FIRST-YEAR WILKINSON MICROWAVE ANISOTROPY PROBE (WMAP)[1] OBSERVATIONS: THE ANGULAR POWER SPECTRUM

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

ABSTRACT

We present the angular power spectrum derived from the first-year Wilkinson Microwave Anisotropy Probe (WMAP) sky maps. We study a variety of power spectrum estimation methods and data combinations and demonstrate that the results are robust. The data are modestly contaminated by diffuse Galactic foreground emission, but we show that a simple Galactic template model is sufficient to remove the signal. Point sources produce a modest contamination in the low-frequency data. After masking $\sim$700 known bright sources from the maps, we estimate that residual sources contribute $\sim$3500 μK^2 at 41 GHz and $\sim$130 μK^2 at 94 GHz to the power spectrum $[\ell(\ell+1)C_\ell/2\pi]$ at $\ell = 1000$. Systematic errors are negligible compared to the (modest) level of foreground emission. Our best estimate of the power spectrum is derived from 28 cross-power spectra of statistically independent channels. The final spectrum is essentially independent of the noise properties of an individual radiometer. The resulting spectrum provides a definitive measurement of the cosmic microwave background (CMB) power spectrum, with uncertainties limited by cosmic variance, up to $\ell \sim 350$. The spectrum clearly exhibits a first acoustic peak at $\ell = 220$ and a second acoustic peak at $\ell \sim 540$ (Page and coworkers), and it provides strong support for adiabatic initial conditions (Spergel and coworkers). Kogut and coworkers analyze the $C_\ell^{\rm TE}$ power spectrum and present evidence for a relatively high optical depth and an early period of cosmic reionization. Among other things, this implies that the temperature power spectrum has been suppressed by $\sim$30% on degree angular scales, as a result of secondary scattering.

Subject headings: cosmic microwave background — cosmological parameters —
cosmology: observations — early universe — large-scale structure of universe —
space vehicles: instruments

1. INTRODUCTION

The Wilkinson Microwave Anisotropy Probe (WMAP) mission was designed to measure the cosmic microwave background (CMB) anisotropy with unprecedented precision and accuracy on angular scales from the full sky to several arcminutes by producing maps at five frequencies from 23 to 94 GHz. The WMAP satellite mission (Bennett et al. 2003a) employs a matched pair of 1.4 m telescopes (Page et al. 2003a) observing two areas on the sky separated by $\sim$141°. A differential radiometer (Jarosik et al. 2003a)

with a total of 10 feeds for each of the two sets of optics (Barnes et al. 2003; Page et al. 2003a) measures the difference in sky brightness between the two sky pixels. The satellite is deployed at the second Earth-Sun Lagrange point, L2, and observes the sky with a compound spin and precession that covers the full sky every 6 months. The differential data are processed on the ground to produce full-sky maps of the CMB anisotropy (Hinshaw et al. 2003).

Full-sky maps provide the smallest record of the CMB anisotropy without loss of information. They permit a wide variety of statistics to be computed from the data: one of the most fundamental is the angular power spectrum of the CMB. Indeed, if the temperature fluctuations are Gaussian, with random phase, then the angular power spectrum provides a complete description of the statistical properties of the CMB. Komatsu et al. (2003) have analyzed the first-year WMAP sky maps to search for evidence of non-Gaussianity and find none, aside from a modest level of point-source contamination that we account for in this paper. Thus, the measured power spectrum may be compared to predictions of cosmological models to develop constraints on model parameters.

This paper presents the angular power spectrum obtained from the first-year WMAP sky maps. Companion papers present the maps and an overview of the basic results (Bennett et al. 2003b) and describe the foreground removal process that precedes the power spectrum analysis (Bennett et al. 2003c). Spergel et al. (2003), Verde et al. (2003), Peiris et al. (2003), Kogut et al. (2003), and Page et al. (2003c)

[2] NASA Goddard Space Flight Center, Code 685, Greenbelt, MD 20771; gary.f.hinshaw@nasa.gov.
[3] Department of Astrophysical Sciences, Princeton University, Peyton Hall, Princeton, NJ 08544.
[4] Science Systems and Applications, Inc. (SSAI), 10210 Greenbelt Road, Suite 600, Lanham, MD 20706.
[5] Departments of Astrophysics and Physics, EFI, and CfCP, University of Chicago, 5640 South Ellis Avenue, Chicago, IL 60637.
[6] Department of Physics, Princeton University, Jadwin Hall, P.O. Box 708, Princeton, NJ 08544.
[7] Department of Physics and Astronomy, University of British Columbia, 6224 Agricultural Road, Vancouver, BC V6T 1Z1, Canada.
[8] National Research Council (NRC) Fellow.
[9] Department of Physics, Brown University, Providence, RI 02912.
[10] Department of Astronomy, UCLA, P.O. Box 951562, Los Angeles, CA 90095-1562.

discuss the implications of the *WMAP* power spectrum for cosmological parameters and carry out a joint analysis of the *WMAP* spectrum together with other CMB data and data from large-scale structure probes. Hinshaw et al. (2003), Jarosik et al. (2003b), Page et al. (2003b), Barnes et al. (2003), and Limon et al. (2003)[11] discuss the data processing, the radiometer performance, the instrument beam characteristics, and the spacecraft in-orbit performance, respectively.

A sky map $\Delta T(\boldsymbol{n})$ defined over the full sky can be decomposed in spherical harmonics

$$\Delta T(\boldsymbol{n}) = \sum_{\ell > 0} \sum_{m=-\ell}^{\ell} a_{\ell m} Y_{\ell m}(\boldsymbol{n}) \qquad (1)$$

with

$$a_{\ell m} = \int d\Omega_{\boldsymbol{n}} \, \Delta T(\boldsymbol{n}) \, Y_{\ell m}^*(\boldsymbol{n}) \, , \qquad (2)$$

where $\boldsymbol{n}$ is a unit direction vector and $Y_{\ell m}(\boldsymbol{n})$ are the spherical harmonic functions evaluated in the direction $\boldsymbol{n}$.

If the CMB temperature fluctuation ΔT is Gaussian distributed, then each $a_{\ell m}$ is an independent Gaussian deviate with

$$\langle a_{\ell m} \rangle = 0 \qquad (3)$$

and

$$\langle a_{\ell m} a_{\ell' m'}^* \rangle = \delta_{\ell \ell'} \delta_{m m'} C_\ell \, , \qquad (4)$$

where C_ℓ is the ensemble average power spectrum predicted by models and δ is the Kronecker symbol. The actual power spectrum realized in our sky is

$$C_\ell^{\rm sky} = \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} |a_{\ell m}|^2 \, . \qquad (5)$$

In the absence of noise, and with full-sky coverage, the right-hand side of equation (5) provides an unbiased estimate of the underlying theoretical power spectrum, which is limited only by cosmic variance. However, realistic CMB anisotropy measurements contain noise and other sources of error that cause the quadratic estimator in equation (5) to be biased. In addition, while *WMAP* measures the anisotropy over the full sky, the data near the Galactic plane are sufficiently contaminated by foreground emission that only a portion of the sky ($\sim$85%) can be used for CMB power spectrum estimation. Thus, the integral in equation (2) cannot be evaluated as such, and other methods must be found to estimate C_ℓ.

In Appendix A we review two methods that have appeared in the literature for estimating the angular power spectrum in the presence of instrument noise and sky cuts. The first (Hivon et al. 2002) is a quadratic estimator that evaluates equation (2) on the *cut* sky yielding a " pseudo–power spectrum " $\tilde{C}_\ell$. The ensemble average of this quantity is related to the true power spectrum, C_ℓ, by means of a mode coupling matrix $G_{\ell \ell'}$ (Hauser & Peebles 1973). The second method (Oh, Spergel, & Hinshaw 1999) uses a maxi-

mum likelihood approach optimized for fast evaluation with *WMAP*-like data. In Appendix A we demonstrate that the two methods produce consistent results.

A quadratic estimator offers the possibility of computing both an " auto-power " spectrum, proportional to $\sum_m |a_{\ell m}|^2$, and a " cross-power " spectrum, proportional to $\sum_m a_{\ell m}^i a_{\ell m}^{j*}$, where the $a_{\ell m}$ coefficients are estimated from two independent CMB maps, i and j. The latter form has the advantage that, if the noise in the two maps is uncorrelated, the quadratic estimator is not biased by the noise. For all cosmological analyses, we use only the cross-power spectra between statistically independent channels. As a result, the angular spectra are, for all intents and purposes, independent of the noise properties of an individual radiometer. This is analogous to interferometric data that exhibit a high degree of immunity to systematic errors. The precise form of the estimator we use is given in Appendix A.

The plan of this paper is as follows. In § 2 we review the properties of the *WMAP* instrument and how they affect the derived power spectrum. In § 3 we present results for the angular power spectra obtained from individual pairs of radiometers, the cross-power spectra, and examine numerous consistency tests. In § 4, in preparation for generating a final combined power spectrum, we present the full covariance matrix of the cross-power spectra. In § 5 we present the methodology used to produce the final combined power spectrum and its covariance matrix. In § 6 we compare the *WMAP* first-year power spectrum to a compilation of previous CMB measurements and to a prediction based on a combination of previous CMB data and the Two-Degree Field Galaxy Redshift Survey (2dFGRS) data. We summarize our results in § 7 and outline the power spectrum data products being made available through the Legacy Archive for Microwave Background Data Analysis (LAMBDA). Appendix A reviews two methods used to estimate the angular power spectrum from CMB maps. Appendix B describes how we account for point-source contamination. Appendix C presents our approach to combining multichannel data. Appendix D describes how the foreground mask correlates multipole moments in the Fisher matrix, and Appendix E collects some useful properties of the spherical harmonics.

2. INSTRUMENTAL PROPERTIES

The *WMAP* instrument is composed of 10 " differencing assemblies " (DAs) spanning five frequencies from 23 to 94 GHz (Bennett et al. 2003a). The two lowest frequency bands (K and Ka) are primarily Galactic foreground monitors, while the three highest (Q, V, and W) are primarily cosmological bands. There are eight high-frequency DAs: Q1, Q2, V1, V2, and W1–W4. Each DA is formed from two differential radiometers that are sensitive to orthogonal linear polarization modes; the radiometers are designated 1 or 2 (e.g., V11 or W12) depending on which polarization mode is being sensed.

The temperature measured on the sky is modified by the properties of the instrument. The most important properties that affect the angular power spectrum are finite resolution and instrument noise. Let $C_\ell^{ii'}$ denote the auto- or cross-power spectrum evaluated from two sky maps, i and i', where i is a DA index. Further, define the shorthand $\boldsymbol{i} \equiv (i, i')$ to denote a pair of indices, e.g., (Q1, V2). This

spectrum will have the form

$$C_\ell^i = w_\ell^i C_\ell^{\mathrm{sky}} + N_\ell^i \,, \qquad (6)$$

where $w_\ell^i \equiv b_\ell^i b_\ell^{i'} p_\ell^2$ is the window function that describes the combined smoothing effects of the beam and the finite sky map pixel size. Here b_ℓ^i is the beam transfer function for DA i, given by Page et al. (2003b) [note that they reserve the term "beam window function" for $(b_\ell^i)^2$], and p_ℓ is the pixel transfer function supplied with the HEALPix package (Górski et al. 1999). N_ℓ^i is the noise spectrum realized in this particular measurement. On average, the observed spectrum estimates the underlying power spectrum, C_ℓ,

$$\langle C_\ell^i \rangle = w_\ell^i C_\ell + \langle N_\ell^i \rangle \delta_{ii'} \,, \qquad (7)$$

where $\langle N_\ell^i \rangle$ is the average noise power spectrum for DA i and the Kronecker symbol indicates that the noise is uncorrelated between DAs. To estimate the underlying power spectrum on the sky, C_ℓ, the effects of the noise bias and beam convolution must be removed. The determination of transfer functions and noise properties is thus a critical component of any CMB experiment.

In § 2.1 we summarize the results of Page et al. (2003b) on the *WMAP* window functions and their uncertainties. We propagate these uncertainties through to the final Fisher matrix for the angular power spectrum. In § 2.2 we present a model of the *WMAP* noise properties appropriate to power spectrum evaluation. For cross-power spectra ($i \neq i'$ above), the noise bias term drops out of equation (7) if the noise between the two DAs is uncorrelated. These cross-power spectra provide a nearly optimal estimate of the true power spectrum, essentially independent of errors in the noise model; thus, we use them exclusively in our final power spectrum estimate. The noise model is primarily used to propagate noise errors through the analysis and to test a variety of different power spectrum estimates for consistency with the combined cross-power spectrum.

2.1. *Window Functions*

As discussed in Page et al. (2003b), the instrument beam response was mapped in flight using observations of the planet Jupiter. The signal-to-noise ratio is such that the response, relative to the peak of the beam, is measured to approximately -35 dB in W band, the band with the highest angular resolution. The beam widths, measured in flight, range from $0°.82$ at K band down to $0°.20$ in some of the W-band channels (FWHM). Maps of the full two-dimensional beam response are presented in Page et al. (2003b) and are available with the *WMAP* first-year data release. The radial beam profiles obtained from these maps have been fitted to a model consisting of a sum of Hermite polynomials that accurately characterize the main Gaussian lobe and small deviations from it. The model profiles are then Legendre transformed to obtain the beam transfer functions b_ℓ^i for each DA i. Full details of this procedure are presented in Page et al. (2003b), and the resulting transfer functions are also provided in the first-year data release. We have chosen to normalize the transfer function to 1 at $\ell = 1$ because *WMAP* calibrates its intensity response using the modulation of the CMB dipole ($\ell = 1$). This effectively partitions calibration uncertainty from window function uncertainties.

The beam processing described above provides a straightforward means of propagating the noise uncertainty directly from the time-ordered data through to the final transfer functions. The result is the covariance matrix $\Sigma_{b,\ell\ell'}^i$ for the normalized transfer function. Plots of the diagonal elements of $\Sigma_{b,\ell\ell'}^i$ are presented in Page et al. (2003b). The fractional error in the transfer functions b_ℓ^i is typically 1%–2% in amplitude. In the end, these window function uncertainties dominate the small off-diagonal elements of the final covariance matrix for the combined power spectrum (see § 4). Additional observations of Jupiter will reduce these uncertainties.

An additional source of error in our treatment of the beam response arises from noncircularity of the main beam. The effects of this noncircularity are mitigated by *WMAP*'s scan strategy, which results in most sky pixels being observed over a wide range of azimuth angles. The effective beam response on the sky is thus largely symmetrized. We estimate that the effects of imperfect symmetrization produce window function errors of less than 1% relative to a perfectly symmetrized beam window function (Page et al. 2003b; Hinshaw et al. 2003). This error is well within the formal uncertainty given in $\Sigma_{b,\ell\ell'}^i$. In § 5 we infer the optimal power spectrum by combining the 28 cross-power spectra measured by the eight high-frequency DAs Q1–W4. As part of this process we marginalize over the window function uncertainty, which automatically propagates these errors into the final covariance matrix for the combined power spectrum. Both the combined power spectrum and the corresponding Fisher matrix are part of the first-year data release.

2.2. *Instrument Noise Properties*

The noise bias term in equation (7) is the noise per $a_{\ell m}$ mode on the sky. If auto-power spectra are used in the final power spectrum estimate, the noise bias term must be known very accurately because it exponentially dominates the convolved power spectrum at high ℓ. If only cross-correlations are used, the noise bias is only required for propagating errors. Our final best spectrum is based only on cross-correlations and is independent of this term. However, as an independent check of our results, we evaluate the maximum likelihood spectrum based on a combined Q+V+W sky map. The noise bias must be estimated accurately for this application.

In the limit that the time-ordered instrument noise is white, the noise bias is a constant, independent of ℓ. If the time-ordered noise has a $1/f$ component, the bias term will rise at low ℓ. In this subsection we estimate the noise bias properties for each of the high-frequency *WMAP* radiometers based on the time-ordered noise properties presented in Jarosik et al. (2003b). While the *WMAP* radiometer noise is nearly white by design (Jarosik et al. 2003a) with $1/f$ knee frequencies of less than 10 mHz for nine out of 10 DAs, one of the radiometers (W41) has a $1/f$ knee frequency of $\sim$45 mHz. The latter is large enough that the deviations of $\langle N_\ell^i \rangle$ from a constant must be accounted for.

The most reliable way to estimate the effects of $1/f$ noise on the measured power spectra is by Monte Carlo simulation. Using the pipeline simulator discussed in Hinshaw et al. (2003), we have generated a library of noise maps with flightlike properties. Specifically, we have included flightlike $1/f$ noise in the simulated time-ordered data and have run

TABLE 1
WMAP Power Spectrum Noise Model

Parameter	Q1	Q2	V1	V2	W1	W2	W3	W4
$\ell < 200$								
$c_0 \, (\times 10^2)$	-6.604	-6.971	-6.823	-6.674	-7.174	-7.439	-7.044	-7.989
$c_1 \, (\times 10^3)$	1.934	4.244	2.497	1.395	3.060	3.721	1.459	4.820
$c_2 \, (\times 10^4)$	-3.798	-8.521	-5.203	-1.483	-6.021	-6.901	-2.677	-5.870
$c_3 \, (\times 10^5)$	2.549	5.921	3.715	0.242	4.147	4.540	1.750	2.167
$200 < \ell < 450$								
$c_0 \, (\times 10^2)$	-6.143	-6.268	-6.320	-6.192	-6.618	-6.833	-6.768	-6.849
$c_1 \, (\times 10^4)$	-1.334	1.341	-0.800	-1.351	0.769	2.582	0.885	2.443
$\ell > 450$								
$c_0 \, (\times 10^2)$	-6.202	-6.167	-6.338	-6.264	-6.564	-6.704	-6.688	-6.772
$c_1 \, (\times 10^5)$	3.288	4.085	3.327	5.979	7.342	12.970	8.851	24.430

Note.—Best-fit coefficients for the noise model in eq. (8). The units of the output noise are mK^2, thermodynamic temperature.

each full-year realization through the map-making pipeline, including the baseline prewhitening discussed in Hinshaw et al. (2003). We evaluate the power spectra of these maps using the quadratic estimator described in Appendix A with three different pixel weighting schemes (see § A1.2 for definitions of the weights and the ℓ range in which each is used). We define the effective noise as a function of ℓ based on fits to these Monte Carlo noise spectra. For the analyses in this paper, we fit the spectra to a model of the form

$$\ln \langle N_\ell^i \rangle = \left[\sum_{n=0}^{n_{\max}} c_n^i (\ln \ell)^n \right]^{-1} , \qquad (8)$$

where the c_n^i are fit coefficients given in Table 1, with $n_{\max} = 3$ for $\ell < 200$ and $n_{\max} = 1$ for $\ell > 200$.

Figure 1 shows the noise spectrum derived from the simulations for each of the eight high-frequency DAs, using uniform weighting over the entire ℓ range. For comparison, we also plot an estimate of the CMB power spectrum from § 5 in gray. Note that the W4 spectrum is the only one of this set to exhibit deviations from white noise in an ℓ range where the signal-to-noise ratio is relatively low, and we believe that this simulation slightly overestimates the $1/f$ noise in the flight W4 DA (Hinshaw et al. 2003).

2.3. *Systematic Errors*

Hinshaw et al. (2003) present limits on systematic errors in the first-year sky maps. They consider the effects of absolute and relative calibration errors, artifacts induced by environmental disturbances (thermal and electrical), errors from the map-making process, pointing errors, and other miscellaneous effects. The combined errors due to relative calibration errors, environmental effects, and map-making errors are limited to less than 15 μK^2 (2 σ) in the quadrupole moment C_2 in any of the eight high-frequency DAs. Tighter limits are placed on higher order moments. We conservatively estimate the absolute calibration uncertainty in the first-year *WMAP* data to be 0.5%.

Random pointing errors are accounted for in the beam mapping procedure; the beam transfer functions presented by Page et al. (2003b) incorporate random pointing errors automatically. A systematic pointing error of $\sim 10''$ at the

spin period is suspected in the quaternion solution that defines the spacecraft pointing. This is much smaller than the smallest beam width ($\sim 12'$ at W band), and we estimate that it would produce less than 1% error in the angular power spectrum at $\ell = 1000$; thus, we do not attempt to correct for this effect. Barnes et al. (2003) place limits on spurious contributions due to stray light pickup through the far sidelobes of the instrument. As shown in Figure 4 of that paper, they place limits of less than 10 μK^2 on spurious contributions to $\ell(\ell+1)C_\ell/2\pi$, at Q–W band, due to far sidelobe pickup.

A detailed model of Galactic foreground emission based on the first-year *WMAP* data is presented by Bennett et al. (2003c) and is summarized in § 3.1. We show that diffuse foreground emission is a modest source of contamination at large angular scales ($\gtrsim 2°$). *Systematic errors on these angular scales are negligible compared to the (modest) level of foreground emission.* On smaller angular scales ($\lesssim 2°$), the 1%–3% uncertainty in the individual beam transfer functions is the largest source of uncertainty, while for multipole moments greater than ~ 600, random white noise from the instrument is the largest source of uncertainty.

3. THE DATA

Figure 2 shows the cross-power spectra obtained from all 28 combinations of the eight DAs Q1–W4 using the quadratic estimator described in § A1. These spectra have been evaluated with the Kp2 sky cut described in Bennett et al. (2003c). The spectra are color coded by effective frequency, $\sqrt{\nu^i \nu^{i'}}$, where ν^i is the frequency of DA i. The low-frequency (41 GHz) data are shown in red, the high-frequency (94 GHz) data in blue, with intermediate frequencies following the colors of the rainbow. The top panel shows $\ell(\ell+1)C_\ell/2\pi$ in μK^2, while the bottom panel plots the ratio of each channel to our final combined spectrum presented in § 5. The top panel shows a very robust measurement of the first acoustic peak with a maximum near $\ell \sim 220$ and a shape that is consistent with the predictions of adiabatic fluctuation models. There is also a clear indication of the rise to a second peak at $\ell \sim 540$. See Page et al. (2003c) for
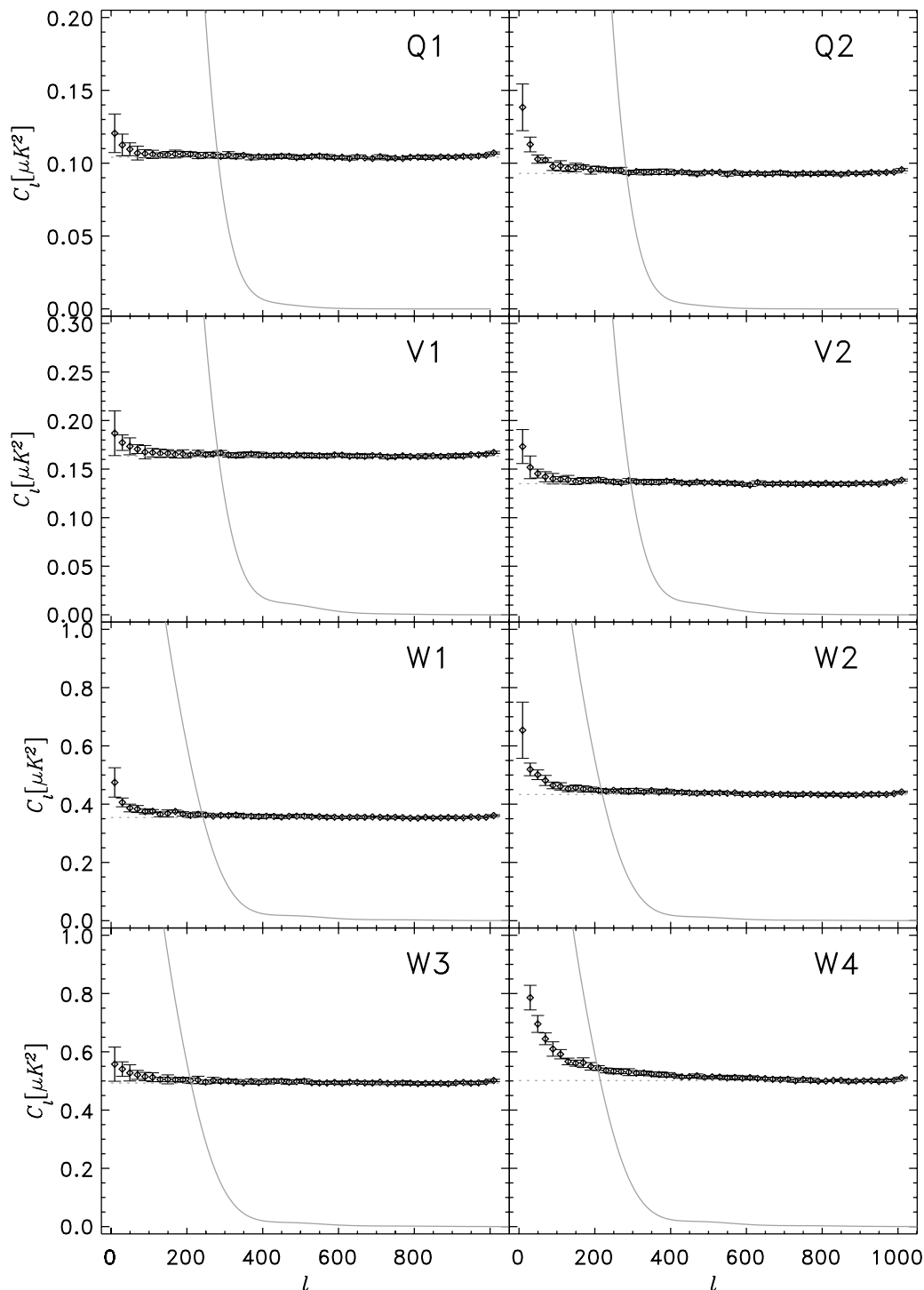
Fig. 1.—Effective noise as a function of ℓ for the eight DAs used in the combined power spectrum analysis. These spectra were computed from end-end simulations of noise maps, as discussed in § 2.2. For illustration, the spectra shown here were computed using the quadratic estimator with uniform pixel weights. The actual noise model uses three separate weighting schemes in three separate ℓ ranges and thus has discontinuities where the effective noise level changes. The weights are defined in § A1.2.

an analysis and discussion of the peaks and troughs in the first-year *WMAP* power spectrum.

The red data in the top panel show very clearly that the low-frequency data are contaminated by diffuse Galactic emission at low ℓ and by point sources at higher ℓ. The higher frequency data show less apparent contamination, consistent with the foreground emission being dominated

by radio emission, rather than thermal dust emission, as expected in this frequency range.

3.1. *Galactic and Extragalactic Foregrounds*

Bennett et al. (2003c) present a detailed model of the Galactic foreground emission based on a maximum entropy analysis of all five *WMAP* frequency bands, in combination
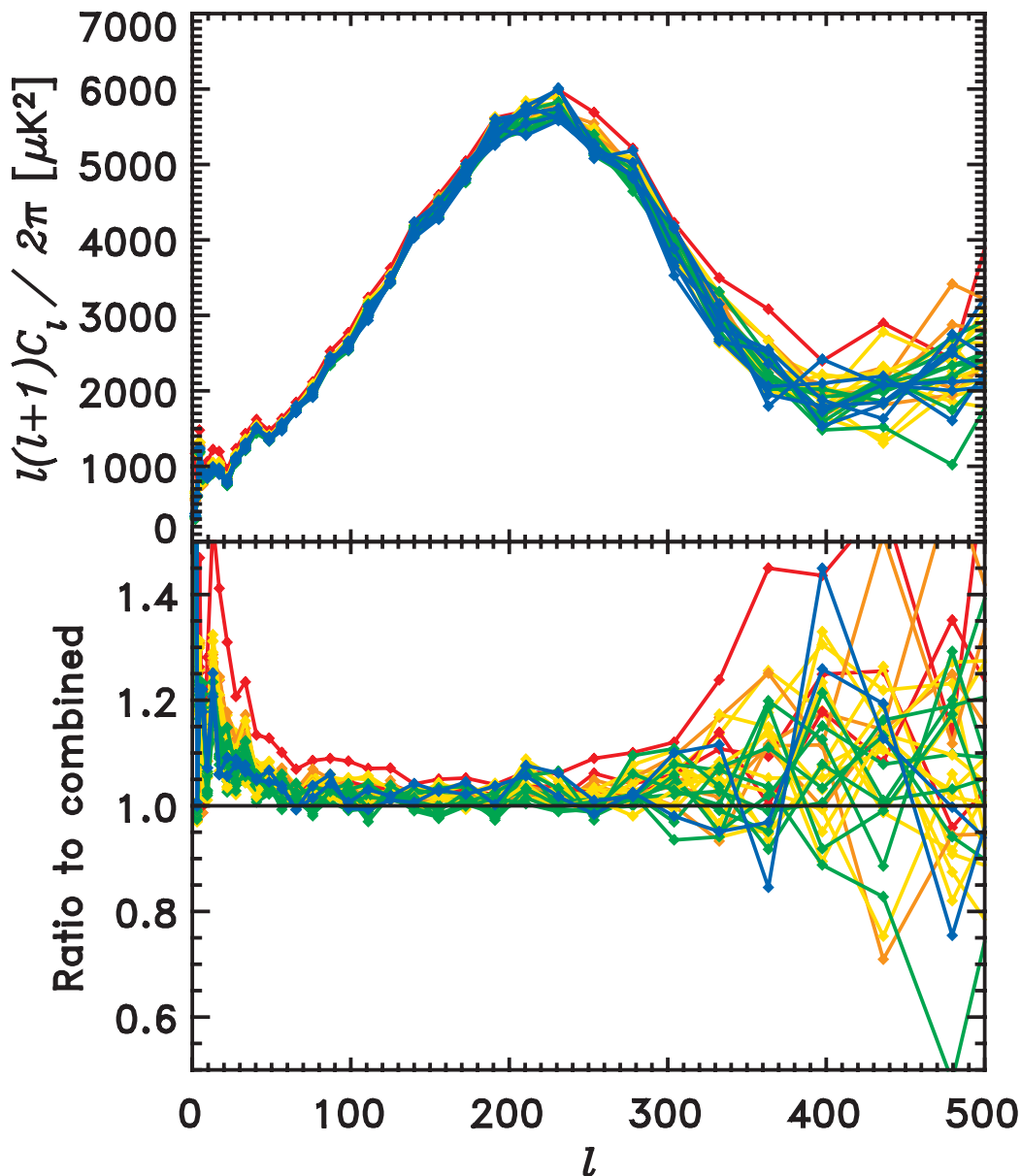
FIG. 2.—Full set of individual cross-power spectra for $\ell < 500$, computed from the eight high-frequency DAs Q1–W4, 28 spectra in all. The spectra were evaluated from the uncorrected sky maps (no Galaxy model subtracted) using the Kp2 sky cut with uniform weighting. The data are plotted in color by effective frequency $\sqrt{\nu_i \nu_j}$ with red corresponding to 41 GHz and blue to 94 GHz. The top panel shows a very robust measurement of the first acoustic peak with a maximum near $\ell \sim 220$. There is also a clear indication of the rise to a second peak at $\ell \sim 540$ as discussed in § 6. The bottom panel shows the ratio of each channel to the combined spectrum presented in § 5. This clearly shows the residual foreground emission due to diffuse Galactic radio emission at low ℓ and to point sources at higher ℓ. The level of contamination, which is strongest at Q band, is consistent with the expected level of foreground emission. See Fig. 3 for the spectra after foreground subtraction.

with external tracer templates. They demonstrate that the emission is well modeled by three distinct emission components: (1) synchrotron emission from cosmic-ray electrons, with a steeply falling spectrum in the *WMAP* frequency range: $T_A(\nu) \propto \nu^\beta$ with $\beta < -3$, steepening with increasing frequency; (2) free-free emission from the ionized interstellar medium that is well traced by Hα emission in regions where the dust extinction is low; and (3) thermal emission from interstellar dust grains with an emissivity index $\beta \sim 2.2$. The model has a Galactic signal minimum between V and W band.

In principle, we could subtract the above model from each *WMAP* channel and recompute the power spectrum.

However, since the model is based on *WMAP* data that have been smoothed to an angular resolution of 1°.0, the resulting maps would have complicated noise properties. For the purposes of power spectrum analysis, we adopt a more straightforward approach based on fitting foreground tracer templates to the Q-, V-, and W-band data. The details of this procedure, the resulting fit coefficients, and a comparison of the fits to the maximum entropy model are given in Bennett et al. (2003c). They estimate that the template model removes ~85% of the foreground emission in Q, V, and W bands and that the remaining emission constitutes less than ~2% of the CMB variance (up to $\ell = 200$) in Q band and $\lesssim 1\%$ of the CMB variance in V and W bands.

The contribution from extragalactic radio sources has been analyzed in three separate ways. Bennett et al. (2003c) directly fit for sources in the sky maps. The result of this analysis is that we have identified 208 sources in the *WMAP* data with sufficient signal-to-noise ratio to pass the detection criterion (we estimate that roughly five of these are likely to be spurious). The derived source count law is consistent with the following model for the power spectrum of the unresolved sources:

$$C_\ell^{i,\mathrm{src}} = A \left(\frac{\nu_i}{\nu_0}\right)^\beta \left(\frac{\nu_{i'}}{\nu_0}\right)^\beta w_\ell^i , \qquad (9)$$

with $A = 0.015$ μK^2 sr (measured in thermodynamic temperature), $\beta = -2.0$, and $\nu_0 \equiv 45$ GHz. Komatsu et al. (2003) evaluate the bispectrum of the *WMAP* data and are able to fit a non-Gaussian source component based on a particular configuration of the bispectrum. They find that the same source model, equation (9), fits the bispectrum data. For the remainder of this section, we adopt this model for correcting the cross-power spectra. At Q band (41 GHz) the correction to $\ell(\ell+1)C_\ell/2\pi$ is 868 and 3468 μK^2 at $\ell = 500$ and 1000, respectively. At W band (94 GHz), the correction is only 31 and 126 μK^2 at the same ℓ-values. For comparison, the CMB power in this ℓ range is $\sim$2000 μK^2. Later, when we derive a final combined spectrum from the multifrequency data, we adopt equation (9) as a model with A as a free parameter. We simultaneously fit for a combined CMB spectrum and source amplitude and marginalize over the residual uncertainty in A. The best-fit source amplitude from this process is consistent with the other two methods.

Figure 3 shows the cross-power spectra obtained from the same 28 combinations as in Figure 2, but this time with the Galactic template model and source model subtracted. The bottom panel of the figure shows the ratio of the 28 channels to the combined spectrum obtained in § 5. The 28 cross-power spectra are consistent with each other at the 5%–20% level over the ℓ range 2–500. Similar scatter is seen in Monte Carlo simulations of an ensemble of 28 cross-power spectra with *WMAP*'s beam and noise properties. The only significant deviation lies in the Q-band data at low ℓ, which are $\sim$10% below the higher frequency bands at $\ell < 20$. This is consistent with the accuracy estimated above for the Galactic template model (see also Fig. 11 of Bennett et al. 2003c for images of the maps after Galactic template subtraction). Since the *WMAP* data are not noise limited at low ℓ, we use only V- and W-band data in the final combined spectrum for $\ell < 100$.

The subtraction of the source model, equation (9), brings the Q-band spectrum into good agreement with the other cross-power spectra up to $\ell \sim 500$. At higher ℓ, the Q-band data contribute very little to the final combined spectrum because the (normalized) Q-band window function has dropped to less than 5% (Page et al. 2003b). As discussed in § 5, we marginalize over the source amplitude uncertainty, δA, when obtaining the final power spectrum estimate and associated covariance matrix. Thus, the uncertainty is also accounted for in subsequent cosmological parameter fits (Verde et al. 2003; Spergel et al. 2003; Peiris et al. 2003).

Figure 4 shows a close-up of the 28 cross-power spectra in Figure 3 up to $\ell = 100$. The top panel shows the raw (unbinned) data, which have correlations of less than 2% between neighboring points in this ℓ range (see § 4).

These data are strikingly consistent with each other and support the conclusion that systematic errors at low ℓ are insignificant. To assess the level of scatter that does exist between the spectra, we have generated a Monte Carlo simulation in which we compute the rms scatter among the 28 spectra at each ℓ, relative to the measured power. The bottom panel of Figure 4 shows the results of this simulation, averaged over 1000 realizations, compared to the relative rms scatter in the data. The agreement is excellent, indicating that the uncertainty in the measured power spectrum in this ℓ range is a few percent and is consistent with a combination of instrument noise and mode coupling due to the 15% sky cut.

Another striking feature is the low amplitude of the observed quadrupole and the sharp rise in power, almost linear in ℓ, to $\ell = 5$. Bennett et al. (2003b) quote a value for the rms quadrupole amplitude, $Q_{\mathrm{rms}} = \sqrt{(5/4\pi)C_2} = 8 \pm 2$ μK, where the uncertainty is largely due to Galactic model uncertainty. This is consistent with the amplitude measured by the *COBE* Differential Microwave Radiometer (DMR) experiment, $Q_{\mathrm{rms}} = 10^{+7}_{-4}$ μK (Bennett et al. 1996). The fast, nearly linear rise to $\ell = 5$ produces an angular correlation function with essentially no power on angular scales $\gtrsim 60°$, again in excellent agreement with the *COBE* DMR correlation function (Bennett et al. 2003b; Hinshaw et al. 1996). In the context of a standard ΛCDM model, the probability of observing this little power on scales greater than $60°$ is $\sim 2 \times 10^{-3}$ (Bennett et al. 2003b).

4. THE FULL COVARIANCE MATRIX

In § 5 we derive our best estimate of the angular power spectrum by optimally combining the 28 cross-power spectra discussed above. The procedure for combining spectra requires the full covariance matrix of the individual cross-power spectra: in this section we outline the salient features of this matrix. There are six principal sources of variance for the measured spectra, C_ℓ^i: cosmic variance, instrument noise, mode coupling due to the foreground mask, point-source subtraction errors, uncertainty in the beam window functions, and an overall calibration uncertainty. We ignore uncertainties in the diffuse foreground correction since they are everywhere subdominant to the cosmic variance uncertainty (see § 3.1).

We may write the covariance matrix as

$$(\Sigma_{\mathrm{full}})_{\ell\ell'}^{ij} = \left\langle \left[C_\ell^i - (C_\ell + AS^i)w_\ell^i \right] \left[C_{\ell'}^j - (C_{\ell'} + AS^j)w_{\ell'}^j \right] \right\rangle , \qquad (10)$$

where the angle brackets represent an ensemble average, C_ℓ is the true underlying power spectrum, w_ℓ^i is the window function of spectrum i, and AS^i is the point-source contribution to spectrum i. Here we have defined a point-source spectral function, S^i, as

$$S^i \equiv \left(\frac{\nu_i}{\nu_0}\right)^\beta \left(\frac{\nu_{i'}}{\nu_0}\right)^\beta , \qquad (11)$$

where ν_i, ν_0, and β are as defined after equation (9). Note that $(\Sigma_{\mathrm{full}})_{\ell\ell'}^{ij}$ is symmetric in both $(\ell\ell')$ and (ij).

In the process of forming the combined spectrum we estimate a best-fit point-source amplitude, $\bar{A}$, and subtract the corresponding source contribution from each spectrum i.
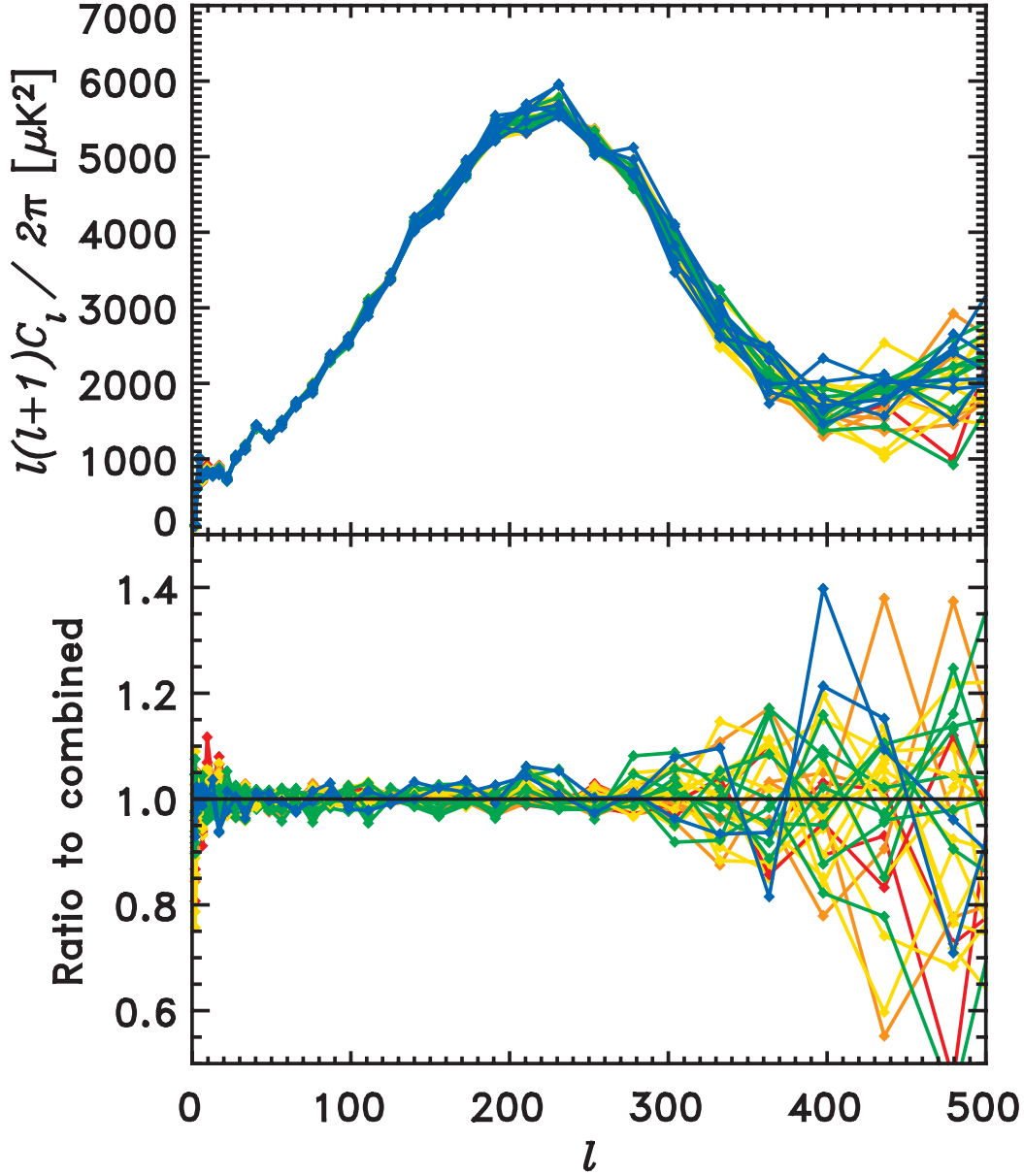
Fig. 3.—Same set of cross-power spectra as shown in Fig. 2. Here the foreground model discussed in § 3.1 has been subtracted from each channel. The bottom panel shows the ratio of each of the 28 cross-power spectra to the combined spectrum presented in § 5. Aside from a ∼10% discrepancy in the Q-band data at $\ell < 20$, the data are consistent with each other to the sensitivity limits of the individual spectra. Because the *WMAP* data are not sensitivity limited at low ℓ, we use only V- and W-band data in the final combined spectrum for $\ell < 100$ (see § 5) to minimize residual Galactic contamination.

We thus rewrite Σ_{full} as

$$(\Sigma_{\text{full}})^{ij}_{\ell\ell'} = \left\langle \left[\bar{C}^i_\ell - (C_\ell + \delta A S^i) w^i_\ell \right] \left[\bar{C}^j_{\ell'} - (C_{\ell'} + \delta A S^j) w^j_{\ell'} \right] \right\rangle , \quad (12)$$

where $\bar{C}^i_\ell \equiv C^i_\ell - \bar{A} S^i w^i_\ell$ is the source-subtracted spectrum and $\delta A \equiv A - \bar{A}$ is the residual source amplitude, which we marginalize over as a nuisance parameter.

We may expand the covariance matrix as

$$\Sigma_{\text{full}} = \Sigma_{\text{cv}} + \Sigma_{\text{mask}} + \Sigma_{\text{src}} + \Sigma_{\text{win}} . \quad (13)$$

We discuss each of these contributions in more detail below.

Cosmic variance, noise, and mode coupling.—The first two terms, $\Sigma_{\text{cv}} + \Sigma_{\text{mask}}$, incorporate the combined uncertainty

due to cosmic variance, instrument noise, and mode coupling due to the foreground mask,

$$(\Sigma_{\text{cv}} + \Sigma_{\text{mask}})^{ij}_{\ell\ell'} = \left\langle (\bar{C}^i_\ell - C_\ell w^i_\ell)(\bar{C}^j_{\ell'} - C_{\ell'} w^j_{\ell'}) \right\rangle , \quad (14)$$

where w^i_ℓ is fixed at its measured value. We have split this contribution into two pieces to mimic the procedure we actually use to compute the covariance matrix. As outlined in § 5.1, we start with Σ_{cv} and then incorporate the effects of point-source error and window function error. We do not add the effects of mode coupling, Σ_{mask}, until the very end of the computation. We consider this term in more detail in Appendix D.

Point-source subtraction errors.—The third term, Σ_{src}, is due to uncertainty in the point-source amplitude
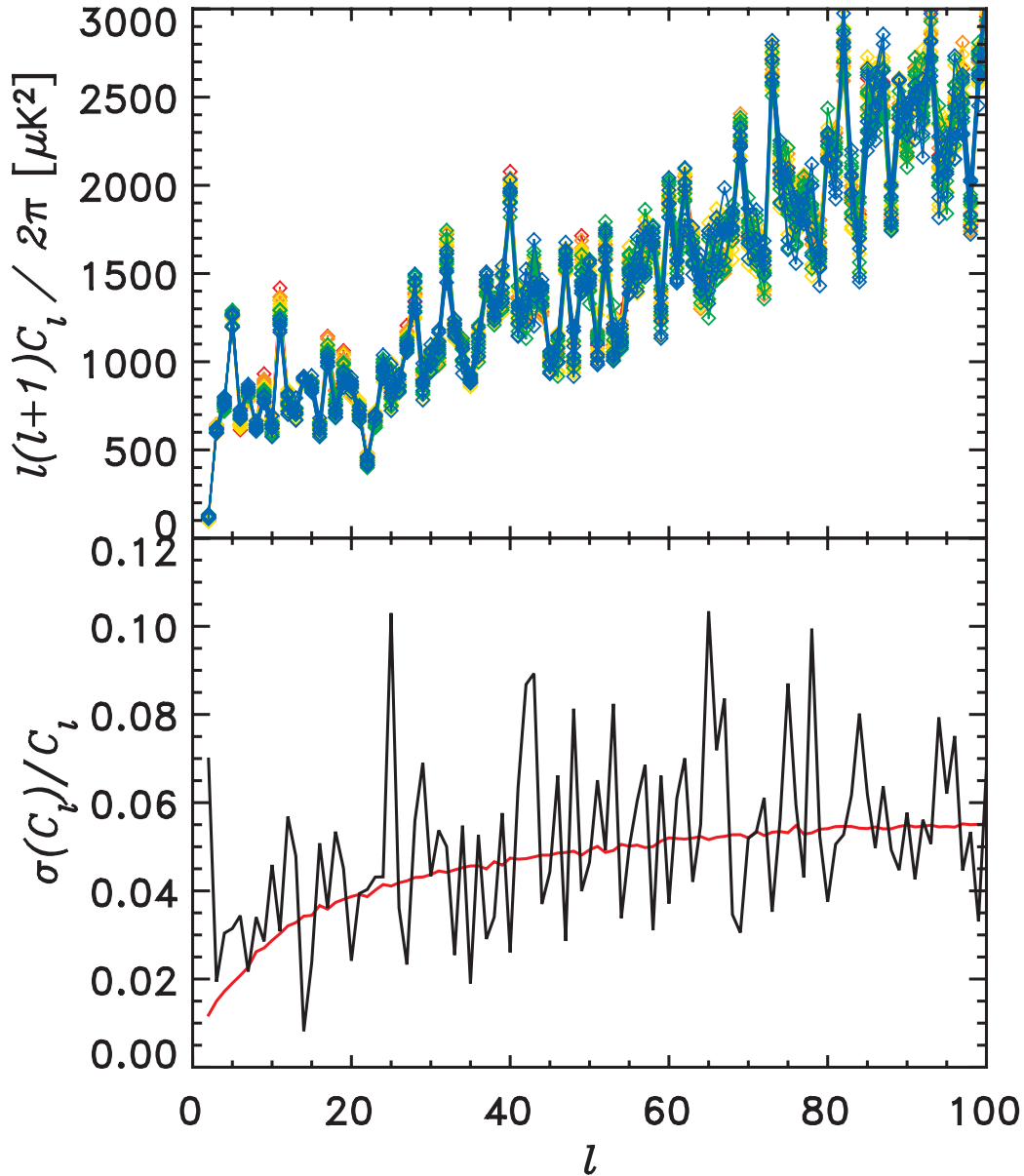
Fig. 4.—Same set of cross-power spectra as shown in Fig. 3, but showing the low-ℓ spectrum unbinned. The agreement between the individual spectra is striking. The low value of the quadrupole moment, C_2, that was first seen by *COBE* DMR is also seen in the *WMAP* data. The steep, nearly linear rise in the spectrum from $\ell = 2$ to 5 translates to a near absence of power in the angular correlation function at separations larger than $\sim 60°$ (Spergel et al. 2003; Bennett et al. 2003b). This was also seen in the *COBE* DMR data, but it is now clear that this is not due to Galaxy modeling errors. The bottom panel shows the fractional rms among the 28 *WMAP* cross-power spectra in black, while the red curve shows the same statistic averaged over an ensemble of 1000 Monte Carlo realizations. Based on this, we estimate the measurement error on the combined spectrum to be less than 2%–5% for $\ell < 100$.

determination

$$(\Sigma_{\rm src})_{\ell\ell'}^{ij} = S^i w_\ell^i \sigma_{\rm src}^2 w_{\ell'}^j S^j , \qquad (15)$$

where $\sigma_{\rm src}^2 \equiv \langle (A - \bar{A})^2 \rangle$ is the variance in the best-fit amplitude $\bar{A}$ and we assume that the frequency dependence, S^i, is perfectly known. In practice, we do not explicitly evaluate $\Sigma_{\rm src}$ as given above; rather, we employ a method based on marginalizing a Gaussian likelihood function, $\mathscr{L}(C_\ell^i, A)$, over a nuisance parameter A. This process, which is discussed in Appendix B, effectively yields $\Sigma_{\rm cv} + \Sigma_{\rm src}$.

Window function and calibration errors.—The fourth term, $\Sigma_{\rm win}$, is due to uncertainty in the beam window function, w_ℓ^i. This term arises from fluctuations in the window

function that cause the measured spectrum, C_ℓ^i, to differ from our estimate of the convolved spectrum, $\tilde{C}_\ell w_\ell^i$, where w_ℓ^i is the estimated window function. This contribution has the form

$$(\Sigma_{\rm win})_{\ell\ell'}^{ij} = C_\ell \langle \Delta w_\ell^i \Delta w_{\ell'}^j \rangle C_{\ell'} . \qquad (16)$$

Recall from § 2 that $w_\ell^i = b_\ell^i b_\ell^{i'} p_\ell^2$, where b_ℓ^i is the beam transfer function for DA i and p_ℓ^2 is the pixel window function. Define $u_\ell^i \equiv \Delta b_\ell^i / b_\ell^i$ to be the fractional error in b_ℓ^i, and then to first order in u_ℓ^i we have

$$\langle \Delta w_\ell^i \Delta w_{\ell'}^j \rangle = w_\ell^i \langle u_\ell^i u_{\ell'}^i + u_\ell^i u_{\ell'}^{i'} + u_\ell^{i'} u_{\ell'}^i + u_\ell^{i'} u_{\ell'}^{i'} \rangle w_{\ell'}^j . \qquad (17)$$

For *WMAP* the uncertainty in b_ℓ^i is uncorrelated between

DAs; thus, the above expression reduces to

$$\langle \Delta w_\ell^i \Delta w_{\ell'}^j \rangle = w_\ell^i \left[\Sigma_{u,\ell\ell'}^i (\delta_{ij} + \delta_{ij'}) + \Sigma_{u,\ell\ell'}^{i'} (\delta_{i'j} + \delta_{i'j'}) \right] w_{\ell'}^j , \qquad (18)$$

where we have defined

$$\Sigma_{u,\ell\ell'}^i \equiv \langle u_\ell^i u_{\ell'}^i \rangle = (b_\ell^i b_{\ell'}^i)^{-1} \Sigma_{b,\ell\ell'}^i , \qquad (19)$$

and where $\Sigma_{b,\ell\ell'}^i$ is the covariance matrix of the beam transfer function for DA i given by Page et al. (2003b). When generating the combined spectrum in the next section, we add Σ_{win} to the above contributions, giving $\Sigma_{\text{cv}} + \Sigma_{\text{src}} + \Sigma_{\text{win}}$.

The *WMAP* absolute calibration uncertainty is 0.5%. We do not explicitly incorporate this contribution in the covariance matrix. Instead, we propagate a 0.5% uncertainty into the normalization of the final power spectrum amplitude (Spergel et al. 2003).

5. THE COMBINED POWER SPECTRUM

In § 3 we demonstrated that the three high-frequency bands of *WMAP* data produced consistent estimates of the angular power spectrum, after a modest correction for diffuse Galactic emission and extragalactic point sources. It is therefore justifiable to combine these data into a single "optimal" estimate of the angular power spectrum of the CMB. In this section we provide an overview of two methods we use to generate a single combined spectrum. The first is a multistep process that simultaneously fits the 28 cross-power spectra presented above to a single CMB power spectrum and a point-source model, equation (9), while correctly propagating beam and residual point-source uncertainties through to a final Fisher matrix. This spectrum constitutes our best estimate of the CMB power spectrum from the first-year *WMAP* data. The second spectrum, which serves as a cross-check of the first, is based on forming a single co-added sky map from the Q1–W4 maps and using the quadratic estimator with noise bias subtraction. We compare the two spectra in § 5.3.

5.1. *Method I: Optimal Combination of Cross-Power Spectra*

Since this method is relatively complicated, we outline the basic procedure here and relegate the details to appendices, as indicated. We present the result in § 5.3. The steps are as follows:

1. Subtract best-fit Galactic foreground templates from each of the maps Q1–W4, using the coefficients given in Table 3 of Bennett et al. (2003c).

2. Evaluate the 28 cross-power spectra from the maps Q1–W4, where each spectrum has been evaluated using the quadratic estimator of § A1 with the weighting scheme defined in § A1.2.

3. Collect the noise bias estimate, $\langle N_\ell^i \rangle$, for each DA from § 2.2. These estimates are used in the calculation of the covariance matrix for the combined spectrum and in setting the relative weight of each cross-power spectrum in the final combined spectrum.

4. Apply the procedure presented in § B1 to obtain an estimate for the point-source amplitude. The value obtained is $\bar{A} = 0.0155 \pm 0.0017$, roughly independent of $\ell_{\max}$ in

equation (B6). This value is in good agreement with an estimate based on the bispectrum (Komatsu et al. 2003) and on an extrapolation of point-source counts (Bennett et al. 2003b). Subtract the point-source contribution from the cross-power spectra: $\bar{C}_\ell^i = C_\ell^i - \bar{A} S^i w_\ell^i$.

5. Compute an approximate form of the full covariance matrix discussed in § 4, $\tilde{\Sigma}_{\text{full}}$. The procedure we use produces a covariance matrix that includes cosmic variance, instrument noise, source subtraction uncertainties, and window function uncertainties. At this stage of the process, it does not include the effects of mode coupling. More details are given in § 4 and Appendix B.

6. Invert the approximate covariance matrix for use in computing the optimal spectrum. This is the most computationally intensive step in the process.

7. Compute the final combined spectrum from the 28 C_ℓ^i as per the procedure given in Appendix C. In particular, assume a fiducial cosmological model (as specified in Appendix C) and use equation (C6), with $\Sigma_{\text{full},\ell\ell'}^{ij} = \tilde{\Sigma}_{\text{full},\ell\ell'}^{ij}$. This produces a final spectrum that is very nearly optimal. Note that for $\ell < 100$ we use a surrogate procedure for computing the combined spectrum. In order to minimize Galactic foreground contamination, we use only V- and W-band data. Moreover, because the statistics of the C_ℓ are mildly non-Gaussian, and because point-source contamination and window function uncertainties are small, the "optimal" machinery developed above is unnecessary. Rather, we simply form a weighted average spectrum from the V- and W-band C_ℓ^i.

8. Compute the approximate inverse covariance matrix, $\tilde{Q}$, for the combined spectrum using equation (C7). This matrix is approximate in two ways: (1) it does not yet incorporate the effects of mode coupling (this is added below), and (2) it has been evaluated for a fiducial cosmological model, $\tilde{Q}(C_\ell^{\text{fid}})$, while in a likelihood application, we need to evaluate $\tilde{Q}(C_\ell^{\text{th}})$ for an arbitrary model, C_ℓ^{th}; we add this next.

9. Introduce the dependence on cosmological model into $\tilde{Q}$ as follows. Invert $\tilde{Q}$ to obtain the approximate covariance matrix of the combined spectrum, $\tilde{\Sigma}$. The off-diagonal terms of $\tilde{\Sigma}$ are small and weakly dependent on cosmological model, so we expand $\tilde{\Sigma}$ as

$$\tilde{\Sigma}_{\ell\ell'} = D_\ell \delta_{\ell\ell'} + \epsilon_{\ell\ell'} , \qquad (20)$$

where $\epsilon_{\ell\ell'}$ encodes the mode coupling due to window function and source subtraction uncertainties. This relation defines $\epsilon_{\ell\ell'}$, which we take to be zero on the diagonal. D_ℓ is dominated by cosmic variance and noise; in order to separate the two contributions, we introduce an *Ansatz* of the form

$$D_\ell \equiv \frac{2}{2\ell + 1} \frac{1}{f_{\text{sky}}} \left(C_\ell^{\text{fid}} + N_\ell^{\text{eff}} \right)^2 , \qquad (21)$$

where C_ℓ^{fid} is the fiducial model used to generate the combined spectrum (see Appendix C) and N_ℓ^{eff} is the effective noise in the combined spectrum, which is defined by this equation. The dependence on cosmological model is introduced in the covariance matrix by recomputing D_ℓ with $C_\ell^{\text{fid}} \to C_\ell^{\text{th}}$, leaving N_ℓ^{eff} fixed.

10. Estimate the coupling induced by the foreground mask as described in Appendix D. The effect of the mask on the off-diagonal elements of the Fisher (inverse covariance)

matrix can be written as

$$\left(\Sigma^{-1}_{\rm mask}\right)_{\ell\ell'} \equiv r_{\ell\ell'}\left(D'_\ell D'_{\ell'}\right)^{-1/2} , \qquad (22)$$

where $D'_\ell \equiv D_\ell f^{-1}_{\rm sky}$. This expression parameterizes, and defines, the off-diagonal mode coupling in the form of a correlation matrix, $r_{\ell\ell'}$. Note that $\tilde{F}_{\ell\ell'}$, defined in § D2, is related to $r_{\ell\ell'}$ by $\tilde{F}_{\ell\ell'} \equiv \delta_{\ell\ell'} + r_{\ell\ell'}$.

11. The final Fisher (curvature) matrix, $Q_{\ell\ell'}$, is obtained using

$$Q_{\ell\ell'} = \left(D'_\ell\right)^{-1}\delta_{\ell\ell'} - \epsilon_{\ell\ell'}\left(D'_\ell D'_{\ell'}\right)^{1/2} + r_{\ell\ell'}\left(D'_\ell D'_{\ell'}\right)^{-1/2} . \quad (23)$$

We further calibrate D'_ℓ with Monte Carlo simulations. This calibration process, as well as a description of how the curvature matrix is used in a maximum likelihood determination of cosmological parameters, is given in § 2 of Verde et al. (2003). As part of the first-year data release, we provide a FORTRAN 90 subroutine to evaluate the likelihood of a given cosmological model, $C^{\rm th}_\ell$, given the *WMAP* data (supplied in the routine). The code also optionally returns the Fisher (inverse covariance) matrix for the combined spectrum.

5.2. *Method II: Combined Sky Map*

Our second approach is to form a single co-added map from the Q1–W4 maps,

$$T = \frac{\sum^{10}_{i=3} T'_i/\sigma^2_{0,i}}{\sum^{10}_{i=3} 1/\sigma^2_{0,i}} , \qquad (24)$$

where T'_i is the sky map for DA i with the best-fit Galactic template model subtracted and $\sigma^2_{0,i}$ is the noise per observation for DA i, given by Bennett et al. (2003b). We evaluate the power spectrum of this map on the Kp2 cut sky using the quadratic estimator in § A1. An effective noise model is obtained by using the same approach as described in § 2.2: we generate co-added noise maps from the library of end-to-end simulations, evaluate their average spectra, and then fit a noise model. The noise bias model is then subtracted from the power spectrum of the combined temperature map. We have performed this analysis with three distinct pixel weighting schemes (see § A1.2) and three corresponding noise models. The results are shown in Figure 5, where it is seen that the three cases are virtually indistinguishable.

In effect, this analysis uses both the auto- and cross-power spectra. We view this as a useful check of the more sophisticated procedure described in § 5.1, but we do not rely on it for a final result. Uncertainties in the noise model only affect the fourth moment of the cross-power spectra, but they affect the second moment of the auto-power spectra and potentially bias the final result. The ~6% sensitivity advantage gained by including auto-power spectra was not deemed worth the effort required to guarantee that the final result was not biased.

5.3. *Comparison of Results*

Figure 6 compares the power spectra obtained from methods I and II above. The combined cross-power spectrum from § 5.1 is shown in black, while the auto-power spectrum obtained in § 5.2 from the co-added map is shown in gray. The two methods agree extremely well, with the only notable deviation being at the highest ℓ range probed
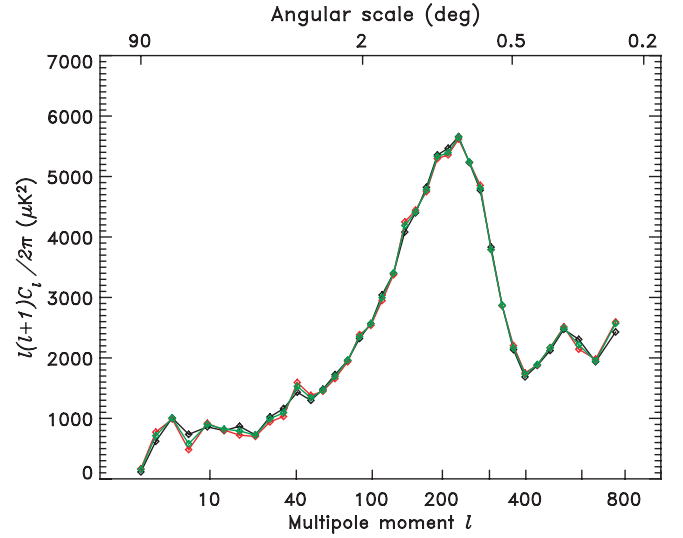

Fig. 5.—Auto-power spectrum of the combined Q+V+W map evaluated with the three weighting schemes defined in § A1.2. In each case, the spectrum was computed over the entire ℓ range; black shows uniform weights, red shows inverse noise weights, and green shows transitional weights. The agreement is excellent.

by the first-year data. This is the regime where the auto-power spectrum will be most sensitive to the noise bias subtraction.

A separate test of robustness is to compute the angular power spectrum in separate regions of the sky to see if the spectrum changes. We have computed the V × W cross-power spectrum in two subsets of the sky: the ecliptic poles $(|\beta| > 30°)$ and the ecliptic plane $(|\beta| < 30°)$, using the
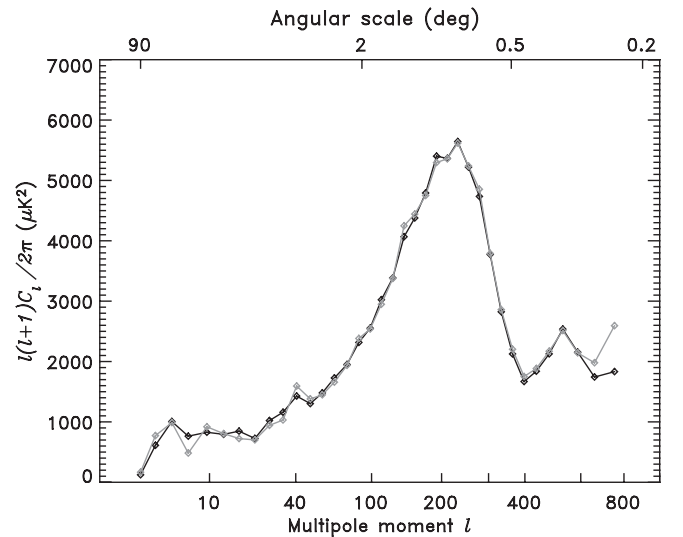

Fig. 6.—Comparison of the auto-power spectrum computed from the combined Q+V+W map (*gray*) to the optimally combined cross-power spectrum (*black*). The two methods agree extremely well, with the only notable deviation being at the highest ℓ range probed by the first-year data. This is the regime where the auto-power spectrum will be most sensitive to the noise bias subtraction. As can be seen in the error estimates shown in Fig. 8, the deviation is less than 1 σ. Note that the auto-power spectrum is not used for parameter estimation.
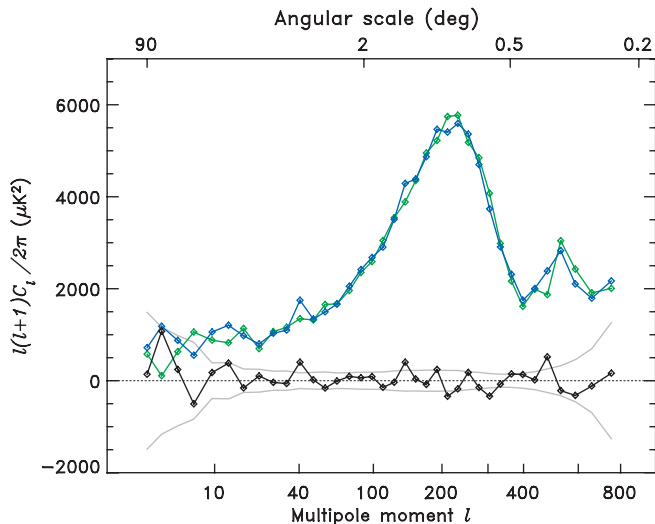
FIG. 7.—Comparison of the V × W cross-power spectrum computed with data from the ecliptic plane ($|\beta| < 30°$; *green*) vs. data from the ecliptic poles ($|\beta| > 30°$; *blue*). The difference between the two spectra is plotted in black, along with a 1 σ error band computed from Monte Carlo simulations that include the effects of sample variance. The χ^2 of the difference spectrum in the range $2 \leq \ell \leq 800$ is 750 for 799 degrees of freedom. There is no evidence that beam ellipticity, which would be more manifest in the plane than in the poles, systematically biases the spectrum. This is consistent with estimates of the effect given by Page et al. (2003b).
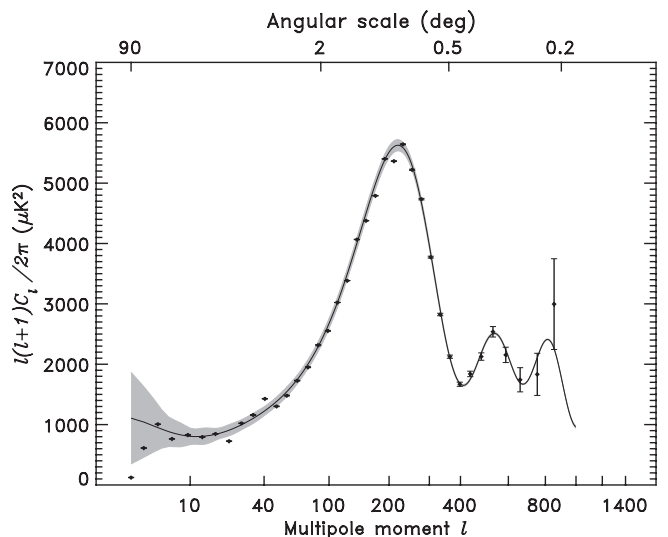

FIG. 8.—Final angular power spectrum, $\ell(\ell + 1)C_\ell/2\pi$, obtained from the 28 cross-power spectra, as described in § 5. The data are plotted with 1 σ measurement errors only, which reflect the combined uncertainty due to noise, beam, calibration, and source subtraction uncertainties. The solid line shows the best-fit ΛCDM model from Spergel et al. (2003). The gray band around the model is the 1 σ uncertainty due to cosmic variance on the cut sky. For this plot, both the model and the error band have been binned with the same boundaries as the data, but they have been plotted as a splined curve to guide the eye. On the scale of this plot the unbinned model curve would be virtually indistinguishable from the binned curve except in the vicinity of the third peak.

quadratic estimator. The results are shown in Figure 7, where the pole data are shown in blue and the plane data in green. The two spectra are very consistent overall; the difference between the two spectra is plotted in black, along with a 1 σ error band computed from Monte Carlo simulations, which include the effects of sample variance. The χ^2 of the difference spectrum in the range $2 \leq \ell \leq 800$ is 750 for 799 degrees of freedom. There is no evidence that beam ellipticity, which would be more manifest in the plane than in the poles, systematically biases the spectrum. This is consistent with estimates of the effect given by Page et al. (2003b).

6. DISCUSSION

Our best estimate of the angular power spectrum of the CMB is shown in Figure 8. Also shown is the best-fit ΛCDM model from Spergel et al. (2003), which is based on a fit to this spectrum plus a compilation of additional CMB and large-scale structure data. The *WMAP* data points are plotted with measurement errors based on the diagonal elements of the Fisher matrix presented in Appendix D. The cosmic variance errors, which include the effects of the sky cut, are plotted as a 1 σ band around the best-fit model. As discussed in Spergel et al. (2003), the model is an excellent fit to the data. The combined spectrum provides a definitive measurement of the CMB power spectrum with uncertainties limited by cosmic variance up to $\ell \sim 350$. The spectrum clearly exhibits a first acoustic peak at $\ell = 220.1 \pm 0.8$ and a second acoustic peak at $\ell = 546 \pm 10$. Page et al. (2003c) present an analysis and interpretation of the peaks and troughs in the first-year *WMAP* power spectrum. We note that the power spectrum errors in the range $\ell \sim 600$ are somewhat larger than one would have forecast based on assuming a Gaussian beam response with an FWHM equal to the measured values

because the beam response is somewhat non-Gaussian (Page et al. 2003b). However, for the same reason, the power spectrum errors at $\ell \gtrsim 1000$ are lower than one would forecast. The latter benefit will not be fully realized until more integration time reduces the noise levels at higher ℓ.

Figure 9 compares the first-year *WMAP* spectrum to a compilation of recent balloon- and ground-based measurements. In order to make this figure meaningful, we plot the best-fit model spectrum to represent the *WMAP* results. The data points are plotted with errors that include both measurement uncertainty and cosmic variance, so no error band is included with the model curve. (Since individual groups report band-power measurement with different bandwidths, it is not possible to represent a single cosmic variance band that applies to all data sets.) The model spectrum fit to *WMAP* agrees very well with the ensemble of previous observations.

Wang et al. (2002a) have recently distilled a CMB power spectrum from an optimal combination of the extant pre-*WMAP* data. In Figure 10 we plot their derived band-power points alongside the *WMAP* data. To make this comparison meaningful, we plot the *WMAP* data with cosmic variance plus measurement errors and omit the error band from the model spectrum. The distilled spectrum is notably lower than the *WMAP* data in the vicinity of the first acoustic peak. In a previous version of this work (Wang, Tegmark, & Zaldarriaga 2002b) the authors noted that the first peak of their combined spectrum was lower than a significant fraction of their input data. They attribute this to their formalism allowing for a renormalization of individual experiments within their respective calibration uncertainties. Figure 1 in Bennett et al. (2003a) presents a similarly
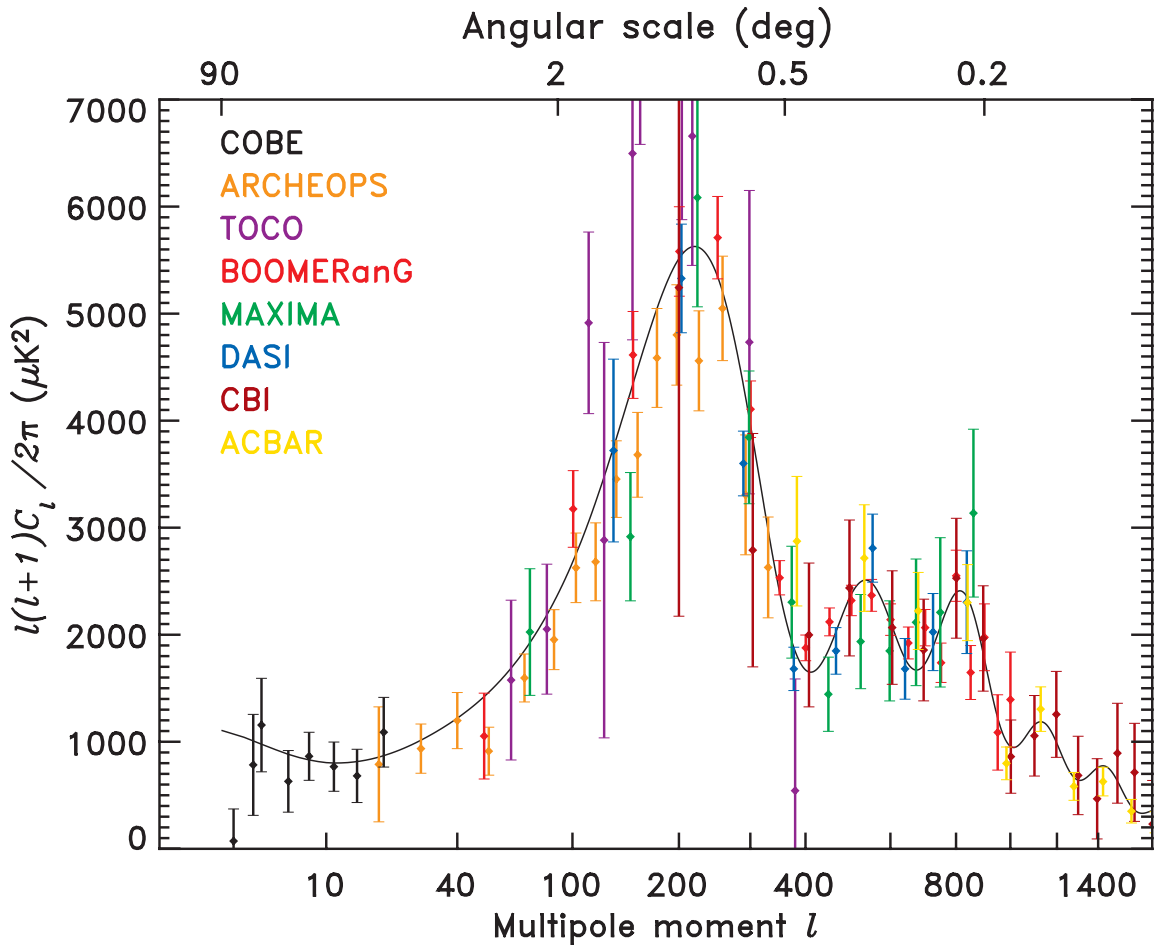
Fig. 9.—Compilation of recent CMB power spectrum measurements compared to the best-fit ΛCDM model from the first-year *WMAP* data. The data points include noise and cosmic variance uncertainty (but not calibration uncertainty); thus, we omit the cosmic variance from the model curve in the figure. On average, the pre-*WMAP* data agree well with the *WMAP* power spectrum. The references for the previous data are as follows: Tegmark (1996) for *COBE*; Benoit et al. (2003) for ARCHEOPS; Miller et al. (2002) for TOCO; Ruhl et al. (2002) for BOOMERANG; Lee et al. (2001) for MAXIMA; Halverson et al. (2002) for DASI; Pearson et al. (2003) for CBI; Kuo et al. (2002) for ACBAR.

distilled spectrum from the data extant in late 2001 and found a first-peak amplitude that was more intuitively consistent with the bulk of the input data and is now seen to be consistent with the *WMAP* power spectrum.

Figure 11 shows the *WMAP* combined power spectrum compared to the locus of predicted spectra, in red, based on a joint analysis of pre-*WMAP* CMB data and 2dFGRS large-scale structure data (Percival et al. 2002). As in Figure 8, the *WMAP* data are plotted with measurement uncertainties, and the best-fit ΛCDM model (Spergel et al. 2003) is plotted with a 1 σ cosmic variance error band. Percival et al. (2002) predict that the location of the first peak should occur at $\ell = 221.8 \pm 2.4$, which is quite consistent with the value reported by Page et al. (2003c) of $\ell = 220.1 \pm 0.8$. The height of the first peak was predicted to be in the range $4920 \pm 170 \, \mu K^2$, while Page et al. (2003c) report a measured height of $5580 \pm 75 \, \mu K^2$, about 13% higher. Unlike the position, the amplitude of the first peak has a complicated dependence on cosmological parameters. Percival et al. (2002) report best-fit parameters for a ΛCDM model that are mostly consistent with those reported by Spergel et al. (2003) for the same class of models. The only mildly disparate comparison lies in the combination of normalization, σ_8, and optical depth, τ. Percival et al. (2002) report the

product $\sigma_8 e^{-\tau} = 0.72 \pm 0.03 \pm 0.02$, where the first error is a "theory" error and the second is measurement error. While Spergel et al. (2003) do not report a maximum likelihood range for this explicit parameter combination, the product of their maximum likelihood values for σ_8 and τ yields $\sigma_8 e^{-\tau} = 0.74$, which is consistent with Percival et al. (2002) but would make the first peak a few percent higher. Small differences in $\Omega_b h^2$, $\Omega_m h^2$, and n_s may also contribute to the difference.

7. CONCLUSIONS

We present measurements of the angular power spectrum of the CMB from the first-year *WMAP* data. The eight high-frequency sky maps from DAs Q1–W4 were used to estimate 28 cross-power spectra, which are largely independent of the noise properties of the experiment. These data were tested for consistency in § 3 and then used in § 5 as input to a final combined spectrum, discussed in § 6. The procedures for estimating the uncertainties in the final combined spectrum were discussed in § 4 and in numerous appendices.

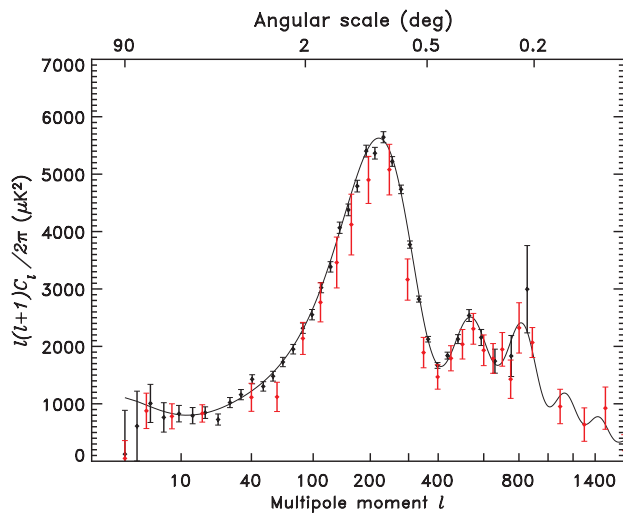The combined spectrum provides a definitive measurement of the CMB power spectrum, with uncertainties

Fig. 10.—*WMAP* combined power spectrum, in black, compared to a compilation of all CMB data published prior to *WMAP* from Wang et al. (2002a), in red. The *WMAP* data are plotted with cosmic variance plus measurement uncertainties here in order to facilitate a comparison with the compiled data that is reported in this way. The data agree well on *COBE* scales, $\ell < 20$ (but note that the *WMAP* cosmic variance errors are computed from the best-fit model rather than the data, thus they appear larger than the *COBE* errors at the quadrupole). However, the overall normalization of the *WMAP* spectrum is ~10% higher on smaller scales.
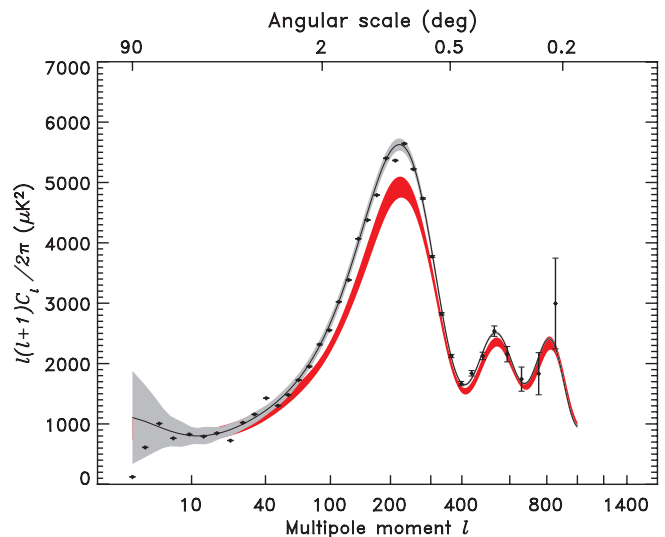

Fig. 11.—*WMAP* combined power spectrum compared to the locus of predicted spectra, in red, based on a joint analysis of pre-*WMAP* CMB data and 2dFGRS large-scale structure data (Percival et al. 2002). As in Fig. 8, the *WMAP* data are plotted with measurement uncertainties, and the best-fit ΛCDM model (Spergel et al. 2003) is plotted with a 1 σ cosmic variance error band. The locus of predicted spectra lies systematically below the *WMAP* data at intermediate ℓ.

limited by cosmic variance up to $\ell \sim 350$ and a signal-to-noise ratio per mode greater than 1 up to $\ell \sim 650$. The spectrum clearly exhibits a first acoustic peak at $\ell = 220.1 \pm 0.8$ and a second acoustic peak at $\ell = 546 \pm 10$. Page et al. (2003c) present an analysis and interpretation of the peaks and troughs in the first-year *WMAP* power spectrum. Spergel et al. (2003), Verde et al. (2003), and Peiris et al. (2003) analyze the combined spectrum in the context of cosmological models. They conclude that the data provide strong support for adiabatic initial conditions, and they give precise measurements of a number of cosmological parameters. Kogut et al. (2003) analyze the correlation between *WMAP*'s temperature and polarization signals, the C_ℓ^{TE} spectrum, and present evidence for a relatively high optical depth and an early period of cosmic reionization. Among other things, this result implies that the temperature power spectrum is suppressed by ~30% on degree angular scales, as a result of secondary scattering.

A variety of first-year *WMAP* data products are being made available by LAMBDA.[12] In addition to the sky maps

and calibrated time-ordered data, we are providing the 28 cross-power spectra used in this paper (with diffuse foregrounds subtracted), the combined spectrum from § 5.1, and a FORTRAN 90 subroutine to compute the likelihood of a given cosmological model. (The code will also optionally return the Fisher [inverse covariance] matrix for the combined spectrum.)

---

[12] The LAMBDA URL is http://lambda.gsfc.nasa.gov.

APPENDIX A

POWER SPECTRUM ESTIMATION METHODS

For the analysis of *WMAP*'s first-year data, we have chosen two distinct methods for inferring the power spectrum. The first is a fast and accurate quadratic method for estimating the power spectrum of a partial sky map (Hivon et al. 2002). We summarize the basic approach here, highlighting the aspects of the method that are especially pertinent to *WMAP*, and refer the reader to Hivon et al. (2002) for details. The second is a maximum likelihood method that provides an independent estimate of the spectrum measured by *WMAP* (Oh et al. 1999).

We have applied both of these methods to the *WMAP* data. The results are shown in Figure 12, which shows spectra estimated from the V-band map, up to $\ell = 200$, for the two methods. The maximum likelihood estimate has slightly lower uncertainties at low ℓ because the method optimally weights the data with a pixel-pixel covariance $\boldsymbol{C} = \boldsymbol{S} + \boldsymbol{N} \approx \boldsymbol{S}$, where $\boldsymbol{S}$ is
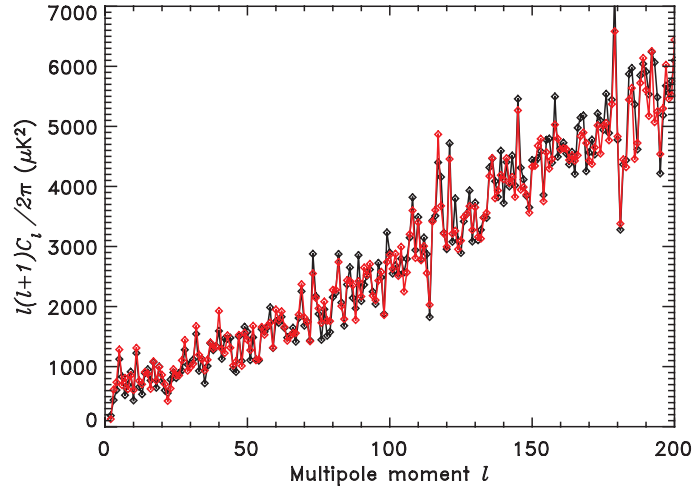
Fɪɢ. 12.—Comparison of two power spectra computed from a single V-band map. The black points result from the maximum likelihood method, the red points from the quadratic estimator computed with uniform pixel weighting. In both cases a noise model has been assumed to treat the noise bias. At higher ℓ, the two spectra would be nearly identical since both impose inverse noise weighting on the data.

the covariance of the CMB signal and N is the covariance of the noise (see § A2). Our quadratic estimator uses uniform pixel weights at low ℓ (see § A1.2) although it is clear from the figure that the difference is not significant. At high ℓ, where the *WMAP* data are noise dominated, the two estimators give essentially identical results because they effectively weight the data in the same way.

To obtain our " best " estimate of the *WMAP* power spectrum, we adopt the quadratic estimator because it can be readily applied to pairs of *WMAP* radiometers in a way that is nearly independent of the properties of the instrument noise. In § 5 we discuss our methodology for combining spectra from pairs of radiometers and present the final combined spectrum.

A1. QUADRATIC ESTIMATION

Hivon et al. (2002) start with the full-sky estimator in equation (2), add a position-dependent weight, $W(\mathbf{n})$, and set W to zero in the regions where the sky is contaminated. In other regions, W is chosen to optimize the sensitivity of the estimator. A temperature map $\Delta T(\mathbf{n})$ on which a weight $W(\mathbf{n})$ is applied can be decomposed in spherical harmonic coefficients as

$$\tilde{a}_{\ell m} = \int d\Omega_{\mathbf{n}} \, \Delta T(\mathbf{n}) \, W(\mathbf{n}) \, Y^*_{\ell m}(\mathbf{n}) \tag{A1}$$

$$\approx \Omega_p \sum_p \Delta T(p) \, W(p) \, Y^*_{\ell m}(p) \, , \tag{A2}$$

where the integral over the sky is approximated by a discrete sum over map pixels, each of which subtends solid angle Ω_p. Hivon et al. (2002) then define the " pseudo–power spectrum " $\tilde{C}_\ell$ as

$$\tilde{C}_\ell = \frac{1}{2\ell + 1} \sum_{m=-\ell}^{\ell} |\tilde{a}_{\ell m}|^2 \, . \tag{A3}$$

The pseudo–power spectrum $\tilde{C}_\ell$, given by the weighted spherical harmonic transform of a map, is clearly different from the full-sky angular spectrum, C^{sky}_ℓ, but the *ensemble averages* of the two spectra can be related by

$$\langle \tilde{C}_\ell \rangle = \sum_{\ell'} G_{\ell\ell'} \langle C^{\mathrm{sky}}_{\ell'} \rangle \, , \tag{A4}$$

where $G_{\ell\ell'}$ describes the mode coupling resulting from the weight function $W(\mathbf{n})$ (Hauser & Peebles 1973). Hivon et al. (2002) give the following expression for the coupling matrix, which depends only on the geometry of the weight function $W(\mathbf{n})$:

$$G_{\ell_1 \ell_2} = \frac{2\ell_2 + 1}{4\pi} \sum_{\ell_3} (2\ell_3 + 1) \mathscr{W}_{\ell_3} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ 0 & 0 & 0 \end{pmatrix}^2 \, , \tag{A5}$$

where the final term in parentheses is the Wigner 3-*j* symbol and $\mathscr{W}_\ell$ is the angular power spectrum of the weight function

$$\mathscr{W}_\ell = \frac{1}{2\ell + 1} \sum_m |w_{\ell m}|^2 \, , \tag{A6}$$

where

$$w_{\ell m} = \int d\Omega_{\boldsymbol{n}} \, W(\boldsymbol{n}) \, Y^*_{\ell m}(\boldsymbol{n}) \; . \tag{A7}$$

Upon inverting the coupling matrix $G_{\ell \ell'}$ and making the identification $\langle C_\ell^{\mathrm{sky}} \rangle = C_\ell$, we obtain the following estimator of the power spectrum:

$$C_\ell^{\mathrm{obs}} = \sum_{\ell'} G_{\ell \ell'}^{-1} \tilde{C}_{\ell'} \; . \tag{A8}$$

The computation of equation (A2) for each (ℓ, m) up to $\ell = \ell_{\max}$ would scale as $N_{\mathrm{pix}} \ell_{\max}^2$ if performed on an arbitrary pixelization of the sphere, where N_{pix} is the number of sky map pixels. However, for a pixelization scheme with isolatitude pixel centers, fast Fourier transform methods may be employed to speed up the evaluation, so it scales like $N_{\mathrm{pix}}^{1/2} \ell_{\max}^2$ (Muciaccia, Natoli, & Vittorio 1997). The *WMAP* sky maps have been produced using the HEALPix layout (Górski et al. 1998), which supports such fast spherical harmonic transforms. In particular, the HEALPix routine map2alm evaluates equation (A2).

A1.1. *Auto- and Cross-Power Spectra from the WMAP Data*

For a multichannel experiment like *WMAP* it is quite powerful to evaluate the power spectra from different maps and compare results. In particular, the quadratic estimator described above may be used on one or two maps at a time by generalizing the expression for the pseudo–power spectrum equation (A3) as

$$\tilde{C}_\ell = \frac{1}{2\ell + 1} \sum_{m=-\ell}^{\ell} \tilde{a}_{\ell m}^i \tilde{a}_{\ell m}^{j*} \; , \tag{A9}$$

where $\tilde{a}_{\ell m}^i$ refers to the transform of map i and $\tilde{a}_{\ell m}^{j*}$ refers to map j, which need not be the same as map i. As discussed in § 2, if $i \neq j$ and the noise in the two maps is uncorrelated, the estimator equation (A9) provides an unbiased estimate of the underlying power spectrum.

We have tested the auto- and cross-power estimator extensively with Monte Carlo simulations of the first-year *WMAP* data. Selected results from this testing are shown in Figure 13. The auto- and cross-power spectra that are obtained from the flight data are presented in detail in § 3. The cross-power spectra form the basis for our final combined spectrum, presented in § 5.

A1.2. *Choice of Weighting*

We seek a weighting scheme that mimics the maximum likelihood estimation (§ A2), which effectively weights the data by the full inverse covariance matrix $\boldsymbol{C}^{-1} = (\boldsymbol{S} + \boldsymbol{N})^{-1}$. For the combined spectrum presented in § 5, we use three distinct weighting functions in three separate ℓ ranges:

1. For $\ell < 200$ we give equal weight to all uncut pixels,

$$W(p) = M(p) \; , \tag{A10}$$

where $M(p)$ is the Kp2 sky mask, defined by Bennett et al. (2003c). It takes values of 0 within the mask and 1 otherwise.

2. For $\ell > 450$ we use inverse noise weighting,

$$W(p) = M(p) N_{\mathrm{obs}}(p) \; , \tag{A11}$$

where $N_{\mathrm{obs}}(p)$ is the number of observations of pixel p.

3. For $200 < \ell < 450$ we use a transitional weighting,

$$W(p) = \frac{M(p)}{1/\langle N_{\mathrm{obs}} \rangle + 1/N_{\mathrm{obs}}(p)} \; , \tag{A12}$$

where $\langle N_{\mathrm{obs}} \rangle$ is the mean of N_{obs} evaluated over the cut sky.

Verde et al. (2003) discuss the choice of weighting in more detail.

A2. MAXIMUM LIKELIHOOD ESTIMATION

This appendix provides a summary of the maximum likelihood approach to power spectrum estimation originally presented by Oh et al. (1999). If the temperature fluctuations are Gaussian and the a priori probability of a given set of cosmological parameters is uniform, then the power spectrum may be estimated by maximizing the multivariate Gaussian likelihood function

$$\mathscr{L}(C_\ell | \boldsymbol{m}) = \frac{\exp\left[-(1/2) \boldsymbol{m}^T \boldsymbol{C}^{-1} \boldsymbol{m} \right]}{\sqrt{\det \boldsymbol{C}}} \; , \tag{A13}$$
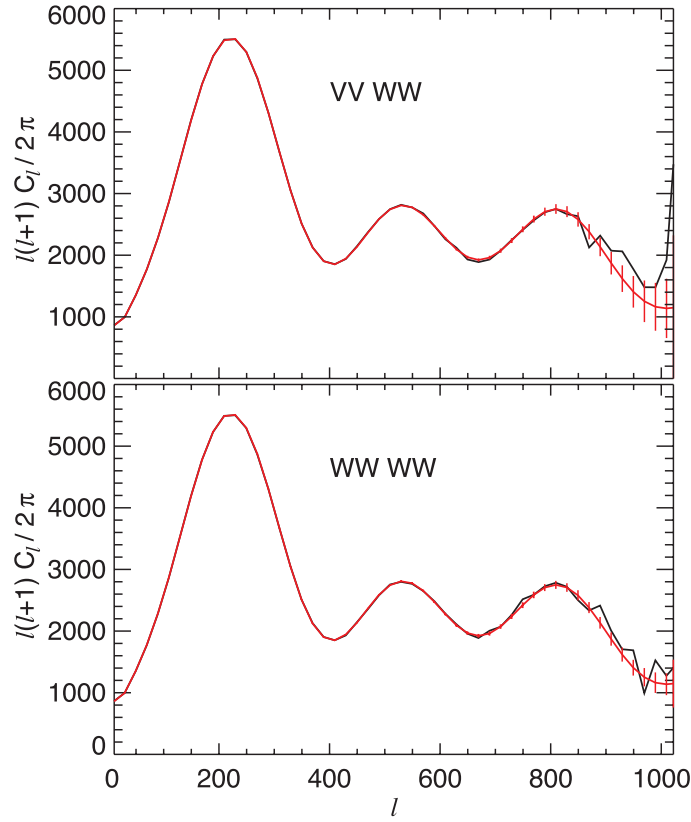
Fig. 13.—Quadratic estimator discussed in § A1, which has been extensively tested with Monte Carlo simulations of first-year *WMAP* data. The top panel shows a model spectrum (*red*) and the mean of 600 realizations of the cross-power spectrum, computed from the V- and W-band–averaged maps (*black*). The red error bars show the uncertainty in the mean in bins with $\Delta\ell = 20$. The cross-power spectrum estimator is unbiased. The bottom panel is similar, but with auto-power spectra computed from the W-band–averaged map. Here the noise bias term was estimated from Monte Carlo simulations of noise maps. This simple auto-power spectrum estimator is unbiased if the noise is white and is suitable for V-band data, but see Fig. 1 for examples of W-band noise spectra.

where m is a data vector (see below) and C is the covariance matrix of the data, which has contributions from both the signal and the instrument noise, $C = S + N$. We can work in whatever basis is most convenient. In the pixel basis the data are the sky map pixel temperatures, and in the spherical harmonic basis the data are the $a_{\ell m}$ coefficients of the map. In the former basis the noise covariance is nearly diagonal, while in the latter basis the signal covariance is. Specifically, in the pixel basis

$$m \rightarrow T_i \, ,$$

$$S \rightarrow \sum_\ell \frac{(2\ell + 1)}{4\pi} C_\ell P_\ell(\cos\theta_{ij}) \, ,$$

$$N \rightarrow \sigma_i^2 \delta_{ij} \, ,$$

while in the spherical harmonic basis

$$m \rightarrow a_{\ell m} \, ,$$

$$S \rightarrow \mathrm{diag}(C_2, C_2, \ldots C_3, C_3, \ldots) \, ,$$

$$N \rightarrow N_{(\ell m)(\ell m)'}$$

For *WMAP*, the length of the data vector, N_{data}, is 2,672,361, the number of 7′ sky map pixels (HEALPix $N_{\mathrm{side}} = 512$) that survive the Galaxy cut, so it is necessary to find methods for evaluating $\mathscr{L}$ that do not require a full inversion of the covariance matrix C, which requires $O(N_{\mathrm{data}}^3)$ operations. We use an iterative method for evaluating the likelihood that exploits the ability to find an approximate inverse $\tilde{C}^{-1}$. The most important features in the data that make this possible are that *WMAP* observes the full sky and the Galaxy cut is predominantly azimuthally symmetric in Galactic coordinates (Bennett et al. 2003b). Of secondary importance for this preconditioner is that *WMAP*'s noise per pixel does not vary strongly across the sky (Bennett et al. 2003a). We discuss the preconditioner in more detail below.

Defining $f \equiv -2\ln\mathscr{L}$ and $P_\ell \equiv \partial C/\partial C_\ell$, we maximize the likelihood by solving

$$\frac{\partial f}{\partial C_\ell} = 0 = m^T C^{-1} P_\ell C^{-1} m + \mathrm{tr}(C^{-1} P_\ell) \, , \tag{A14}$$

using a Newton-Raphson root finding method that generates an iterative estimate of the angular power spectrum (Bond, Jaffe, & Knox 1998)

$$C_\ell^{(n+1)} = C_\ell^{(n)} - \frac{1}{2} \sum_{\ell'} F_{\ell\ell'} \frac{\partial f}{\partial C_{\ell'}} \tag{A15}$$

at each step. Here $F_{\ell\ell'}$ is the Fisher matrix

$$F_{\ell\ell'} = -\left\langle \frac{\partial^2 \ln \mathscr{L}}{\partial C_\ell \partial C_{\ell'}} \right\rangle = \frac{1}{2} \text{tr}(C^{-1} P_\ell C^{-1} P_{\ell'}) \ . \tag{A16}$$

To implement the solution in equation (A15), we need a fast way to evaluate the following three components of $\sum_{\ell'} F_{\ell\ell'}(\partial f/\partial C_{\ell'})$: (1) $m^T C^{-1} P_\ell C^{-1} m$, (2) $\text{tr}(C^{-1} P_\ell)$, and (3) $\text{tr}(C^{-1} P_\ell C^{-1} P_{\ell'})$. We use the spherical harmonic basis in which the data vector consists of the $a_{\ell m}$ coefficients of the map obtained by least-squares fitting on the cut sky. The signal covariance is diagonal in this basis, while the noise matrix is obtained from the normal equations for the $a_{\ell m}$ fit

$$\sum_{(\ell m)'} N_{(\ell m)(\ell m)'}^{-1} a_{(\ell m)'} = y_{(\ell m)} \ , \tag{A17}$$

where

$$N_{(\ell m)(\ell m)'}^{-1} \equiv \sum_i \frac{Y_{(\ell m)}(n_i) Y_{(\ell m)'}(n_i)}{\sigma_i^2} \ , \tag{A18}$$

$$y_{(\ell m)} \equiv \sum_i \frac{T_i Y_{(\ell m)}(n_i)}{\sigma_i^2} \ . \tag{A19}$$

The sums are over all sky map pixels that survive the Galaxy cut, and we have assumed that the noise is uncorrelated from pixel to pixel.

A2.1. *Evaluation of $C^{-1} m$*

The term $C^{-1} m$ appears repeatedly in the evaluation of equation (A15). We compute this by solving $Cz = (S + N)z = m$ for z. A more numerically tractable system is obtained by multiplying both sides by $S^{1/2} N^{-1}$, so

$$(I + S^{1/2} N^{-1} S^{1/2}) S^{1/2} z = S^{1/2} N^{-1} m = S^{1/2} y \ , \tag{A20}$$

where y is the spherical harmonic transform of the map, defined in equation (A18). Note that y can be quickly computed in any pixelization scheme that has isolatitude pixel centers with fixed longitude spacing. We then solve equation (A20) using an iterative conjugate gradient method with a preconditioner for the matrix $A \equiv (I + S^{1/2} N^{-1} S^{1/2})$. We find the following block-diagonal form of A to be a good starting point:

$$\tilde{A} = \begin{pmatrix} I + S^{1/2} \tilde{N}^{-1} S^{1/2} & 0 \\ 0 & \text{diag}(I + S^{1/2} \tilde{N}^{-1} S^{1/2}) \end{pmatrix} \ , \tag{A21}$$

where $\tilde{N}^{-1}$ is a block-diagonal approximation of the noise matrix discussed below. The lower right block of $\tilde{A}$ occupies the high-ℓ portion of the matrix where the signal-to-noise ratio $S^{1/2} N^{-1} S^{1/2}$ is low, so a diagonal approximation is adequate. The upper left block occupies the low-ℓ portion of the matrix where the signal dominates the noise, so we need a better estimate of N^{-1}. In practice, we find that this split works well at $\ell = 512$ for the *WMAP* noise levels. As for the approximate form of N^{-1}, defined in equation (A19), note that the dominant off-diagonal contributions arise from the azimuthally symmetric Galaxy cut, which couples different ℓ-modes, but not m-modes. Thus, N^{-1} is approximately block-diagonal, with perturbations induced by the nonuniform sky coverage of *WMAP*. We therefore use a block-diagonal form of N^{-1} as the preconditioner,

$$\tilde{N}_{(\ell m)(\ell m)'}^{-1} = N_{(\ell m)(\ell m)'}^{-1} \delta_{mm'} \ . \tag{A22}$$

Using the preconditioner equation (A21), we find that the conjugate gradient solution of equation (A20) converges in approximately six iterations and requires only CPU minutes of processing on an SGI Origin 2000.

APPENDIX B

POINT-SOURCE SUBTRACTION

In this appendix we describe the procedure we use to estimate and subtract the point-source contribution directly from the multifrequency cross-power spectra. We then show how we incorporate the source model uncertainty into the covariance

matrix of the source-subtracted spectra by marginalizing a Gaussian likelihood function over the source model amplitude parameter.

This marks the starting point of the multifrequency analysis that will lead to the combined power spectrum, discussed in § 5.1. In order to generate the combined spectrum, we need the full covariance matrix of the cross-power spectra (see § 4). Our approach to generating the full covariance is to start with the ideal, full-sky form, which only includes cosmic variance and instrument noise, and then we incorporate additional effects step by step, as outlined in § 5.1 and in these appendices. For an ideal experiment with full-sky coverage, no point-source contamination, and no beam uncertainty, the covariance matrix is

$$\Sigma_{\ell\ell'}^{ii'jj'} = \frac{1}{(2\ell+1)}\left[(C_\ell^{\mathrm{th}}w_\ell^{ij} + n^i n^j \delta_{ij})(C_\ell^{\mathrm{th}}w_\ell^{i'j'} + n^{i'}n^{j'}\delta_{i'j'}) + (C_\ell^{\mathrm{th}}w_\ell^{ij'} + n^i n^{j'}\delta_{ij'})(C_\ell^{\mathrm{th}}w_\ell^{i'j} + n^{i'}n^j \delta_{i'j})\right], \tag{B1}$$

where δ_{ij} denotes the Kronecker symbol, $w_\ell^{ij} \equiv b_\ell^i b_\ell^j p_\ell^2$ is the window function, and $n^i n^{i'}\delta_{ii'} = N^i$.

B1. ESTIMATING THE POINT-SOURCE AMPLITUDE

We start by assuming a Gaussian likelihood for the sky model, given the *WMAP* data,

$$-2\ln \mathscr{L}(A, C_\ell | C_\ell^i) = \sum_{ij\ell}\left[C_\ell^i - (C_\ell + AS^i)w_\ell^i\right](\Sigma^{-1})_\ell^{ij}\left[C_\ell^j - (C_\ell + AS^j)w_\ell^j\right], \tag{B2}$$

where C_ℓ^i is cross-power spectrum i, w_ℓ^i is the window function for spectrum i, C_ℓ is the true CMB power spectrum, and AS^i is the source model defined in equations (9) and (11). Here we assume the diagonal form of the covariance matrix in equation (B1).

To determine the best-fit source amplitude, A, we marginalize this likelihood over the CMB spectrum, C_ℓ. First we expand equation (B2) as

$$-2\ln \mathscr{L} = \sum_\ell \sum_{ij}(C_\ell^i - AS^i w_\ell^i)(\Sigma^{-1})_\ell^{ij}(C_\ell^j - AS^j w_\ell^j) - \sum_\ell 2C_\ell \sum_{ij}(C_\ell^i - AS^i w_\ell^i)(\Sigma^{-1})_\ell^{ij}w_\ell^j + \sum_\ell C_\ell^2 \sum_{ij} w_\ell^i(\Sigma^{-1})_\ell^{ij}w_\ell^j, \tag{B3}$$

which is a quadratic form in C_ℓ, $\mathscr{L} \propto \prod_\ell \exp[-\frac{1}{2}(aC_\ell^2 - 2bC_\ell + c)]$, with

$$a \equiv \sum_{ij} w_\ell^i (\Sigma^{-1})_\ell^{ij}w_\ell^j,$$

$$b \equiv \sum_{ij}(C_\ell^i - AS^i w_\ell^i)(\Sigma^{-1})_\ell^{ij}w_\ell^j,$$

$$c \equiv \sum_{ij}(C_\ell^i - AS^i w_\ell^i)(\Sigma^{-1})_\ell^{ij}(C_\ell^j - AS^j w_\ell^j). \tag{B4}$$

We wish to evaluate the marginalized likelihood, $\mathscr{L}_{C_\ell}(A) = \int dC_\ell \mathscr{L}(A, C_\ell)$. This is readily evaluated using the substitution $C_\ell \to C_\ell - b/a$, giving $\mathscr{L}_{C_\ell} \propto \prod_\ell \exp[-\frac{1}{2}(c - b^2/a)]$, where we drop a multiplicative term proportional to a, which is independent of A. The marginalized likelihood function is thus

$$-2\ln \mathscr{L}_{C_\ell} = \sum_\ell \sum_{ij}(C_\ell^i - AS^i w_\ell^i)(\Sigma^{-1})_\ell^{ij}(C_\ell^j - AS^j w_\ell^j) - \sum_\ell \frac{1}{\sum_{i'j'} w_\ell^{i'}(\Sigma^{-1})_\ell^{i'j'}w_\ell^{j'}}\left[\sum_{ij}(C_\ell^i - AS^i w_\ell^i)(\Sigma^{-1})_\ell^{ij}w_\ell^j\right]^2. \tag{B5}$$

Setting $\partial \mathscr{L}_{C_\ell}(A)/\partial A = 0$ gives the most likely point-source amplitude

$$\bar{A} = \frac{\sum_{ij\ell} C_\ell^j (\Sigma^{-1})_\ell^{ij} H_\ell^i w_\ell^i}{\sum_{ij\ell} S^j w_\ell^j (\Sigma^{-1})_\ell^{ij} H_\ell^i w_\ell^i}, \tag{B6}$$

where

$$H_\ell^i = S^i - \frac{\sum_{i'j'} S^{i'}(\Sigma^{-1})_\ell^{i'j'}w_\ell^{j'}}{\sum_{i'j'} w_\ell^{i'}(\Sigma^{-1})_\ell^{i'j'}w_\ell^{j'}}. \tag{B7}$$

The standard error on the best-fit value of A is

$$\sigma_{\mathrm{src}}^2 = \left[\sum_{ij\ell} H_\ell^i w_\ell^i (\Sigma^{-1})_\ell^{ij} S_\ell^j w_\ell^j\right]^{-1}. \tag{B8}$$

For *WMAP*, the off-diagonal terms in the covariance matrix are small. Here neglecting them changes the inferred point-source amplitude by less than 1%.

B2. MARGINALIZING OVER POINT-SOURCE AMPLITUDE

The source subtraction procedure discussed above is uncertain. In this section we incorporate this uncertainty into the full covariance matrix for the cross-power spectra. We again assume a Gaussian likelihood function of the form

$$-2\ln\mathscr{L} = \sum_{ij\ell\ell'}\big[\bar{C}^i_\ell - \big(C^{\mathrm{th}}_\ell + \delta A S^i_\ell\big)w^i_\ell\big]\big(\Sigma^{-1}\big)^{ij}_{\ell\ell'}\big[\bar{C}^j_{\ell'} - \big(C^{\mathrm{th}}_{\ell'} + \delta A S^j_{\ell'}\big)w^j_{\ell'}\big]\,, \tag{B9}$$

where $\bar{C}^i_\ell$ is the source-subtracted cross-power spectrum for DA pair i, obtained above, w^i_ℓ is the window function for spectrum i, C^{th}_ℓ is now a fixed CMB model spectrum, and δA is the residual source amplitude.

We can marginalize the likelihood function over the residual point-source amplitude as follows. Expand equation (B9) as

$$-2\ln\mathscr{L} = \sum_{ij\ell\ell'}\big(\bar{C}^i_\ell - C^{\mathrm{th}}_\ell w^i_\ell\big)\big(\Sigma^{-1}\big)^{ij}_{\ell\ell'}\big(\bar{C}^j_{\ell'} - C^{\mathrm{th}}_{\ell'}w^j_{\ell'}\big) - 2(\delta A)\sum_{ij\ell\ell'}\big(\bar{C}^i_\ell - C^{\mathrm{th}}_\ell w^i_\ell\big)\big(\Sigma^{-1}\big)^{ij}_{\ell\ell'}S^j w^j_{\ell'} + (\delta A)^2\sum_{ij\ell\ell'}S^i w^i_\ell\big(\Sigma^{-1}\big)^{ij}_{\ell\ell'}S^j w^j_{\ell'}\,, \tag{B10}$$

which is a quadratic form in δA, $\mathscr{L}\propto\exp\{-\tfrac{1}{2}[a(\delta A)^2 - 2b(\delta A) + c]\}$, with

$$a \equiv \sum_{ij\ell\ell'}S^i w^i_\ell\big(\Sigma^{-1}\big)^{ij}_{\ell\ell'}S^j w^j_{\ell'}\,,$$

$$b \equiv \sum_{ij\ell\ell'}\big(\bar{C}^i_\ell - C^{\mathrm{th}}_\ell w^i_\ell\big)\big(\Sigma^{-1}\big)^{ij}_{\ell\ell'}S^j w^j_{\ell'}\,,$$

$$c \equiv \sum_{ij\ell\ell'}\big(\bar{C}^i_\ell - C^{\mathrm{th}}_\ell w^i_\ell\big)\big(\Sigma^{-1}\big)^{ij}_{\ell\ell'}\big(\bar{C}^j_{\ell'} - C^{\mathrm{th}}_{\ell'}w^j_{\ell'}\big)\,. \tag{B11}$$

We wish to evaluate the marginalized likelihood, $\mathscr{L}_A = \int\mathscr{L}\,d(\delta A)$. This is readily evaluated using the substitution $\delta A \to \delta A - b/a$, giving $\mathscr{L}_A \propto \exp[-\tfrac{1}{2}(c - b^2/a)]$, where we drop a multiplicative term proportional to a, which is independent of $\bar{C}^i_\ell$. The marginalized likelihood function is thus

$$-2\ln\mathscr{L}_A = \sum_{ij\ell\ell'}\big(\bar{C}^i_\ell - C^{\mathrm{th}}_\ell w^i_\ell\big)\big(\Sigma^{-1}\big)^{ij}_{\ell\ell'}\big(\bar{C}^j_{\ell'} - C^{\mathrm{th}}_{\ell'}w^j_{\ell'}\big) - \frac{1}{a}\sum_{ij\ell\ell'}\big(\bar{C}^i_\ell - C^{\mathrm{th}}_\ell w^i_\ell\big)\big(\Sigma^{-1}\big)^{ij}_{\ell\ell'}S^j w^j_{\ell'}\sum_{i'j'\ell''\ell'''}\big(\bar{C}^{i'}_{\ell''} - C^{\mathrm{th}}_{\ell''}w^{i'}_{\ell''}\big)\big(\Sigma^{-1}\big)^{i'j'}_{\ell''\ell'''}S^{j'}w^{j'}_{\ell'''}\,. \tag{B12}$$

This expression may be recast in the form

$$-2\ln\mathscr{L}_A = \sum_{ij\ell\ell'}\big(\bar{C}^i_\ell - C^{\mathrm{th}}_\ell w^i_\ell\big)(F^{\mathrm{src}})^{ij}_{\ell\ell'}\big(\bar{C}^j_{\ell'} - C^{\mathrm{th}}_{\ell'}w^j_{\ell'}\big)\,, \tag{B13}$$

where $(F^{\mathrm{src}})^{ij}_{\ell\ell'}$ is

$$(F^{\mathrm{src}})^{ij}_{\ell\ell'} \equiv \big(\Sigma^{-1}\big)^{ij}_{\ell\ell'} - \frac{1}{a}B^{ij}_{\ell\ell'}\,, \tag{B14}$$

with

$$B^{ij}_{\ell\ell'} \equiv \sum_{i'j'\ell''\ell'''}\big(\Sigma^{-1}\big)^{ii'}_{\ell\ell''}S^{i'}w^{i'}_{\ell''}\big(\Sigma^{-1}\big)^{jj'}_{\ell'\ell'''}S^{j'}w^{j'}_{\ell'''}\,. \tag{B15}$$

The superscript "src" indicates that the Fisher matrix so obtained includes point-source subtraction uncertainty, in addition to whatever effects have been included in $(\Sigma^{-1})^{ij}_{\ell\ell'}$ already, in this case only cosmic variance and noise. Note that equation (B14) neglects a term proportional to $\ln(\det a)$ that has a weak dependence on cosmological parameters. In the actual calculation, as in the previous section, we assume the diagonal form of the covariance matrix in equation (B1).

APPENDIX C

OPTIMAL WEIGHTING OF MULTICHANNEL SPECTRA

We use the eight high-frequency DAs Q1–W4 to generate the final combined spectrum. In this appendix we show how we combine these spectra into a single estimate of the angular power spectrum.

The ultimate goal of the *WMAP* analysis is to produce a likelihood function for a set of cosmological parameters, $\{a\}$, given the data, C^i_ℓ. Specifically,

$$\mathscr{P}\big(a|C^i_\ell\big) = \mathscr{L}\big[C^i_\ell|C^{\mathrm{th}}_\ell(a)\big]\mathscr{P}(a)\,, \tag{C1}$$

where $\mathscr{P}(a|C^i_\ell)$ is the probability of $\{a\}$ given the data, $\mathscr{L}[C^i_\ell|C^{\mathrm{th}}_\ell(a)]$ is the likelihood of the data given the model, $C^{\mathrm{th}}_\ell(a)$, and

$\mathscr{P}(\boldsymbol{a})$ is the prior probability of the parameter set (Verde et al. 2003). To this end, we seek a combined spectrum, $\hat{C}_\ell$, that estimates the power spectrum in our sky, C_ℓ^{sky}, with the property that $\mathscr{P}(\boldsymbol{a}|\hat{C}_\ell) = \mathscr{P}(\boldsymbol{a}|C_\ell^i)$, and hence $\mathscr{L}[\hat{C}_\ell|C_\ell^{\text{th}}(\boldsymbol{a})] = \mathscr{L}[C_\ell^i|C_\ell^{\text{th}}(\boldsymbol{a})]$.

To estimate the combined spectrum, we approximate the likelihood function for the cross-power spectra as Gaussian,

$$-2\ln\mathscr{L}(\bar{C}_\ell^i|C_\ell^{\text{th}}) = \sum_{ij\ell\ell'}(\bar{C}_\ell^i - C_\ell^{\text{th}}w_\ell^i)(\Sigma_{\text{full}}^{-1})_{\ell\ell'}^{ij}(\bar{C}_{\ell'}^j - C_{\ell'}^{\text{th}}w_{\ell'}^j),\tag{C2}$$

where $\bar{C}_\ell^i$ is the spectrum with the best-fit source model subtracted, defined after equation (12), w_ℓ^i is the window function of spectrum i, defined after equation (6), and Σ_{full} is the covariance matrix of the 28 cross-power spectra. Note that the treatment in the remainder of this section does not depend on any specific property of the covariance, so we use generic notation for readability. However, when we generate the *WMAP* first-year combined spectrum, the actual form of the covariance used at this step is $(\tilde{\Sigma}_{\text{full}})_{\ell\ell'}^{ij}$, where $\tilde{\Sigma}$ indicates the approximate covariance, obtained in § 4, that has not yet had the effects of the foreground mask accounted for.

We seek a spectrum $\hat{C}_\ell$ such that

$$-2\ln\mathscr{L}(\hat{C}_\ell|C_\ell^{\text{th}}) \equiv \sum_{\ell\ell'}(\hat{C}_\ell - C_\ell^{\text{th}})Q_{\ell\ell'}(\hat{C}_{\ell'} - C_{\ell'}^{\text{th}}) = -2\ln\mathscr{L}(\bar{C}_\ell^i|C_\ell^{\text{th}}),\tag{C3}$$

where $Q_{\ell\ell'}$ is the inverse covariance matrix of the combined spectrum that comes with the estimate of $\hat{C}_\ell$. Suppose, for simplicity, that Σ_{full} is diagonal, $(\Sigma_{\text{full}})_{\ell\ell'}^{ij} = (\Sigma_{\text{full}})_\ell^{ij}\delta_{\ell\ell'}$; then it is straightforward to show that the deconvolved, weighted-average spectrum

$$\hat{C}_\ell = \frac{\sum_{ij}\bar{C}_\ell^i(\Sigma_{\text{full}}^{-1})_\ell^{ij}w_\ell^j}{Q_\ell},\tag{C4}$$

where

$$Q_\ell = \sum_{ij} w_\ell^i(\Sigma_{\text{full}}^{-1})_\ell^{ij}w_\ell^j\tag{C5}$$

is the desired spectrum. Substituting equations (C4) and (C5) into equation (C3) produces equation (C2) up to a term that has a weak dependence on cosmological model, which we ignore. This combined spectrum is equivalent to the result we would obtain using the " optimal data compression " of Tegmark, Taylor, & Heavens (1997).

If the inverse covariance matrix is not diagonal in ℓ, it can be shown that the optimal combined spectrum is given by

$$\hat{C}_\ell = \frac{\sum_{ij\ell'}\bar{C}_\ell^i(\Sigma_{\text{full}}^{-1})_{\ell\ell'}^{ij}C_{\ell'}^{\text{fid}}w_{\ell'}^j}{\sum_{\ell'}Q_{\ell\ell'}C_{\ell'}^{\text{fid}}},\tag{C6}$$

with

$$Q_{\ell\ell'} = \sum_{ij} w_\ell^i(\Sigma_{\text{full}}^{-1})_{\ell\ell'}^{ij}w_{\ell'}^j,\tag{C7}$$

and where C_ℓ^{fid} is a fiducial cosmological model that we take to be a flat ΛCDM model with $\Omega_b h^2 = 0.021$, $\Omega_c h^2 = 0.129$, $\Omega_{\text{tot}} = 1$, $h = 0.68$, $n_s = 1.2$, and $\tau = 0.2$. While this model has parameter values that are substantially different than the best-fit *WMAP* model obtained (afterward) by Spergel et al. (2003), the parameter degeneracies are such that C_ℓ^{fid} is close to the best-fit model for $\ell > 100$ where this estimator is actually used (see § 5.1). The combined spectrum is optimal if the fiducial model chosen is the correct one; otherwise, it is still unbiased but slightly suboptimal (Gupta & Heavens 2002).

APPENDIX D

CUT-SKY FISHER MATRIX

The *WMAP* sky maps have nearly diagonal pixel-pixel noise covariance (Hinshaw et al. 2003), which greatly simplifies the properties of the power spectrum Fisher matrix. In this appendix we present an analytic derivation of the effect of a sky cut and nonuniform pixel weighting. In § D1 we assume that we have an optimal estimator of the power spectrum. In the noise-dominated limit, we can obtain an exact expression, while in the signal-dominated limit, we need to approximate the signal correlation matrix to obtain an analytic expression. In § D2 we interpolate the Fisher matrix between the signal- and noise-dominated regimes and show that it agrees with numerical estimates. In § D3 we estimate the power spectrum covariance matrix from Monte Carlo simulations of the sky and calibrate the interpolation formula.

D1. CUT-SKY FISHER MATRIX: ANALYTIC EVALUATION

We equate the Fisher matrix of the power spectrum to the curvature of the likelihood function, equation (A16), and then develop an approximate form for the covariance matrix, $\boldsymbol{C} = \boldsymbol{S} + \boldsymbol{N}$, where $\boldsymbol{S}$ is the signal matrix and $\boldsymbol{N}$ is the noise matrix.

We split the noise matrix into two pieces: a weight term and a mask term. In the limit that the pixel noise is diagonal, the weight term has the form

$$\left(N_w^{-1}\right)_{ij} = \frac{n_i}{\sigma_0^2}\delta_{ij} \equiv w_i\delta_{ij} \ , \tag{D1}$$

where i and j are pixel indices, n_i is the number of observations of pixel i, σ_0 is the rms noise of a single observation, and, by definition, w_i is the weight of pixel i. The mask, M_i, is defined so that M_i equals 0 in pixels that are not used as a result of foreground contamination and equals 1 otherwise. The noise matrix can then be written as the product of the two terms

$$N^{-1} = N_w^{-1}M = MN_w^{-1} = \frac{\hat{n}_i}{\sigma_0^2}\delta_{ij} = \hat{w}_i\delta_{ij} \ , \tag{D2}$$

where $\hat{n}_i = n_i$ and $\hat{w}_i = w_i$ in the unmasked pixels and are set to 0 otherwise. We thus define the covariance matrix over the full sky, which allows us to exploit the orthogonality properties of the spherical harmonics. Note that $M^2 = M$.

The inverse of the full covariance matrix can now be written as

$$C^{-1} = (S+N)^{-1} = N^{-1}\left(SN^{-1}+I\right)^{-1} \tag{D3}$$

$$= M^{-1}\left(SM^{-1}+N_w\right)^{-1} \ . \tag{D4}$$

The covariance matrix has two limits. In the noise-dominated limit, $SN^{-1} \ll I$,

$$C^{-1} \to N^{-1} \ . \tag{D5}$$

In the signal-dominated limit, $SM^{-1} \gg N_w$, only the mask alters the covariance matrix, so

$$C^{-1} \to MS^{-1}M \ , \tag{D6}$$

where we have set the inverse of the mask to zero where there are no data, i.e., $M^{-1} \equiv M$.

D1.1. *Cut-Sky Fisher Matrix in the Noise-dominated Limit*

In the noise-dominated limit, the Fisher matrix, equation (A16), is approximately

$$F_{\ell\ell'}^N = \tfrac{1}{2}\mathrm{tr}\left(C^{-1}P_\ell C^{-1}P_{\ell'}\right) \to \tfrac{1}{2}\mathrm{tr}\left(N^{-1}P_\ell N^{-1}P_{\ell'}\right) \ . \tag{D7}$$

Using

$$(P_\ell)_{ij} \equiv \frac{\partial C}{\partial C_\ell} = \frac{(2\ell+1)}{4\pi}w_\ell P_\ell(\cos\theta_{ij}) = w_\ell\sum_m Y_{\ell m}^*(p_i)\,Y_{\ell m}(p_j) \ , \tag{D8}$$

where θ_{ij} is the angle between pixels i and j, $Y_{\ell m}(p_i)$ is a spherical harmonic evaluated in the direction of pixel i, and we have employed the addition theorem for spherical harmonics in the last equality. Then the Fisher matrix takes the form

$$F_{\ell\ell'}^N = \tfrac{1}{2}w_\ell w_{\ell'}\sum_{mm'}\sum_{ij}\hat{w}_i\,Y_{\ell m}^*(p_i)\,Y_{\ell m}(p_j)\,\hat{w}_j\,Y_{\ell'm'}^*(p_j)\,Y_{\ell'm'}(p_i) \ . \tag{D9}$$

Now expand the weight array as

$$\hat{w}_i = \sum_{\ell m}\hat{w}_{\ell m}Y_{\ell m}(p_i) \tag{D10}$$

and substitute this into equation (D9) to yield

$$F_{\ell\ell'}^N = \tfrac{1}{2}w_\ell w_{\ell'}\sum_{mm'}\sum_{\ell''m''}\sum_{\ell'''m'''}\hat{w}_{\ell''m''}\hat{w}_{\ell'''m'''}\sum_{ij}Y_{\ell''m''}(p_i)\,Y_{\ell m}^*(p_i)\,Y_{\ell m}(p_j)\,Y_{\ell'''m'''}(p_j)\,Y_{\ell'm'}^*(p_j)\,Y_{\ell'm'}(p_i) \ . \tag{D11}$$

Since $\hat{w}_i$ was defined over the full sky, the sum over pixels may be expressed in terms of products of Wigner 3-j symbols. As shown in Appendix E, we may use an orthogonality property of the 3-j symbols to reduce the expression for the Fisher matrix to

$$F_{\ell\ell'}^N = \frac{1}{2}w_\ell w_{\ell'}\frac{(2\ell+1)(2\ell'+1)}{4\pi\Omega_p^2}\sum_{\ell''}\hat{W}_{\ell''}\begin{pmatrix}\ell & \ell' & \ell'' \\ 0 & 0 & 0\end{pmatrix}^2 \ , \tag{D12}$$

where $\hat{W}_\ell \equiv \sum_m|\hat{w}_{\ell m}|^2$ and Ω_p is the solid angle per pixel.

D1.2. *Cut-Sky Fisher Matrix in the Signal-dominated Limit*

In the signal-dominated limit, the Fisher matrix, equation (A16), is approximately

$$F_{\ell\ell'}^S = \tfrac{1}{2}\mathrm{tr}\left(\boldsymbol{C}^{-1}\boldsymbol{P}_\ell\boldsymbol{C}^{-1}\boldsymbol{P}_{\ell'}\right) \to \tfrac{1}{2}\mathrm{tr}\left(\hat{\boldsymbol{S}}^{-1}\boldsymbol{P}_\ell\hat{\boldsymbol{S}}^{-1}\boldsymbol{P}_{\ell'}\right), \tag{D13}$$

where $\hat{\boldsymbol{S}}^{-1} \equiv M_i S_{ij}^{-1} M_j$. Using equation (D8), we can write $F_{\ell\ell'}^S$ in the form

$$F_{\ell\ell'}^S = \tfrac{1}{2}w_\ell w_{\ell'} \sum_{mm'}\sum_{ijkp} M_i M_j S_{ij}^{-1} Y_{\ell m}^*(p_j)\, Y_{\ell m}(p_k)\, M_k M_p S_{kp}^{-1} Y_{\ell'm'}^*(p_p)\, Y_{\ell'm'}(p_i). \tag{D14}$$

Substituting a normalized expression for $\boldsymbol{S}^{-1}$ in the pixel basis,

$$S_{ij}^{-1} = \Omega_p^2 \sum_{\ell m} \frac{1}{C_\ell w_\ell} Y_{\ell m}^*(p_i)\, Y_{\ell m}(p_j), \tag{D15}$$

into equation (D14), one obtains a term

$$\Omega_p^2 \sum_{\ell''} \frac{1}{C_{\ell''}w_{\ell''}} \sum_{m''}\sum_i M_i Y_{\ell'm'}(p_i) Y_{\ell''m''}^*(p_i) \sum_j M_j Y_{\ell''m''}(p_j) Y_{\ell m}^*(p_j). \tag{D16}$$

Since our mask cuts only $\sim$15% of the sky, the coupling sum, $\sum_i M_i Y_{\ell'm'}(p_i) Y_{\ell''m''}^*(p_i)$, peaks very sharply at $|\ell' - \ell''| \ll \ell'$. Therefore, one may approximate equation (D16) with

$$\frac{\Omega_p^2}{(C_\ell w_\ell C_{\ell'} w_{\ell'})^{1/2}} \sum_{\ell''m''}\sum_i M_i Y_{\ell'm'}(p_i) Y_{\ell''m''}^*(p_i) \sum_j M_j Y_{\ell''m''}(p_j) Y_{\ell m}^*(p_j) = \frac{\Omega_p}{(C_\ell w_\ell C_{\ell'} w_{\ell'})^{1/2}} \sum_i M_i Y_{\ell'm'}(p_i) Y_{\ell m}^*(p_i), \tag{D17}$$

where, in the last equality, we have used the completeness relation for the spherical harmonics

$$\sum_{\ell m} Y_{\ell m}^*(p_i)\, Y_{\ell m}(p_j) = \Omega_p^{-1}\delta_{ij}. \tag{D18}$$

With this approximation equation (D14) reduces to

$$F_{\ell\ell'}^S \simeq \frac{1}{2}\frac{\Omega_p^2}{C_\ell C_{\ell'}} \sum_{mm'}\sum_{ij} M_i Y_{\ell m}^*(p_i)\, Y_{\ell m}(p_j)\, M_j Y_{\ell'm'}^*(p_j)\, Y_{\ell'm'}(p_i). \tag{D19}$$

Following the same steps outlined above for the derivation of equation (D12) yields

$$F_{\ell\ell'}^S = \frac{1}{2}\frac{1}{C_\ell C_{\ell'}}\frac{(2\ell+1)(2\ell'+1)}{4\pi} \sum_{\ell''} M_{\ell''} \begin{pmatrix} \ell & \ell' & \ell'' \\ 0 & 0 & 0 \end{pmatrix}^2, \tag{D20}$$

where $M_\ell \equiv \sum_m |m_{\ell m}|^2$ and $m_{\ell m}$ is the spherical harmonic transform of the mask. Note that $M_0 = 4\pi f_{\mathrm{sky}}^2$, where f_{sky} is the fraction of the sky that survives the mask.

D2. INTERPOLATING THE CUT-SKY FISHER MATRIX

We can combine the two limiting cases of the Fisher matrix, obtained in the previous section, into a single expression

$$F_{\ell\ell'}^{\mathrm{mask}} = \frac{\sqrt{(2\ell+1)(2\ell'+1)}f_{\mathrm{sky}}^2}{2(C_\ell+N_\ell)(C_{\ell'}+N_{\ell'})} \tilde{F}_{\ell\ell'}, \tag{D21}$$

where $N_\ell \equiv \Omega_p \sigma_0^2/(\bar{n}_{\mathrm{obs}} w_\ell)$ is the deconvolved noise power spectrum. Using equations (D12) and (D20) derived in the previous section, $\tilde{F}_{\ell\ell'}$ can be expressed in the low-ℓ (signal-dominated) and high-ℓ (noise-dominated) limits as

$$\tilde{F}_{\ell\ell'}^S = \frac{\sqrt{(2\ell+1)(2\ell'+1)}}{4\pi f_{\mathrm{sky}}^2} \sum_{\ell''} M_{\ell''} \begin{pmatrix} \ell & \ell' & \ell'' \\ 0 & 0 & 0 \end{pmatrix}^2 \tag{D22}$$

and

$$\tilde{F}_{\ell\ell'}^N = \frac{\sqrt{(2\ell+1)(2\ell'+1)}}{4\pi f_{\mathrm{sky}}^2 \bar{n}_{\mathrm{obs}}^2/\sigma_0^4} \sum_{\ell''} \hat{W}_{\ell''} \begin{pmatrix} \ell & \ell' & \ell'' \\ 0 & 0 & 0 \end{pmatrix}^2. \tag{D23}$$

By construction, these matrices are normalized to 1 on the diagonal, $\tilde{F}_{\ell\ell}^S = \tilde{F}_{\ell\ell}^N = 1$, since $M_0 = 4\pi f_{\mathrm{sky}}^2$ and $\hat{W}_0 = 4\pi f_{\mathrm{sky}}^2 \bar{n}_{\mathrm{obs}}^2/\sigma_0^4$.
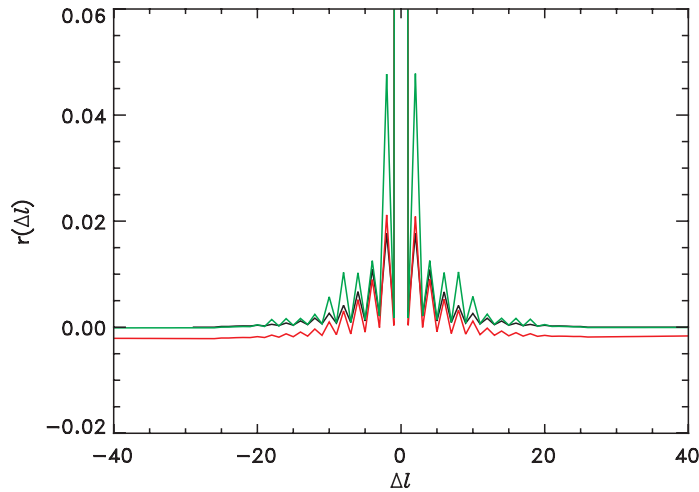
Fig. 14.—Slices of the Fisher (or curvature) matrix normalized as $F_{\ell\ell'}/\sqrt{F_{\ell\ell}F_{\ell'\ell'}}$, plotted vs. $\Delta\ell = \ell - \ell'$, from the Fisher matrix for the combined spectrum. Black is $\ell = 30$, red is $\ell = 300$, and green is $\ell = 600$.

We have computed $F_{\ell\ell'}^{\text{mask}}$ directly, for selected ℓ-values, by evaluating $Y_{\ell m}\mathbf{C}^{-1}Y_{\ell'm'}$ using the preconditioner code described by Oh et al. (1999). We find that deviations between the numerical and analytical results are consistent with numerical noise in the Fisher matrix estimate. Figure 14 shows the estimated Fisher matrix for the Kp2 cut in the noise- and signal-dominated limits.

We interpolate between the two regimes with the following expression:

$$F_{\ell\ell'}^{\text{mask}} = \frac{\left(C_\ell\sqrt{F_{\ell\ell'}^S} + N_\ell\sqrt{F_{\ell\ell'}^N}\right)\left(C_{\ell'}\sqrt{F_{\ell\ell'}^S} + N_{\ell'}\sqrt{F_{\ell\ell'}^N}\right)}{(C_\ell + N_\ell)(C_{\ell'} + N_{\ell'})} \ . \tag{D24}$$

Note that the form of the Fisher matrix primarily depends on $\Delta\ell$ but is weakly dependent on ℓ; there is more coupling in the noise-dominated limit.

D3. POWER SPECTRUM COVARIANCE: MONTE CARLO EVALUATION

As discussed in § A1.2, we compute C_ℓ using three different pixel weightings. The uniform weighting and the inverse noise weighting are optimal in the signal-dominated regime and the noise-dominated regime, respectively. In between these limits the transitional weighting performs better. In order to determine which ranges in ℓ correspond to which regimes and to calibrate our *Ansatz* for the covariance matrix, equation (D21), we proceed as follows. Using 100,000 Monte Carlo simulations of signal plus noise (with *WMAP* noise levels and symmetrized beams), we compute the diagonal elements of the covariance matrix for the three different weighting schemes, evaluated with the Kp2 sky cut. We denote these estimates D_ℓ^{sim}.

We find that the uniform weighting produces the smallest D_ℓ^{sim} below $\ell = 200$. Inverse noise weighting is the best above $\ell = 450$, and the transitional weighting produces the lowest variance in between. In each of these regimes we use the resulting D_ℓ^{sim} to calibrate our *Ansatz* for the diagonal elements of the covariance matrix

$$D_\ell = \frac{2}{2\ell + 1}\frac{1}{f_{\text{sky}}^{\text{eff}}}\left(C_\ell + N_\ell^{\text{eff}}\right)^2 , \tag{D25}$$

as illustrated in Verde et al. (2003). These calibrations produce a smooth correction to equation (D24) of at most 6%. No correction at all is required in the signal-dominated regime.

APPENDIX E

SOME USEFUL PROPERTIES OF SPHERICAL HARMONICS

The derivation of the form of the Fisher matrix in the signal- and noise-dominated limits led to expressions that included a term

$$\sum_{mm'}\sum_{ij} Y_{\ell''m''}(p_i)\, Y_{\ell m}^*(p_i)\, Y_{\ell m}(p_j)\, Y_{\ell'''m'''}(p_j)\, Y_{\ell'm'}^*(p_j)\, Y_{\ell'm'}(p_i) \ , \tag{E1}$$

where the sum is effectively a double integral over the full sky. This can be evaluated in terms of the Wigner 3-*j* symbols,

defined as

$$\int d\Omega_{\boldsymbol{n}} \, Y_{\ell m}(\boldsymbol{n}) \, Y^*_{\ell' m'}(\boldsymbol{n}) \, Y_{\ell'' m''}(\boldsymbol{n}) \equiv (-1)^{m'} \sqrt{\frac{(2\ell+1)(2\ell'+1)(2\ell''+1)}{4\pi}} \begin{pmatrix} \ell & \ell' & \ell'' \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \ell & \ell' & \ell'' \\ m & -m' & m'' \end{pmatrix}. \tag{E2}$$

Substituting this into equation (E1) gives

$$\sum_{mm'} (-1)^{m+m'} \frac{(2\ell+1)(2\ell'+1)}{4\pi} \sqrt{(2\ell''+1)(2\ell'''+1)} \begin{pmatrix} \ell & \ell' & \ell'' \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \ell & \ell' & \ell''' \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \ell & \ell' & \ell'' \\ m & -m' & m'' \end{pmatrix} \begin{pmatrix} \ell & \ell' & \ell''' \\ m & -m' & m''' \end{pmatrix}. \tag{E3}$$

We simplify this using the Wigner 3-*j* orthogonality condition

$$\sum_{mm'} (-1)^{m+m'} \begin{pmatrix} \ell & \ell' & \ell'' \\ -m & m' & m'' \end{pmatrix} \begin{pmatrix} \ell & \ell' & \ell''' \\ m & -m' & m''' \end{pmatrix} = \frac{1}{2\ell''+1} \delta_{\ell''\ell'''} \delta_{m''m'''} \delta(\ell, \ell', \ell'') \tag{E4}$$

where $\delta(\ell, \ell', \ell'') = 1$ for $|\ell - \ell'| \le \ell'' \le \ell + \ell'$ and is 0 otherwise. This reduces equation (E1) to

$$\frac{(2\ell+1)(2\ell'+1)}{4\pi\Omega_p^2} \begin{pmatrix} \ell & \ell' & \ell'' \\ 0 & 0 & 0 \end{pmatrix}^2, \tag{E5}$$

where the factor of Ω_p^2 accounts for the fact that equation (E1) is a double sum over pixels, instead of a double integral.

REFERENCES

Barnes, C., et al. 2003, ApJS, 148, 51
Bennett, C. L., et al. 1996, ApJ, 464, L1
———. 2003a, ApJ, 583, 1
———. 2003b, ApJS, 148, 1
———. 2003c, ApJS, 148, 97
Benoit, A., et al. 2003, A&A, submitted
Bond, J. R., Jaffe, A. H., & Knox, L. 1998, Phys. Rev. D, 57, 2117
Górski, K. M., Wandelt, B. D., Hansen, F. K., Hivon, E., & Banday, A. J. 1999, preprint (astro-ph/9905275)
Gupta, S., & Heavens, A. F. 2002, MNRAS, 334, 167
Halverson, N. W., et al. 2002, ApJ, 568, 38
Hauser, M. G., & Peebles, P. J. E. 1973, ApJ, 185, 757
Hinshaw, G., Branday, A. J., Bennett, C. L., Górski, K. M., Kogut, A., Lineweaver, C. H., Smoot, G. F., & Wright, E. L. 1996, ApJ, 464, L25
Hinshaw, G. F., et al. 2003, ApJS, 148, 63
Hivon, E., Górski, K. M., Netterfield, C. B., Crill, B. P., Prunet, S., & Hansen, F. 2002, ApJ, 567, 2
Jarosik, N., et al. 2003a, ApJS, 145, 413
———. 2003b, ApJS, 148, 29
Kogut, A., et al. 2003, ApJS, 148, 161
Komatsu, E., et al. 2003, ApJS, 148, 119
Kuo, C. L., et al. 2002, ApJ, submitted
Lee, A. T., et al. 2001, ApJ, 561, L1
Limon, M., et al. 2003, *Wilkinson Microwave Anisotropy Probe Explanatory Supplement*
Miller, A., et al. 2002, ApJS, 140, 115
Muciaccia, P. F., Natoli, P., & Vittorio, N. 1997, ApJ, 488, L63
Oh, S. P., Spergel, D. N., & Hinshaw, G. 1999, ApJ, 510, 551
Page, L., et al. 2003a, ApJ, 585, 566
———. 2003b, ApJS, 148, 39
———. 2003c, ApJS, 148, 233
Pearson, T. J., et al. 2003, ApJ, 591, 556
Peiris, H., et al. 2003, ApJS, 148, 213
Percival, W. J., et al. 2002, MNRAS, 337, 1068
Ruhl, J. E., et al. 2002, preprint (astro-ph/0212229)
Spergel, D. N., et al. 2003, ApJS, 148, 175
Tegmark, M. 1996, ApJ, 464, L35
Tegmark, M., Taylor, A. N., & Heavens, A. F. 1997, ApJ, 480, 22
Verde, L., et al. 2003, ApJS, 148, 195
Wang, X., Tegmark, M., Jain, B., & Zaldarriaga, M. 2002a, Phys. Rev. D, submitted
Wang, X., Tegmark, M., & Zaldarriaga, M. 2002b, Phys. Rev. D, 65, 123001