-3w} + \Omega_m x + \Omega_{\rm rad}\right] \right\}^{-1/2} dx \tag{26}$$

is the sound horizon at decoupling, and

$$d_A(a_{\rm dec}) = \frac{c}{H_0} \int_{a_{\rm dec}}^1 \left[(1-\Omega)x^2 + \Omega_\Lambda x^{1-3w} + \Omega_m x + \Omega_{\rm rad}\right]^{-1/2} dx \tag{27}$$

is the angular diameter distance at decoupling, where H_0 denotes the Hubble constant and c is the speed of light. Here $\Omega_m = \Omega_c + \Omega_b$, Ω_Λ denotes the dark energy density parameter, w is the equation of state of the dark energy component, $\Omega = \Omega_m + \Omega_\Lambda$, and the radiation density parameter $\Omega_{\rm rad} = \Omega_\gamma + \Omega_\nu$, where Ω_γ and Ω_ν are the photon and neutrino density parameters, respectively. For reionization we use the physical parameter $\mathscr{X} \equiv \exp(-2\tau)$, where τ denotes the optical depth to the last scattering surface (not the decoupling surface). The remaining two core parameters are the spectral slope of the scalar primordial density perturbation power spectrum, n_s, and the overall amplitude of the primordial power spectrum, A. Both are normalized at $k = 0.05\ {\rm Mpc}^{-1}$ ($\ell \sim 700$).

For more complex models we add other parameters as described in Spergel et al. (2003) and Peiris et al. (2003) and in § 5. To investigate nonflat models, we use the vacuum energy, $\omega_\Lambda \equiv \Omega_\Lambda h^2$. Other examples include the tensor index, n_t, the tensor-to-scalar ratio, r, and the running of the scalar spectral index, $dn_s/d\ln k$.

Here we relate the input parameter for the overall normalization, A, as in the CMBFAST code (ver. 4.1 with UNNORM option), to the amplitude of primordial comoving curvature perturbations $\mathscr{R}$, $\Delta^2_{\mathscr{R}}(k_0) \equiv (k^3/2\pi^2)\langle|\mathscr{R}|^2\rangle$. We also relate our convention for the tensor perturbations to the one in the code. CMBFAST calculates

$$C_\ell^S = (4\pi) T_0^2 \int \frac{dk}{k} \Delta^2_\Psi(k) \left[g^\Psi_{T\ell}(k)\right]^2 \tag{28}$$

$$C_\ell^T = (4\pi) T_0^2 \int \frac{dk}{k} \tilde{\Delta}^2_h(k) \left[g^h_{T\ell}(k)\right]^2 , \tag{29}$$

where Ψ is the Newtonian potential, $g_{T\ell}(k)$ is the radiation transfer function, and $T_0 = 2.725 \times 10^6$ is the CMB temperature in units of μK. The tilde denotes that $\tilde{\Delta}^2_h(k)$ is used in CMBFAST but differs from our convention, $\Delta^2_h(k)$, where $\tilde{\Delta}^2_h = \Delta^2_h/16$. The comoving curvature perturbation, $\mathscr{R}$, is related to Ψ by $\Psi = -\frac{3}{5}\mathscr{R}$; thus, $\Delta^2_{\mathscr{R}}(k) = (25/9)\Delta^2_{\Psi}(k)$. Note that this relation holds from radiation domination to matter domination with accuracy better than 0.5%.

CMBFAST uses A to parameterize $\Delta^2_{\Psi}(k_0)$. The tensor perturbations are calculated accordingly. The relations are

$$\Delta^2_{\Psi}(k_0) = \frac{800\pi^2}{T_0^2} A \, , \tag{30}$$

$$\tilde{\Delta}^2_h(k_0) = \frac{1}{16}\Delta^2_h(k_0) = \frac{r}{16}\Delta^2_{\mathscr{R}}(k_0) = \frac{25r}{144}\Delta^2_{\Psi}(k_0) \, . \tag{31}$$

Therefore, one obtains

$$\Delta_{\mathscr{R}}(k_0) = 2.95 \times 10^{-9} A \, . \tag{32}$$

The amplitude A is normalized at $k_1 = 0.05$ Mpc^{-1} and the tensor-to-scalar ratio r is evaluated at $k_0 = 0.002$ Mpc^{-1}, unless otherwise specified. To convert $A(k_0)$ to $A(k_1)$, we use

$$A(k_1) = A(k_0)\left(\frac{k_1}{k_0}\right)^{n_s(k_0)-1+(1/2)(dn_s/d\ln k)\ln(k_1/k_0)} \, . \tag{33}$$

3.4.2. *Step Size Optimization*

The choice of the step size in the Markov chain is crucial to improve the chain efficiency and speed up convergence. If the step size is too big, the acceptance rate will be very small; if the step size is too small, the acceptance rate will be high, but the chain will exhibit poor mixing. Both situations will lead to slow convergence. For our initial step sizes for each parameter we use the standard deviation for each parameter when all the other parameters are held fixed at the maximum likelihood value. These are easy to find once a preliminary chain has been run and the likelihood surface has been fitted, as explained in § 3.4.3. If a given parameter is roughly orthogonal to all the other parameters, it is not necessary to adjust the step size further; in the presence of severe degeneracies the step size estimate needs to be increased by a "banana correction" factor, which is approximately the ratio of the projection of the 1 σ error along the degeneracy to the projection perpendicular to the degeneracy.

With these optimizations the convergence criterion is satisfied for the four chains after roughly 30,000 steps each ($2N = 30,000$) for a model with six parameters. On an Origin 300 machine this takes roughly 32 hr running each chain on eight processors. These numbers serve only as a rough indication: convergence speed depends on the model and on the data set. For a fixed number of parameters, convergence can be significantly slower if there are severe degeneracies among the parameters. Adding more data sets might slow down the evaluation of a single step in the chain but can also speed up convergence by breaking degeneracies.

3.4.3. *Likelihood Surface Fitting*

The likelihood surface explored by the MCMC was found to be functionally well approximated by a quartic expansion of the cosmological parameters (for example, $\{\alpha_i\} = \{\omega_b, \omega_c, n, \theta_A, \mathscr{Z}, A\}$):

$$y \equiv \log \mathscr{L} = q_0 + \sum_i q_1^i \delta_i + \sum_{i\le j} q_2^{ij}\delta_i\delta_j + \sum_{i\le j\le k} q_3^{ijk}\delta_i\delta_j\delta_k + \sum_{i\le j\le k\le \ell} q_4^{ijk\ell}\delta_i\delta_j\delta_k\delta_\ell \, . \tag{34}$$

Here q are fit coefficients and δ_i are related to the cosmological parameters via $\delta_i = (\alpha_i - \alpha_i^0)/\alpha_i$, where α_i^0 is the maximum likelihood value of the parameter. Lower order expansions were unable to reproduce the likelihood surface. With six parameters there are $M_f = 210$ fit coefficients. Writing equation (34) as $y = \boldsymbol{q}\cdot\boldsymbol{x}$, the minimum least-squares estimator for $\boldsymbol{q}$ is

$$\boldsymbol{q} = \left(\boldsymbol{X}^T\boldsymbol{X}\right)^{-1}\boldsymbol{X}^T\boldsymbol{y} \, , \tag{35}$$

where $\boldsymbol{X}$ is the $N \times M_f$ matrix $X_{ij} = x_j^{(i)}$ and N is the number of unique points in the chain.

We run preliminary MCMC chains with "guesstimated" step sizes until there are $\sim$1000 unique points in total. Then we use equation (34) to cut through the likelihood surface at the maximum likelihood value to find the 1 σ level in each parameter direction (see § 3.4.2). This defines our "step size" for subsequent chains.

3.5. *The Choice of Priors*

From Bayes's theorem (eq. [19]) we can infer $\mathscr{P}(\alpha_i|x)$, the probability of the model parameters α_i given the event x (i.e., our observation of the power spectra), from the likelihood function once the prior is specified. It is reasonable to take prior probabilities to be equal when nothing is known to the contrary (Bayes's postulate). Unless otherwise stated, we assume uniform priors on the parameters given in the following list:

1. $0 \le \omega_c \le 1$.
2. $0 \le \omega_b \le 1$.
3. $0.005 \le \theta_A \le 0.1$.
4. $0 \le \tau \le 0.3$.
5. $0.5 \le A \le 2.5$.
6. $0 \le n_s|_{k_0} \le 2$.
7. $0 \le n_{\rm iso} \le 2$.
8. $0 \le f_{\rm iso} \le 5000$.
9. $-0.5 \le dn/d\ln k \le 0.5$.
10. $0 \le r \le 2.5$.
11. $0 \le \omega_\nu \le 1$.
12. -3.2 $(-1.2) \le w \le 0$ (we will present two sets of results, one with the prior $w \ge -1.2$ and the other with $w \ge -3.2$).
13. $0 \le \omega_\Lambda \le 1$.

Note that we assume uniform priors on ω_c, ω_b, and θ_A rather than uniform priors on Ω_m, Ω_b, and H_0.

Except for the priors on τ and w (the equation of state of the dark energy component), the MCMCs never hit the imposed boundaries; thus, most of our choices for priors have no effect on the outcome. A detailed discussion about the prior on τ is presented in Spergel et al. (2003).

We set a lower bound on w at -3.2 (-1.2), but we discard the region of parameter space where $w < -3$ $(w < -1)$. This is necessary because our best-fit value for this parameter is close to the boundary. If we had instead set the prior to be $w \ge -3$ $(w \ge -1)$, then the chains would fail to be a fair representation of the posterior distribution in the region of

parameter space where the distance from the boundary is comparable to the step size.

3.6. *MCMC Output Analysis*

We merge the four converged MCMCs ($\gtrsim$120,000 points) into one. From this we give the cosmological parameters that yield our best estimate of $\mathscr{C}_\ell$ and the marginalized distribution of the parameters. We compute the marginalized distribution for one parameter and the joint distribution for two parameters, obtained marginalizing over all the other parameters. Since the MCMC passes objective tests for convergence and mixing, the density of points in parameter space is proportional to the posterior probability of the parameters.

The marginalized distribution is obtained by projecting the MCMC points. For the marginalized parameter values $\bar{\alpha}_i$, Spergel et al. (2003) quote the expectation value of the marginalized likelihood, $\int \mathscr{L}\alpha_i \, d\alpha_i = 1/N \sum_t \alpha_{t,i}$. Here N is the number of points in the merged chain and $\alpha_{t,i}$ denotes the value of parameter α_i at the tth step of the chain. The last equality becomes clear if we consider that the MCMC gives to each point in parameter space a "weight" proportional to the number of steps the chain has spent at that particular location. The $100(1-2p)\%$ confidence interval $[c_p, c_{1-p}]$ for a parameter is estimated by setting c_p to the pth quantile of $\alpha_{t,i}$, $t = 1, \ldots, N$ and c_{p-1} to the $(1-p)$th quantile. The procedure is similar for multidimensional constraints: the density of points in the n-dimensional space is proportional to the likelihood, and multidimensional confidence levels can be found as illustrated in § 15.6 of Press et al. (1992).

We note that the global maximum likelihood value for the parameters does not necessarily coincide with the expectation value of their marginalized distribution if the likelihood surface is not a multivariate Gaussian. We find that, for most of the parameters, the maximum likelihood values of the global joint fit are consistent with the expectation values of the marginalized distribution.

A virtue of the MCMC method is that the addition of extra data sets in the joint analysis can efficiently be done with minimal computational effort from the MCMC output if the inclusion of extra data sets does not require the introduction of extra parameters or does not drive the param eters significantly away from the current best fit. For example, we add Lyα power spectrum constraints to MCMC's outputs, but we cannot do this for the 2dFGRS, since this requires the introduction of two extra parameters (β and σ_p; see § 5.1 for more details).

If the likelihood surface for a subset of parameters from an external (independent) data set is known, or if a prior needs to be added a posteriori, the joint likelihood surface can be obtained by multiplying the likelihood with the posterior distribution of the MCMC output. In Spergel et al. (2003) we follow this method to obtain the joint constraint of CMB with Type Ia supernova (Riess et al. 1998, 2001) data and CMB with Hubble Key Project Hubble constant (Freedman et al. 2001) determination.

There is yet another advantage of the MCMC technique. The current version of CMBFAST with the nominal interpolation settings is accurate to 1%, but *random* numerical errors can sometimes exceed this. As the precision of the CMB measurements improves, these effects can become problematic for any approach that calculates derivatives as a function of parameters. Because MCMC calculations average over $\sim$100,000 CMB calculations, the MCMC technique is much less sensitive than either grid-based likelihood calculations or methods that numerically calculate the Fisher matrix.

4. EXTERNAL CMB DATA SETS

The Cosmic Background Imager (CBI; Mason et al. 2002; Sievers et al. 2002; Pearson et al. 2002) and the Arcminute Cosmology Bolometer Array Receiver (ACBAR; Kuo et al. 2002) experiments complement *WMAP* by probing the amplitude of CMB temperature power spectrum at $\ell > 900$. These observations probe the Silk damping tail and improve our analysis in two ways: (1) they improve our ability to constrain the baryon density, the amplitude of fluctuations, and the slope of the matter power spectrum, and (2) they improve convergence by preventing the chains from spending long periods of time in large, moderately low likelihood regions of parameter space.

The CBI data set is described in Mason et al. (2002), in Pearson et al. (2002), and on their Web site.[17] We use data from the CBI *mosaic* data set (Pearson et al. 2002) and do not include the deep data set as the two data sets are not independent. We use the three band powers from the even binning at central ℓ-values of 876, 1126, and 1301, thus ensuring that the chosen band power can be considered independent from the *WMAP* data. At $\ell \gtrsim 1500$, the CBI experiment detected excess power. If the rms amplitude of mass fluctuations on scales of $8\,h^{-1}$ Mpc is $\sigma_8 \sim 1$, then this excess power can be interpreted as due to Sunayev-Zeldovich distortion from undetected galaxy clusters (Mason et al. 2002; Bond et al. 2002; Komatsu & Seljak 2002). We simplify our analyses by not using the CBI data on scales where this effect can be important. The correlations between different band powers are taken into account with the full covariance matrix; we use the lognormal form of the likelihood (as in Pearson et al 2002). In addition, we marginalize over a 10% calibration uncertainty (CBI beam uncertainties are negligible).

The ACBAR data set is described in Kuo et al. (2002). We use the seven band powers at multipoles 842, 986, 1128, 1279, 1426, 1580, and 1716. As shown in Figure 4, these points do not overlap with the *WMAP* power spectrum except at $\ell \sim 800$, where *WMAP* is noise dominated. As shown in Figure 4, the ACBAR experiment is less sensitive to Sunyaev-Zeldovich contamination than CBI. We compute the likelihood analysis for cosmological parameters for the ACBAR data set following Goldstein et al. (2002) and using the error bars given on the ACBAR Web site.[18] In addition, we marginalize over conservative beam and calibration uncertainties (B. Holzapfel 2002, private communication). In particular, we assume a calibration uncertainty of 20% (the double of the nominal value) and 5% beam uncertainty (60% larger than the nominal value).

The ACBAR and CBI data are completely independent from each other (they map different regions of the sky) and from the *WMAP* data (the band powers we consider span different ℓ ranges). To perform the joint likelihood analysis, we simply multiply the individual likelihoods.

[17] See http://www.astro.caltech.edu/~tjp/CBI/data/index.html (last update 2002 August).

[18] See http://cosmology.berkeley.edu/group/swlh/acbar/data.

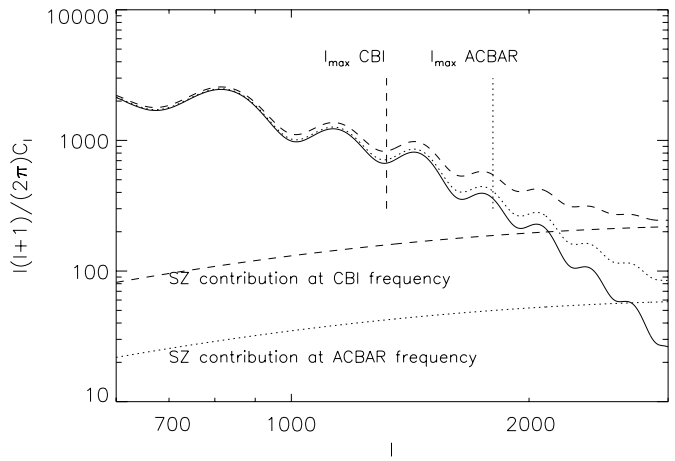

FIG. 4.—CMB angular power spectrum (in μK^2) for our best-fit ΛCDM model for $\ell > 800$ and the Sunayev-Zeldovich contribution for $\sigma_8 = 0.98$ for CBI wavelengths (*dotted line*) and for ACBAR (*dashed line*). The vertical line shows the adopted cutoff for CBI and ACBAR.

5. ANALYSIS OF LARGE-SCALE STRUCTURE DATA

We can enhance the scientific value of the CMB data from $z \sim 1089$ by combining it with measurements of the low-redshift universe. Galaxy redshift surveys allow us to measure the galaxy power spectrum at $z \sim 0$, and observations of Lyα absorption of about 50 quasar spectra (Lyα forest) allow us to probe the dark matter power spectrum at redshift $z \sim 3$.

We use the Anglo-Australian Telescope Two-Degree Field Galaxy Redshift Survey (2dFGRS; Colless et al. 2001) as compiled in 2001 February. This survey probes the universe at redshift $z_{\rm eff} \sim 0.1$ and probes the power spectrum on scales corresponding to $0.022 < k < 0.2$ (where k is in units of h Mpc^{-1}). The anticipated Sloan Digital Sky Survey (Gunn & Knapp 1993) power spectrum will be an important complement to 2dFGRS. We also use the linear matter power spectrum as recovered by Croft et al. (2002) from Lyα forest observations. This power spectrum is reconstructed at an effective redshift $z \sim 2.72$ and probes scales $k > 0.2\,h$ Mpc^{-1}. Together these data sets allow us to probe not only a wide range of physical scales—from $k \sim 1 \times 10^{-4}$ (30,000 Mpc h^{-1}) to $k \sim 1$ (3 Mpc h^{-1}) (see Fig. 5)—but also the evolution of a given scale with redshift.

When including LSS data sets, one should keep in mind that the underlying physics for these data sets is much more complicated and less well understood than for *WMAP* data, and systematic and instrumental effects are much more important. We attempt here to include all the known (up-to-date) uncertainties and systematics in our analysis. In what follows, we illustrate our modeling of the "real-world" effects of LSS surveys and how we propagate systematic and statistical uncertainties into the parameter estimation. The goal of our modeling is to relate not just the *shape* but also the *amplitude* of the observed power spectrum to that of the linear matter power spectrum as constrained by CMB data. The reason for this will be clear in § 5.1.3; by using the information in the power spectrum amplitude, we can break some of the degeneracies among cosmological parameters.

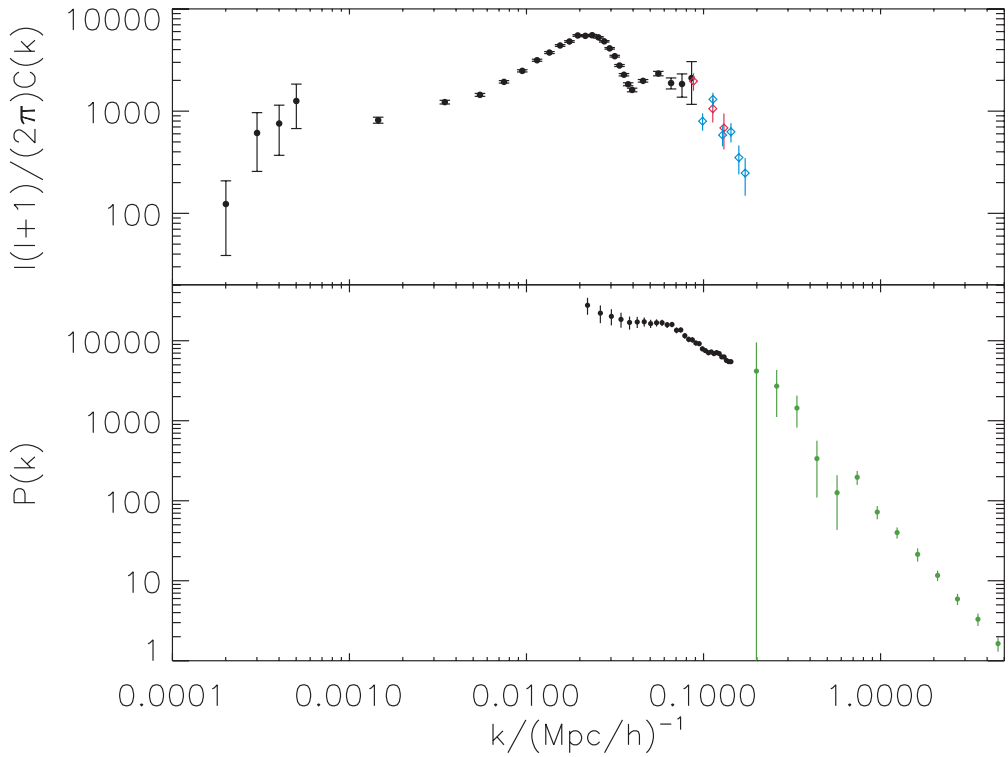

FIG. 5.—Combined CMB and LSS data set. *Top:* CMB angular power spectrum in μK^2 as a function of k, where k is related to ℓ by $\ell = \eta_0 k$ (where $\eta_0 \sim 14,400$ Mpc is the distance to the last scattering surface). Black points are the *WMAP* data, red points CBI, blue points ACBAR. *Bottom:* LSS data. Black points are the 2dFGRS measurements, and green points are the Lyα measurements. Both LSS power spectra are in units of (Mpc h^{-1})^3 and have been rescaled to $z = 0$. This plot only illustrates the scale coverage of all the data sets we consider. The various LSS power spectra as plotted here cannot be directly compared with the theory because of the effects outlined in § 5 (e.g., redshift-space distortions, nonlinearities, bias and window function effect).

5.1. *The 2dFGRS Power Spectrum*

The 2dFGRS power spectrum, as released in 2002 June, has been calculated from the 2001 February catalog that includes 140,000 galaxies (Percival et al. 2001). The full survey is composed of 220,000 galaxies but is not yet available. The sample is magnitude limited at $b_J = 19.45$ and thus probes the universe at $z_{\rm eff} \sim 0.1$ and the power spectrum on scales corresponding to $k > 0.015\ h\ {\rm Mpc}^{-1}$. The input catalog is an extended version of the Automatic Plate Machine (APM) galaxy catalog (Maddox et al. 1990b; Maddox, Efstathiou, & Sutherland 1990a, 1996), which includes about 5 million galaxies to $b_J = 20.5$. The APM catalog was used previously to recover the three-dimensional power spectrum of galaxies by inverting the clustering properties of the two-dimensional galaxy distribution (Baugh & Efstathiou 1993; Efstathiou & Moody 2001). These techniques, however, are affected by sample variance and uncertainties in the photometry; a full three-dimensional analysis is thus more reliable.

The power spectrum of the galaxy distribution as measured by LSS surveys, such as the 2dFGRS, cannot be directly compared to that of the initial density fluctuations, as predicted by theory, or recovered from *WMAP* or the combination of *WMAP*+CBI+ACBAR data sets. This is due to a number of intervening effects that can be broadly divided into two classes: effects due to the survey geometry (i.e., window function, selection function effects) and effects intrinsic to the galaxy distribution (e.g., redshift-space distortions, bias, nonlinearities).

5.1.1. *Survey Geometry*

Galaxy surveys such as the 2dFGRS are magnitude limited rather than volume limited; thus, most nearby galaxies are included in the catalog while only the brighter of the more distant galaxies are selected. The selection function accounts for the fact that fewer galaxies are included in the survey as the distance (or the redshift) increases. An additional effect arises from the fact that the clustering properties of bright galaxies might be different from the average clustering properties of the galaxy population as a whole. The selection function does not take this into account (we return to this point in § 5.1.2).

Moreover, the completeness across the sky is not constant, and the survey can only cover a fraction of the whole sky, sometimes with a very complicated geometry described by the window function. In particular, for the data we use, unobserved fields make the survey completeness a strongly varying function of position. The measured Fourier coefficients are therefore the true coefficients of the galaxy distribution convolved by the Fourier transform of the selection function (in the direction of the line of sight) and of the window function (on the plane of the sky). In this section we follow the standard notation used in LSS analyses and refer to all of these effects as window effects.

The window not only modifies the measured power spectrum but also introduces spurious correlations between Fourier modes (see Percival et al. 2001 for more details). For the 2dFGRS these effects have been quantified by Monte Carlo simulations of mock catalogs of the survey.[19] We include them in our analysis by convolving the theory power spectrum with the window "kernel" and by including off-diagonal terms in the covariance matrix.

5.1.2. *Effects Intrinsic to the Galaxy Distribution*

Linear gravitational evolution modifies the amplitude but not the shape of the underlying power spectrum. However, in the nonlinear regime (where the amplitude of fluctuations is $\delta\rho/\rho \sim 1$) this is no longer the case. Nonlinear gravitational evolution changes the shape of the power spectrum and introduces correlations between Fourier modes. This effect becomes important on scales $k \sim 0.1\ h\ {\rm Mpc}^{-1}$, but the exact scale at which it appears and its detailed characteristic depend on cosmological parameters. Most of the clustering signal from galaxy surveys such as 2dFGRS comes from the regime where nonlinearities are nonnegligible because shot noise is the dominant source of error at $k \gtrsim 0.5\ h\ {\rm Mpc}^{-1}$ and the number density of modes scales as k^3. These nonlinearities encode additional information about cosmology and motivate their inclusion in the present analysis. This approach is complicated by the fact that an accurate description of the fully nonlinear evolution of the *galaxy* power spectrum is complicated. In the literature, there are several different approaches to modeling the nonlinear evolution of the underlying *dark matter* power spectrum in real space: (1) linear (and extended) perturbation theory, (2) semianalytical modeling, and (3) numerical simulation. All of these approaches yield consistent results on the scales used in our analysis. We will use the semianalytical approach developed by Hamilton et al. (1991) and Peacock & Dodds (1996). In particular, we use the Ma et al. (1999) formulation of the nonlinear power spectrum. Figure 6 shows the effect of nonlinearities on the matter power spectrum on the scales of interest (compare solid and dashed lines).

Theory predicts the statistical properties of the continuous matter distribution, while observations are concerned with the galaxy distribution, which is discrete. Moreover, galaxies might not be faithful tracers of the mass distribution (i.e., the galaxy distribution might be *biased*). In the analysis of galaxy surveys it is assumed that galaxies form a *Poisson sampling* of an underlying continuous field that is related to the matter fluctuation field via the *bias*. It is possible to formally relate the discrete galaxy field and its continuous counterpart. For the power spectrum, this consists of the subtraction from the measured galaxy power spectrum of the shot-noise contribution. The published power spectra from galaxy surveys already have this contribution subtracted but are still biased with respect to the underlying mass power spectra.

The idea that galaxies are biased tracers of the mass distribution even on large scales was introduced by Kaiser (1984) to explain the properties of Abell clusters. Nevertheless, the fact that galaxies of different morphologies have different clustering properties (hence different power spectra) was known much before (e.g., Hubble 1936; Dressler 1980; Postman & Geller 1984). Since the clustering properties of different types of galaxies are different, they cannot all be good tracers of the underlying mass distribution.[20]

[19] For *WMAP* data, we deconvolve the raw measured $\tilde{C}_\ell$ by the effect of the window (the mask), thus leaving the effect of the window function and the mask only in the Fisher matrix. For LSS we will convolve the theory with the window, project the power spectrum into redshift space, and compare this to the observed power spectrum.

[20] Galaxies are likely to be formed in the very high density regions of the matter fluctuation field; thus, they are formed very biased at $z \gg 0$ (e.g., Lyman break galaxies), but then gravitational evolution should make the galaxy distribution less and less biased as time goes on (e.g., Fry 1996).

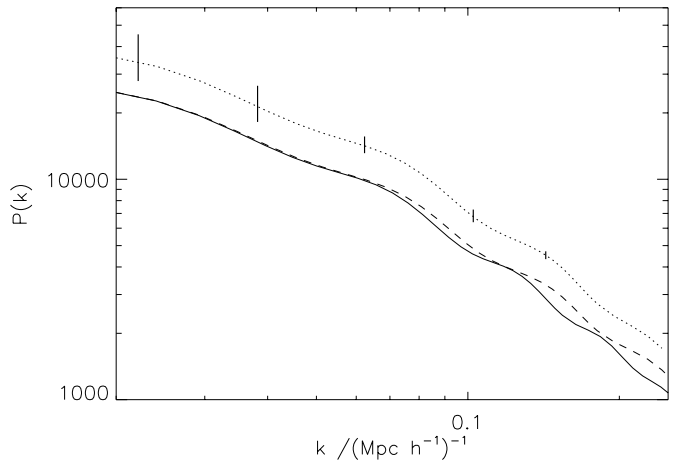

FIG. 6.—Matter power spectrum in (Mpc h^{-1})^3, linear in real space (*solid line*), nonlinear in real space (*dashed line*), and nonlinear in redshift space (*dotted line*). The error bars on the dotted line show the size of the statistical error bars per k bin of the 2dFGRS galaxy power spectrum. The power spectrum is in units of (Mpc h)^3.

In the simplest biasing model, the linear bias model, the mass and galaxy fractional overdensity fields δ and δ_g are related by $\delta_g(\boldsymbol{x}) = b\delta(\boldsymbol{x})$. This implies that on all scales

$$P_g(k) = b^2 P(k) \ . \tag{36}$$

This simple model (although justified by the Kaiser 1984 assumption that galaxies form on the highest peaks of the mass distribution) cannot be true in detail for two reasons. The first is that, on a fundamental level, the galaxy fluctuation field on small smoothing scales could become $\delta_g < -1$, which corresponds to a negative galaxy density. The second is that, from an observational point of view, this scheme leaves the shape of the power spectrum unchanged while not all galaxy populations have the same observed power spectrum shape, although the differences are not large (e.g., Peacock & Dodds 1994; Norberg et al. 2001). Many different and more complicated biasing schemes have been introduced in the literature. For our purposes it is important to note that the bias of a sample of galaxies depends on the sample selection criteria and on the weighting scheme used in the analysis. Thus, different surveys will have different biases, and care must be taken when comparing the different galaxy power spectra.

There are several indications that large-scale galaxy bias is scale independent on large scales (e.g., Hoekstra et al. 2002; Verde et al. 2002). This justifies adopting equation (36). For the 2dFGRS, the bias of galaxies has been measured by Verde et al. (2002), by using higher order correlations of the galaxy fluctuation field. They assume a generalization of the simple linear biasing scheme, $\delta_g = b_1\delta + b_2/2\delta^2$. They find no evidence for scale-dependent bias at least on linear and mildly nonlinear scales (i.e., $k < 0.4\,h\,\mathrm{Mpc}^{-1}$) and b_2 consistent with 0. This finding further supports the use of equation (36). In particular, they find $b_1 = 1.06 \pm 0.11$. In our analysis we assume linear biasing.

The Verde et al. (2002) bias measurement has to be interpreted with care. It applies to 2dFGRS galaxies weighted with a modification of the Feldman, Kaiser, & Peacock (1994) weighting scheme as described in Percival et al. (2001). It is important to note that, close to the observer, dim galaxies are included in the survey; the galaxy density is high, but a small volume of the sky is covered. On the other hand, far away from the observer, only very bright galaxies are included in the survey; a large volume is probed, but the galaxy density is low. As a consequence, clustering of dim galaxies in a small volume close to the observer contains most of the signal for the power spectrum at small scales. While rare, bright galaxies in a large volume enclose most of the information about the power spectrum on large scales. An "optimal" weighting scheme would thus weight dim galaxies on small scales and bright galaxies on large scales. This weighting scheme is, unfortunately, biased. Bright galaxies are more strongly clustered (i.e., more biased) than dim ones. This effect is known as "luminosity bias." The power spectrum recovered from such a weighting scheme will have optimal error bars but will exhibit scale-dependent bias. The weighting scheme used in Percival et al. (2001) is not optimal but is virtually unaffected by luminosity bias (Percival 2003). The power spectrum so obtained is that of two $L*$ galaxies on virtually all scales, and the effective redshift for the power spectrum is $z_{\mathrm{eff}} = 0.17$, slightly larger than the effective redshift of the survey as defined by the selection function (Percival et al. 2001; Peacock et al. 2001).

The final complication is that galaxy catalogs use the redshift as the third spatial coordinate. In a perfectly homogeneous Friedman universe, redshift would be an accurate distance indicator. Inhomogeneities, however, perturb the Hubble flow and introduce peculiar velocities. As Kaiser (1987) emphasized, the peculiar velocities distort the clustering pattern not only on small scales, where virialized objects produce "Fingers of God," but also on large scales, where coherent flows produce large-scale distortion components.

On large (linear) scales the redshift-space effect on an individual Fourier component of the density fluctuation field $\delta_{\boldsymbol{k}}$ can be modeled by

$$\delta_{\boldsymbol{k}} \rightarrow \delta^s_{\boldsymbol{k}} = \delta_{\boldsymbol{k}}\left(1 + \beta\mu^2\right) \ , \tag{37}$$

where the superscript s refers to the quantity in redshift space and μ is the cosine of the angle between the $\boldsymbol{k}$-vector and the line of sight. The Kaiser factor, β, is the linear redshift-space distortion parameter. One defines $\beta = f/b$, where $f = d\ln\delta/d\ln a$, with $\delta = \delta\rho/\rho$ and $a = (1+z)^{-1}$; b is the linear bias parameter. The expression for $f(z)$ is a known function of Ω_m, Λ, and z (Lahav et al. 1991),

$$f(\Omega_m, \Lambda, z) = X^{-1}\left[\frac{\Lambda}{(1+z)^2} - \frac{\Omega_m}{2}(1+z)\right] - 1 + (1+z)^{-1}X^{-3/2}\left[\int_0^{1/(1+z)} X^{-2/3}\,da\right]^{-1} , \tag{38}$$

where $X = 1 + \Omega_m z + \Lambda(a^2 - 1)$ and can be approximated by[21] $\beta \simeq \Omega_m^{0.6}/b$. The analysis of the 2dFGRS (Peacock et al. 2001; Percival et al. 2001) constrains f at the effective redshift of the survey. The effective redshift of the survey depends on the galaxy weighting scheme adopted to compute the power spectrum for the above work ($z_{\mathrm{eff}} \sim 0.17$). This peculiar velocity infall causes the overdensity to appear

[21] In our analysis we use the exact expression for β as in eq. (38).

squashed along the line of sight. The net effect on the angle-averaged power spectrum in the small-angle approximation is

$$P^s(k) = P(k)\left(1 + \tfrac{2}{3}\beta + \tfrac{1}{5}\beta^2\right) . \quad (39)$$

Thus, on large scales the redshift-space distortions boost the power spectrum if $\beta > 0$.

On smaller scales, virialized motions produce a radial smearing and the associated "Fingers of God" effect contaminates the wavelengths we are interested in. This is difficult to treat exactly, but as it is a smearing effect, it produces a mild damping of the power, acting in the opposite direction to the large-scale boosting by the Kaiser effect (see, e.g., Matsubara 1994). On these scales, the redshift-space correlation function is well modeled as a convolution of the real-space isotropic correlation function with some distribution function for the line-of-sight velocities (e.g., Davis & Peebles 1983; Cole, Fisher, & Weinberg 1994; Fisher 1995). Since the convolution in real space is equivalent to multiplication in Fourier space, the redshift-space power spectrum on small scales is multiplied by the square of the Fourier transform of the velocity distribution function (e.g., Peacock & Dodds 1994),

$$P^s(k,\mu) = P(k)\left(1 + \beta\mu^2\right)^2 D(k\sigma_p\mu) , \quad (40)$$

where σ_p denotes the line-of-sight pairwise velocity dispersion. If the pairwise velocity distribution is taken to be an exponential (e.g., Ballinger, Heavens, & Taylor 1995; Ballinger, Peacock, & Heavens 1996; Hatton & Cole 1998), which seems to be supported by simulations (e.g., Zurek et al. 1994) and observations (e.g., Marzke et al. 1995), then the damping factor is a Lorentzian (see also Kang et al. 2002),

$$D(k\sigma_p\mu) = \frac{1}{1 + (k^2\sigma_p^2\mu^2)/2} . \quad (41)$$

We adopt this functional form as it is used by Peacock et al. (2001) in determining the redshift-space distortion parameters β and σ_p from the 2dFGRS. The overall effect for the power spectrum in a thin shell in k-space is given by

$$P^s(k) = \left[4\frac{(\sigma_p^2k^2 - \beta)\beta}{\sigma_p^4k^4} + \frac{2\beta^2}{3\sigma_p^2k^2} + \frac{\sqrt{2}(k^2\sigma_p^2 - 2\beta)^2 \arctan\left(k\sigma_p/\sqrt{2}\right)}{k^5\sigma_p^5}\right] P(k) \quad (42)$$

obtained by averaging over μ in equation (40) with the damping factor given by equation (41). Figure 6 shows the effect of redshift-space distortions (eq. [42]) on the scales of interest.

This model is simplistic for several reasons. The most important is that, because of the complicated geometry of the survey, the simple angle average operation performed to obtain equation (42) might not be strictly correct. In addition, equation (42) is obtained in the plane-parallel (also known as small-angle) approximation (i.e., as if the lines of sight to different galaxies on the sky were parallel).

We have performed extensive testing of equation (42) using mock 2dFGRS catalogs obtained from the Hubble volume simulation. We find that the simulations' redshift-space power spectrum is consistent, given the errors, with equation (42), where $P(k)$ is the simulations' real-space power spectrum up to $k < 0.4\,h\,\mathrm{Mpc}^{-1}$, even for the complicated geometry of the 2dFGRS. This means that up to $k \sim 0.4$ the systematics introduced by equation (42) is smaller than the statistical errors; in the analysis we use only $k \lesssim 0.15$. However, the value for β in equation (42) needs to be calibrated on Monte Carlo realizations of the survey. We find that $\beta^{\mathrm{eff}} = 0.85\beta$. We have verified that our results for the cosmological parameters are insensitive to the exact choice of the correction factor. Peacock et al. (2001) measured the parameters β and σ_p and their joint probability distribution from the survey, obtaining $\beta = 0.43$ and $\sigma_p = 385$ km s^{-1}. This measurement has been obtained by using the full angular dependence of the power spectrum and therefore recovers directly β and not β^{eff}. Hawkins et al. (2002) measured these parameters from a larger sample than the one from Peacock et al. (2001), obtaining a slightly different result. This is mostly due to a shift in the recovered value for σ_p. Since most of the galaxies in the Hawkins et al. (2002) sample are in the Peacock et al. (2001) sample, we conservatively extend our error bars on β and σ by 10% and 30%, respectively, to include the new value within the 1 σ marginalized confidence contour and to include a possible error in the determination of β^{eff}. Figure 6 illustrates the importance of including all the above effects in our analysis.

In our analysis we consider data in the k range $0.02\,h\,\mathrm{Mpc}^{-1} < k < 0.2\,h\,\mathrm{Mpc}^{-1}$. On large scales the limit is set by the accuracy of the window function model; on small scales the limit is set by where the covariance matrix has been extensively tested. In this regime we also have a weak dependence on the velocity dispersion parameter σ_p, the parameter with the largest systematic uncertainty.

5.1.3. *Motivation for this Modeling*

The motivation behind the complicated modeling of §§ 5.1.1 and 5.1.2 is to be able to infer the amplitude of the matter power spectrum from the observed galaxy clustering properties.

Figures 7 and 8 illustrate how the modeling of §§ 5.1.1 and 5.1.2 helps in breaking degeneracies among cosmological parameters. For illustration, we consider two cases below: the degeneracy of the dark energy equation of state, w (Huey et al. 1999), with Ω_b and n_s and the $\omega_\nu - h$ degeneracy, where $\omega_\nu = \Omega_\nu h^2$.

Figure 7 shows two models that are virtually indistinguishable with CMB data, but which predict different amplitudes for the matter power spectra at $z \sim 0$. This is because the linear growth factor and the shape parameter Γ are different for the two cases. The two models differ in the values of ω_b, n_s, and w. The solid line is a model with $w = -0.4$, while the dotted line is a model with $w = -1$.

In Figure 8 we show two sets of cosmological parameters that differ only in the values of the neutrino mass and the Hubble constant. These two models are virtually indistinguishable with CMB observations. However, the matter power spectrum in the two cases is different in shape and amplitude. Since redshift-space distortions and window function affect the power spectrum shape, extra information about cosmological parameters is encoded in its amplitude. By using this information, Spergel et al. (2003) obtain a cosmological upper bound on the neutrino mass that is $\sim$4 times better than current cosmological constraints (Elgarøy et al. 2002).

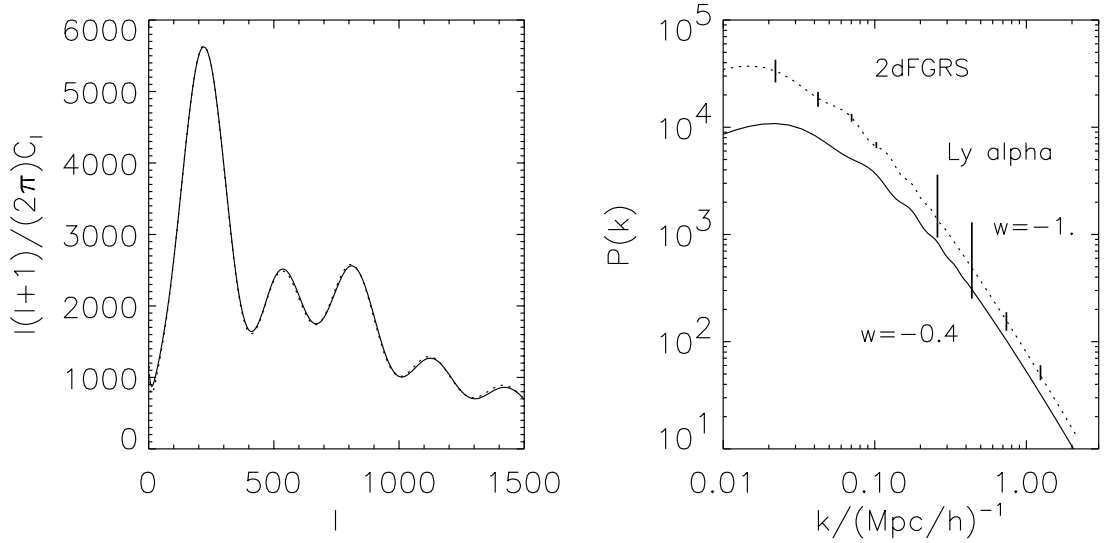

FIG. 7.—Two cosmological models: $\omega_b = 0.0235$, $\omega_m = 0.143$, $n_s = 0.978$, $\tau = 0.11$, $w = -0.426$, $h = 0.53$ (*solid line*) and $\omega_b = 0.0254$, $\omega_m = 0.137$, $n_s = 1.024$, $\tau = 0.08$, $w = -1$, $h = 0.77$ (*dotted line*). The two models are indistinguishable within current error bars from the CMB angular power spectrum (*left-hand panel*; units for the power spectrum are μK^2). However, they can easily be distinguished if we can relate the observed power spectrum to the underlying matter power spectrum [*right-hand panel*; units for the power spectrum are $(\mathrm{Mpc}\, h^{-1})^3$]. The error bars on the solid line are examples of the size of the 2dFGRS and Lyα power spectra statistical error bars for one data point at different scales. There are four error bars for 2dFGRS and four for Lyα .

For completeness, we have shown the power spectrum also for scales probed by the Lyα forest (see § 6). The error bars in Figures 7 and 8 are examples of the size of the 2dFGRS and Lyα power spectra statistical uncertainties in one data point, showing that the two models can be distinguished if the observed power spectrum can be related to the linear matter power spectrum without introducing large additional uncertainties.

5.1.4. *Practical Approach*

The procedure we adopt in order to compare the observed galaxy power spectrum with the theory predictions is outlined below (the published 2dFGRS galaxy power spectrum has been already corrected for shot noise). For a given set of cosmological parameters and a pairwise velocity dispersion parameter we proceed as follows:

1. The MCMC selects a set of cosmological parameters and values for β and σ_p. CMBFAST computes the theoretical linear matter power spectrum at $z = 0$.
2. We evolve the theoretical linear matter power spectrum to obtain the nonlinear matter power spectrum at the effective redshift of the survey, following the prescription of Ma et al. (1999).

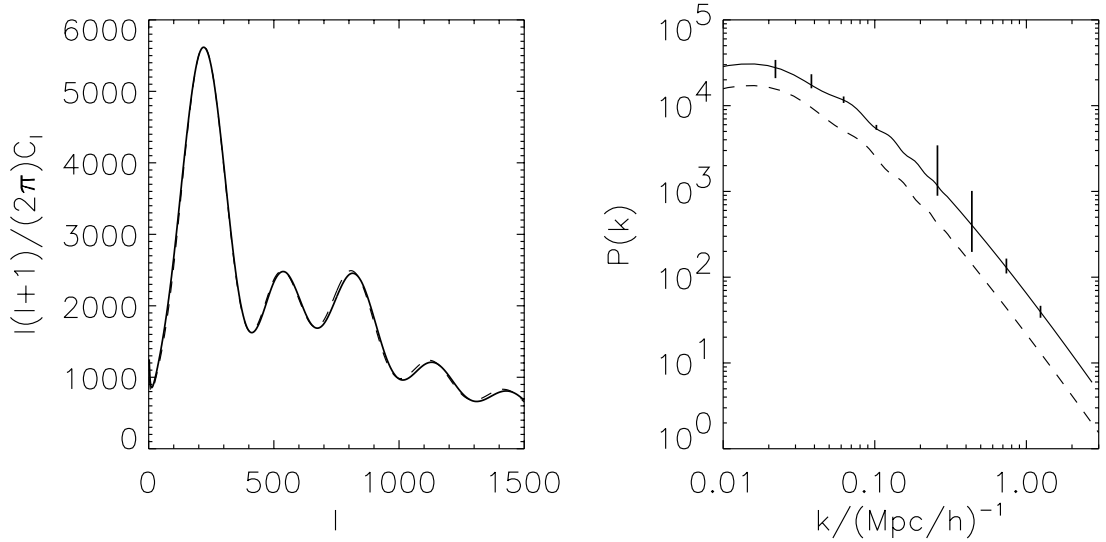

FIG. 8.—Two cosmological models: $\Omega_m = 0.26$, $\omega_b = 0.02319$, $\tau = 0.12$, $n_s = 0.953$, $\omega_\nu = 0$, $h = 0.714$ (*solid line*) and $\Omega_m = 0.26$, $\omega_b = 0.02319$, $\tau = 0.12$, $n_s = 0.953$, $\omega_\nu = 0.02$, $h = 0.6$ (*dashed line*). As before the two models are virtually indistinguishable from the CMB angular power spectrum (*left-hand panel*; units for the power spectrum are μK^2), but they can easily be distinguished if the matter power spectrum amplitude is known [*right-hand panel*; units for the power spectrum are $(\mathrm{Mpc}\, h^{-1})^3$]. The error bars on the solid line are examples of the size of the 2dFGRS and Lyα power spectra statistical error bars for one data point. There are four error bars for 2dFGRS and four for Lyα .

3. We then obtain the redshift-space power spectrum for the mass by using equation (42) with $\beta^{\rm eff}$ calibrated from Monte Carlo realizations of the catalog.

4. The bias is computed from β and Ω_m using equation (38). The galaxy power spectrum is obtained by correcting for bias, equation (36).

5. The resulting power spectrum is convolved with the galaxy window function. We use the routine provided on the 2dFGRS Web site to perform this numerically. This is the power spectrum that can be compared with the quantity measured from a galaxy survey.

6. We can now evaluate the likelihood using the full covariance matrix as provided by the 2dFGRS team. We approximate the likelihood to be Gaussian as it was done by the team. In principle, this is not strictly correct since in the linear regime the power spectrum is an exponential distribution and in the nonlinear regime the distribution has contributions from higher orders correlations. However, because of the size of the survey we are considering, the central limit theorem ensures that the likelihood is well described by the Gaussian approximation (e.g., Scoccimarro, Zaldarriaga, & Hui 1999). Moreover, the covariance matrix for the 2dFGRS power spectrum has been computed by the 2dFGRS team under the assumption that the likelihood is Gaussian.

We assume that the likelihood for the bias parameter is Gaussian, centered on $b = 1.04$ with dispersion $\sigma_b = 0.11$. This is a conservative overestimate of the error on the bias parameter, as noted in Verde et al. (2002). The determination of b is correlated with β and σ_p, and the error quoted has already been marginalized over the uncertainties in these two parameters. In practice, for each step in the Markov chain we compute the likelihood according to items 1–6 above. The bias parameter is determined once β, σ_p, and the other cosmological parameters are chosen. We then multiply the likelihood by the joint likelihood for β and σ_p, as in Figure 4 of Peacock et al. (2001), and by the likelihood for the bias parameter. In effect, we use the determination of β, σ_p, and b as priors. By multiplying the likelihood, we assume that the *measurements* of the redshift-space distortion parameters, bias, and the 2dFGRS power spectrum are independent. We justify this assumption below.

The parameters needed to map the real-space nonlinear matter power spectrum onto the redshift-space galaxy spectrum are β, σ_p, and b. These three parameters are not independent: not only is $\beta \propto 1/b$, but, more importantly, the three parameters are measured from the same catalog that we are using to constrain other cosmological parameters. However, the information we use to constrain cosmological parameters is all encoded in the shape and amplitude of the angle-averaged power spectrum. The information used to measure β and σ_p is all encoded in the dependence of the Fourier coefficients (i.e., of the power spectrum) on the angle from the line of sight. Thus, we can treat the determinations of β and σ_p as independent from the likelihood for cosmological parameters. The analysis of Verde et al. (2002) to measure the bias parameter from the 2dFGRS uses both information about the amplitude of the Fourier coefficients and their angular dependence. This dependence, however, is not that introduced by redshift-space distortions but is the configuration dependence of the bispectrum. Thus, in principle, we should not treat this measurement as completely independent. However, most of the signal for the bias measurement comes from the k range of $0.2\ h\ {\rm Mpc}^{-1} < k < 0.4\ h\ {\rm Mpc}^{-1}$, while the signal for the present analysis comes from $k < 0.2\ h\ {\rm Mpc}^{-1}$. Note that the configuration dependence of the bispectrum is largely independent of cosmology. This allows us to easily include a prior for the bias parameter in the analysis.

6. Lyα FOREST DATA

The Lyα forest traces the fluctuations in the neutral gas density along the line of sight to distant quasars. Since most of this absorption is produced by low-density unshocked gas in the voids or in mildly overdense regions that are thought to be in ionization equilibrium, this gas is assumed to be an accurate tracer of the large-scale distribution of dark matter. In this epoch and on these scales the clustering of dark matter is still in the linear regime.

Since the Lyα forest observations are probing the distribution of matter at $z \sim 3$, they are an important complement to the CMB data and the galaxy survey data. Because of their importance, there has been extensive numerical and observational work testing the notion that they trace the LSS. In our analyses, we find that the addition of Lyα forest data appears to confirm trends seen in other data sets and tightens cosmological constraints. However, more observational and theoretical work is still needed to confirm the validity of the emerging consensus that the Lyα forest data trace the LSS.

Recent papers use two different approaches for analysis of the Lyα forest power spectrum data. McDonald et al. (2000) and Zaldarriaga, Hui, & Tegmark (2001) directly compare the observed transmission spectra to the predictions from cosmological models. We follow the approach of Croft et al. (2002) and Gnedin & Hamilton (2002, hereafter GH), who use an analytical fitting function to recover the matter power spectrum from the transmission spectrum.[22]

GH factorize the linear power spectrum into four terms,

$$P_L(k) = P^{\rm fact}(k) Q_\Omega Q_T Q_\tau\ , \tag{43}$$

where $P^{\rm fact}(k)$ is a quantity that is independent of modeling and is almost directly measured. The other parameters convert this quantity into the linear matter power spectrum and encode the dependence on cosmology and the modeling of the intergalactic medium (IGM). In our treatment, we use the values of $P^{\rm obs}(k)$ (the estimator from Lyα forest observations of $P^{\rm fact}$) from GH and their parameterization in terms of equation (43) because it allows us to explicitly include the dependence of the recovered linear matter power spectrum on the cosmological parameters. Q_Ω encodes the dependence of the recovered linear power spectrum on the matter density parameter at $z = 2.72$. For Q_Ω we use the GH *Ansatz* of

$$Q_\Omega = \left(\frac{2.4}{1 + \Omega_m^{0.6} 1.4}\right)^2 . \tag{44}$$

[22] After the present paper was submitted, a preprint appeared (Seljak et al. 2003) claiming that the treatment of GH and Croft et al. (2002) significantly underestimates the errors. Given the importance of this data set to tighten cosmological constraints, the Lyα forest community should reach a consensus on the interpretation of these observations and on the level of systematic contamination.

$Q_T = 20,000\ \mathrm{K}/T_0$ ($T_0 \sim 20,000$ K) parameterizes the dependence on the mean temperature of the IGM, and $Q_\tau \sim 1.11$ parameterizes the dependence on the assumed mean optical depth. In addition to the statistical errors, GH quote a systematic uncertainty that we add to the statistical one. Finally, the uncertainties in Q_Ω, Q_T, and Q_τ contribute to the overall normalization uncertainty. We use the Croft et al. (2002) prescription to parameterize this uncertainty as $\ln \mathscr{P}(A) = -1/2(A-1)^2/\sigma^2_{\mathrm{Ly}\alpha}$, where if $A \leq 1$, then $\sigma_{\mathrm{Ly}\alpha} = 0.25$, while if $A > 1$, $\sigma_{\mathrm{Ly}\alpha} = 0.29$.

N-body simulations are used to convert the flux power spectrum to the dark matter power spectrum and calibrate the form of equation (43). The two different groups, GH and Croft et al. (2002), have done this independently. The resulting power spectra agree well within the 1 σ errors for all data points except the last three. We thus increase the 1 σ uncertainties on the last three data points to make the two determinations of $P_L(k)$ consistent and use this as a rough measure of the intrinsic systematic uncertainties in the Lyα data.

GH point out that the correlation in flux measured from the Lyα forest samples power over a finite band of wavenumbers. The effective band-power windows are rather broad as a result of the peculiar velocities that smear power on scales comparable to the one-dimensional velocity dispersion. Thus, the recovered linear power spectrum is effectively smoothed with a window that becomes broader at smaller scales. In principle, the resulting covariance between estimates of power at different k needs to be taken into account to do a full likelihood analysis to extract cosmological parameters. However, the full covariance matrix is not available. Since the Lyα data are such a powerful tool, we just perform a simple χ^2 fit and caution the reader that interpreting the reduced χ^2 as a measure of goodness of fit for this data set is not meaningful since the data are strongly correlated.

To marginalize over the overall normalization uncertainty, we take advantage of the MCMC approach. In principle, we could marginalize over it analytically, as we do for the calibration uncertainty. Instead, at each step of the chain we compute the best-fit amplitude $\bar{A}$ as done for point sources (Hinshaw et al. 2003b),

$$\bar{A} = \left[\sum_k \frac{P^{\mathrm{obs}}(k) P_L(k)}{\sigma_k^2}\right] \left[\sum_k \frac{P^{\mathrm{obs}}(k)^2}{\sigma_k^2}\right]^{-1}. \quad (45)$$

The likelihood for the Lyα data for the model is given by $\ln \mathscr{L}_{\mathrm{Ly}\alpha} = \ln \mathscr{L}(P^{\mathrm{obs}}|\bar{A}, P_L) + \ln \mathscr{P}(\bar{A})$. The marginalization is then automatically obtained from the MCMC output. In other words, the analytic marginalization computes $\int \mathscr{P}(\mathrm{data}|\mathrm{model})\mathscr{P}(A)dA$, while we compute an estimator of this given by $\int \mathscr{P}(\mathrm{data}|\mathrm{model})\mathscr{P}(\bar{A})d\bar{A}$.

7. CONCLUSIONS

In this paper we have presented the basic formalism that we use for our likelihood analysis. This paper shows the final step on the path from time-ordered data to cosmological parameters. It provides the framework for the analysis of cosmological parameters and their implications for cosmology.

The unprecedented quality of the *WMAP* data and the tight constraints on cosmological parameters that are derived require a rigorous analysis so that the approximations made in the modeling do not propagate into significant biases and systematic errors. We have derived an approximation to the exact likelihood function for the $\mathscr{C}_\ell$ that is accurate to better than 0.1%, and we have carefully calibrated the temperature power spectrum covariance matrix with Monte Carlo simulations. This enables us to use the effective χ^2 per degree of freedom as a tool to test whether or not a model is an acceptable fit to the data.

We implement our likelihood analysis with the MCMC. We have concentrated on the issue of convergence and mixing, emphasizing how important these issues are in recovering cosmological parameter values and their confidence levels from the MCMC output.

To the *WMAP* data sets (TT and TE angular power spectra) we have added the CBI and ACBAR measurement of the CMB on smaller angular scales, the 2dFGRS galaxy power spectrum at $z \sim 0$, and the Lyα forest matter power spectrum at $z \sim 3$. These external data sets significantly enhance the scientific value of the *WMAP* measurement, by allowing us to break parameter degeneracies. While the underlying physics for these data sets is much more complicated and less well understood than for *WMAP* data, and systematic and instrumental effects are much more important, we feel we have made a significant step toward improving the rigor of the analysis of these data sets. We have included a detailed modeling of galaxy bias, redshift distortions, and the nonlinear growth of structure. We also include known (as to the present day) systematic and statistical uncertainties intrinsic to these other data sets.

We thank Bill Holzapfel for invaluable discussions about the ACBAR data. We thank the 2dFGRS team for giving us access to the Monte Carlo realizations of the 2dFGRS. The mock catalogs of the 2dFGRS were constructed at the Institute for Computational Cosmology at Durham. We thank Will Percival for discussions and for providing us with the covariance matrix of the 2dFGRS power spectrum. L. V. is supported by NASA through Chandra Fellowship PF2-30022 issued by the Chandra X-Ray Observatory Center, which is operated by the Smithsonian Astrophysical Observatory and on behalf of NASA under contract NAS 8-39073. L. V. also acknowledges Rutgers University for support during the initial stages of this work. H. V. P. is supported by a Dodds Fellowship granted by Princeton University. The *WMAP* mission is made possible by the support of the Office of Space Sciences at NASA Headquarters and by the hard and capable work of scores of scientists, engineers, technicians, machinists, data analysts, budget analysts, managers, administrative staff, and reviewers.

REFERENCES

Ballinger, W. E., Heavens, A. F., & Taylor, A. N. 1995, MNRAS, 276, L59
Ballinger, W. E., Peacock, J. A., & Heavens, A. F. 1996, MNRAS, 282, 877
Barnes, C., et al. 2003, ApJS, 148, 51
Baugh, C. M., & Efstathiou, G. 1993, MNRAS, 265, 145
Bennett, C. L., et al. 2003a, ApJ, 583, 1
Bennett, C. L., et al. 2003b, ApJS, 148, 97
Bond, J. R., Crittenden, R., Davis, R. L., Efstathiou, G., & Steinhardt, P. J. 1994, Phys. Rev. Lett., 72, 13
Bond, J. R., & Efstathiou, G. 1984, ApJ, 285, L45
Bond, J. R., Jaffe, A. H., & Knox, L. 1998, Phys. Rev. D, 57, 2117

Bond, J. R., Jaffe, A. H., & Knox, L. 2000, ApJ, 533, 19
Bond, J. R., et al. 2002, preprint (astro-ph/0205386)
Bridle, S. L., Crittenden, R., Melchiorri, A., Hobson, M. P., Kneissl, R., & Lasenby, A. N. 2002, MNRAS, 335, 1193
Christensen, N., & Meyer, R. 2000, preprint (astr-ph/0006401)
Christensen, N., Meyer, R., Knox, L., & Luey, B. 2001, Classical Quantum Gravity, 18, 2677
Cole, S., Fisher, K. B., & Weinberg, D. H. 1994, MNRAS, 267, 785
Colless, M., et al. 2001, MNRAS, 328, 1039
Croft, R. A. C., Weinberg, D. H., Bolte, M., Burles, S., Hernquist, L., Katz, N., Kirkman, D., & Tytler, D. 2002, ApJ, 581, 20
Davis, M., & Peebles, P. J. E. 1983, ApJ, 267, 465
Dressler, A. 1980, ApJ, 236, 351
Efstathiou, G., & Bond, J. R. 1999, MNRAS, 304, 75
Efstathiou, G., & Moody, S. J. 2001, MNRAS, 325, 1603
Elgarøy, Ø., et al. 2002, Phys. Rev. Lett., 89, 61301
Feldman, H. A., Kaiser, N., & Peacock, J. A. 1994, ApJ, 426, 23
Fisher, K. B. 1995, ApJ, 448, 494
Freedman, W. L., et al. 2001, ApJ, 553, 47
Fry, J. N. 1996, ApJ, 461, L65
Gelman, A., & Rubin, D. 1992, Stat. Sci., 7, 457
Gilks, W. R., Richardson, S., & Spiegelhalter, D. J. 1996, Markov Chain Monte Carlo in Practice (London: Chapman and Hall)
Gnedin, N. Y., & Hamilton, A. J. S. 2002, MNRAS, 334, 107 (GH)
Goldstein, J. H., et al. 2002, ApJ, submitted
Gunn, J. E., & Knapp, G. R. 1993, in ASP Conf. Ser. 43, Sky Surveys: Protostars to Protogalaxies, ed. B. T. Soifer (San Francisco: ASP), 267
Hamilton, A. J. S., Matthews, A., Kumar, P., & Lu, E. 1991, ApJ, 374, L1
Hatton, S., & Cole, S. 1998, MNRAS, 296, 10
Hawkins, E., et al. 2002, MNRAS, submitted
Hinshaw, G. F., et al. 2003a, ApJS, 148, 63
———. 2003b, ApJS, 148, 135
Hoekstra, H., van Waerbeke, L., Gladders, M. D., Mellier, Y., & Yee, H. K. C. 2002, ApJ, 577, 604
Hu, W. 2001, Phys. Rev. D, 64, 83005
Hubble, E. 1936, ApJ, 84, 517
Huey, G., Wang, L., Dave, R., Caldwell, R. R., & Steinhardt, P. J. 1999, Phys. Rev. D, 59, 63005
Jarosik, N., et al. 2003a, ApJS, 145, 413
———. 2003b, ApJS, 148, 29
Kaiser, N. 1984, ApJ, 284, L9
———. 1987, MNRAS, 227, 1
Kang, X., Jing, Y. P., Mo, H. J., & Börner, G. 2002, MNRAS, 336, 892
Kass, R. E., Carlin, B. P., Gelman, A., & Neal, R. M. 1997, J. Am. Stat. Assoc., 52, 93
Knox, L., Christensen, N., & Skordis, C. 2001, ApJ, 563, L95
Kogut, A., et al. 2003, ApJS, 148, 161
Komatsu, E., & Seljak, U. 2002, MNRAS, 336, 1256
Komatsu, E., et al. 2003, ApJS, 148, 119
Kosowsky, A., Milosavljevic, M., & Jimenez, R. 2002, Phys. Rev. D, 66, 63007
Kuo, C. L., et al. 2002, ApJ, submitted
Lahav, O., Rees, M. J., Lilje, P. B., & Primack, J. R. 1991, MNRAS, 251, 128
Lewis, A., & Bridle, S. 2002, Phys. Rev. D, 66, 103511
Ma, C., Caldwell, R. R., Bode, P., & Wang, L. 1999, ApJ, 521, L1
Maddox, S. J., Efstathiou, G., & Sutherland, W. J. 1990a, MNRAS, 246, 433
———. 1996, MNRAS, 283, 1227
Maddox, S. J., Efstathiou, G., Sutherland, W. J., & Loveday, J. 1990b, MNRAS, 243, 692
Marzke, R. O., Geller, M. J., da Costa, L. N., & Huchra, J. P. 1995, AJ, 110, 477
Mason, B. S., et al. 2002, ApJ, submitted
Matsubara, T. 1994, ApJ, 424, 30
McDonald, P., Miralda-Escudé, J., Rauch, M., Sargent, W. L. W., Barlow, T. A., Cen, R., & Ostriker, J. P. 2000, ApJ, 543, 1
Norberg, P., et al. 2001, MNRAS, 328, 64
Page, L., et al. 2003a, ApJ, 585, 566
Page, L., et al. 2003b, ApJS, 148, 39
Panter, B., Heavens, A. F., & Jimenez, R. 2002, MNRAS, submitted
Peacock, J. A., & Dodds, S. J. 1994, MNRAS, 267, 1020
———. 1996, MNRAS, 280, L19
Peacock, J. A., et al. 2001, Nature, 410, 169
Pearson, T. J., et al. 2002, ApJ, submitted
Peebles, P. J. E., & Yu, J. T. 1970, ApJ, 162, 815
Peiris, H., et al. 2003, ApJS, 148, 213
Percival, W. J. 2003, MNRAS, submitted
Percival, W. J., et al. 2001, MNRAS, 327, 1297
Postman, M., & Geller, M. J. 1984, ApJ, 281, 95
Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 1992, Numerical Recipes in C (2d ed.; Cambridge: Cambridge Univ. Press)
Riess, A. G., et al. 1998, AJ, 116, 1009
———. 2001, ApJ, 560, 49
Scoccimarro, R., Zaldarriaga, M., & Hui, L. 1999, ApJ, 527, 1
Seager, S., Sasselov, D. D., & Scott, D. 1999, ApJ, 523, L1
Seljak, U., & Zaldarriaga, M. 1996, ApJ, 469, 437
Seljak, U., et al. 2003, preprint (astro-ph/0302571)
Sievers, J. L., et al. 2002, preprint (astro-ph/0205387)
Spergel, D. N., et al. 2003, ApJS, 148, 175
Verde, L., et al. 2002, MNRAS, 335, 432
Zaldarriaga, M., Hui, L., & Tegmark, M. 2001, ApJ, 557, 519
Zaldarriaga, M., & Seljak, U. 2000, ApJS, 129, 431
Zurek, W. H., Quinn, P. J., Salmon, J. K., & Warren, M. S. 1994, ApJ, 431, 559